

UNIVERSITAT POLITÈCNICA DE VALÈNCIA
ESCOLA TÈCNICA SUPERIOR D'ENGINYERIA
AGRONÒMICA I DEL MEDI NATURAL



**UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA**

Grado en Biotecnología

Origen del microbioma de placas solares

Trabajo Fin de Grado (TFG)

AUTOR: Adriel Latorre Pérez

TUTOR UPV: José Gadea Vacas

COTUTOR ACADÉMICO: Manuel Porcar Miralles

COTUTORA COLABORADORA: Cristina Vilanova Serrador

Curso Académico: 2016/2017

Valencia, Junio de 2017



Datos personales

Nombre y apellidos: Adriel Latorre Pérez

Datos del Trabajo Fin de Grado

Título: Origen del microbioma de placas solares

Titulación: Grado en Biotecnología

Tutor académico: José Gadea Vacas

Cotutor académico: Manuel Porcar Miralles

Cotutora externa colaboradora: Cristina Vilanova Serrador

Localidad y fecha de lectura: Valencia, Junio 2017

Resumen

Durante los últimos años, los avances en las técnicas metagenómicas y bioinformáticas han permitido explorar la biodiversidad microbiana presente en distintos tipos de ambientes. Entre todos ellos, los ambientes extremos y artificiales han despertado un especial interés en la comunidad científica debido a su alto valor ecológico y su enorme potencial como fuente de microorganismos con aplicaciones industriales.

Las placas solares son una de las estructuras hechas por el hombre más extendidas y se caracterizan por combinar estas dos propiedades: son ambientes artificiales sometidos a condiciones fisicoquímicas extremas. En un trabajo anterior, nuestro grupo caracterizó este ecosistema utilizando una aproximación metagenómica en placas solares de la ciudad de Valencia, y determinó que la comunidad microbiana presente tenía más relación con ambientes naturales extremos, como desiertos, que con otro tipo de ambientes relacionados con los humanos o las áreas urbanas.

El objetivo de este trabajo es discernir el origen de esta comunidad microbiana. Para ello, se partió de la hipótesis de que el microbioma de placas solares dependía en gran medida de las migraciones microbianas desde los desiertos a través de las tormentas de polvo. Se analizaron diferentes perfiles taxonómicos extraídos de la bibliografía y correspondientes a ecosistemas susceptibles de ser el germen de la biocenosis presente en los paneles fotovoltaicos (aire sin polvo, aire con polvo desértico y suelo desértico), y se compararon con muestras de placas solares de Valencia y Berkeley secuenciadas por nuestro grupo.

Los resultados demostraron que la comunidad microbiana presente en los paneles solares no es fruto de una simple deposición de microorganismos aéreos, sino que se trata de una comunidad compleja en la que se seleccionan aquellos organismos mejor adaptados. Se encontró que tanto el aire como los desiertos ejercían una importante influencia en la formación del microbioma de placas solares. Además, las comunidades microbianas que habitan estas estructuras presentaban un elevado grado de similitud entre las dos localizaciones estudiadas, y un reducido núcleo de esta población se encontraba significativamente infrarrepresentado en los otros ecosistemas analizados. Este *core* se compone de bacterias resistentes a la desecación y la radiación que se caracterizan por su capacidad de sintetizar carotenoides.

Palabras clave

Placas solares, microbioma, metagenómica comparativa, ecología microbiana

Abstract

During the last years, new advances in metagenomic and bioinformatic techniques have allowed the exploration of the biodiversity present in different types of environments. Among them, extreme and artificial environments have been found of special interest by the scientific community due to their high ecological value and their potential as a source of microorganisms with industrial applications.

Solar panels are one of most widespread man-made structures that combine these two characteristics: they are artificial environments that are subjected to extreme physicochemical conditions. In a previous work, our group characterized this ecosystem by means of a metagenomic approach performed on solar panels in the city of Valencia. The microbial community present in the surface of these structures proved more similar to that associated to other extreme environments, such as deserts, than to other human-associated or urban-associated areas.

The objective of this work is to explore the origin of this microbial community, under the hypothesis that the microbiome present on solar panels is largely dependent on microbial migrations coming from deserts through dust storms. Different taxonomic profiles that corresponded to the ecosystems susceptible to be the germ of the solar panels biocenosis (air without dust, air with dust, and desert soil) were retrieved from the literature and compared with those of solar panels sampled in Valencia and Berkeley that had been previously sequenced by our group.

Our results showed that the microbial community present in the surface of solar panels was not a simple deposition of air microorganisms, but a complex community in which resistant microorganisms are selected. Air and deserts proved to have an important influence on the formation of the solar panels microbiome. Furthermore, the microbial communities inhabiting solar panels were found to be strikingly similar in the two locations analyzed, and a reduced core of those communities was found to be significantly underrepresented in the other ecosystems studied. This core is composed of desiccation- and radiation-resistant bacteria, characterized by the ability to synthesize carotenoids.

Keywords

Solar panels, microbiome, comparative metagenomics, microbial ecology

Tipo de Licencia:

Licencia Creative Commons: "Reconocimiento – No Comercial – Sin Obra Derivada (by-nc-nd)"

Agradecimientos

A día de hoy, tengo claro que el trabajo y el esfuerzo personal no siempre es suficiente si no se cuenta con el apoyo de ciertas personas. Por tanto, este proyecto no puede entenderse al completo sin tener en cuenta a toda esa gente que me ha acompañado en su desarrollo:

En el ámbito académico, agradecer a Pepe Gadea por sus consejos y por ser el inicio de este trabajo. A Manel por darme la oportunidad de desarrollar el mismo, ofrecerme su ayuda en todo lo necesario y depositar su confianza en mí desde el primer momento. Quisiera dar las gracias especialmente a Cristina, ahora ya la Dra. Vilanova, por estar siempre al pie del cañón con su ayuda, por su especial ambientación del laboratorio y por enseñarme tanto en tan poco tiempo. Sin ti seguramente nada de esto habría sido posible.

También quisiera hacer especial mención a todas las personas que forman parte del equipo del I2SysBio y que han contribuido a que ir al laboratorio se convirtiera en una motivación y no una obligación. Kristie, Esther, Àngela, Louise, Alberto, Carmen, Christian, José Manuel y Juli: gracias.

En lo personal, quisiera dar las gracias a mi padre y a mi hermano por todo el apoyo y el cariño. A Irene, por haber recorrido junto a mí este largo y duro camino que ha sido la carrera y haberme permitido crecer a tu lado. Y en especial gracias a ti, mamá, por enseñarme a ser valiente.

ÍNDICE

1. INTRODUCCIÓN	1
1.1 La era del microbioma.....	1
1.2 Técnicas para el análisis masivo de comunidades microbianas.....	2
1.2.1 <i>Metagenómica y next generation sequencing (NGS)</i>	2
1.2.2 <i>Estrategias: secuenciación dirigida y metagenómica</i>	3
1.3 El microbioma ambiental.....	4
1.4 El caso particular de placas solares.....	5
1.4.1 <i>Las placas solares como biotopos</i>	5
1.4.2 <i>Estudios previos</i>	6
1.4.3 <i>Origen del microbioma de placas solares: hipótesis</i>	7
2. OBJETIVOS	10
3. MATERIALES Y MÉTODOS	11
3.1 Diseño experimental.....	11
3.2 Selección de muestras.....	11
3.2.1 <i>Secuencias propias</i>	11
3.2.2 <i>Estudio comparativo de placas solares: Valencia</i>	12
3.2.3 <i>Estudio comparativo de placas solares: Berkeley</i>	12
3.2.4 <i>Estudio comparativo global</i>	14
3.3 Análisis bioinformático.....	14
3.3.1 <i>Pre-procesamiento de secuencias</i>	14
3.3.2 <i>Asignación taxonómica</i>	14
3.3.3 <i>Comparación de los perfiles taxonómicos</i>	15
3.3.4 <i>Cuantificación de la fuente de origen del microbioma</i>	16
3.3.5 <i>Inferencia funcional</i>	16
3.3.6 <i>Adaptación de datos de WGS</i>	17
4. RESULTADOS Y DISCUSIÓN	18
4.1 Estudio comparativo de placas solares: Valencia.....	18
4.1.1 <i>Comparación de los perfiles taxonómicos</i>	18
4.1.2 <i>Cuantificación de la fuente de origen del microbioma</i>	20
4.1.3 <i>Comparación de los perfiles funcionales inferidos</i>	21
4.2 Estudio comparativo de placas solares: Berkeley.....	22
4.2.1 <i>Comparación de los perfiles taxonómicos</i>	22
4.2.2 <i>Cuantificación de la fuente de origen del microbioma</i>	24
4.3 Estudio comparativo global.....	25
4.3.1 <i>La biocenosis de placas solares</i>	26
4.4 Discusión general.....	27
4.4.1 <i>Limitaciones del análisis</i>	27
4.4.2 <i>El origen del microbioma de placas solares</i>	27
4.4.3 <i>Plan de trabajo futuro</i>	28
5. CONCLUSIONES	29
6. REFERENCIAS BIBLIOGRÁFICAS	30
7. ANEXO	36

ÍNDICE DE FIGURAS

- Figura 1.1.** Coste de la secuenciación y su impacto en el campo de la metagenómica. Eje y (negro): coste por Mb secuenciada en escala logarítmica. Eje y (azul): el número total de publicaciones en PubMed encontradas bajo el término de búsqueda “Metagenomics”. *Adaptado de Temperton y Giovannoni, 2012*..... 2
- Figura 1.2.** Esquema general de las dos principales estrategias empleadas en los estudios metagenómicos: a) Secuenciación de genomas completos b) Secuenciación basada en genes marcadores. *Adaptado de Yoo et al., 2017*..... 4
- Figura 1.3.** Análisis de Coordenadas Principales (PCoA) a partir del perfil taxonómico extraído de un conjunto de metagenomas de diferentes ecosistemas. *Adaptado de Dorado-Morales et al., 2016*..... 6
- Figura 1.4.** Mapa de la distribución global media de las medidas diarias del Índice de Aerosoles capturadas por el satélite TOMS (NASA). Los cuadrados indican las zonas que se proponen como las principales fuentes mundiales de polvo: 1. Sáhara 2. Arabia 3. Asia 3b. Asia del Este 4. América del Norte 5. América del Sur 6. Sudáfrica 7. Australia. *Adaptado de Varga, 2012*..... 8
- Figura 1.5.** Fuentes principales del polvo desértico y su distribución atmosférica global. Las líneas naranjas muestran las rutas del polvo africano, las amarillas del proveniente de los desiertos de Asia Central, las marrones muestran las rutas del polvo proveniente de otros desiertos, y las líneas negras punteadas representan los movimientos del aire. *Adaptado de Griffin, 2007*..... 8
- Figura 3.1.** Desiertos de EE. UU. Solo se incluye la parte estadounidense de los desiertos de Sonora y Chihuahua. La flecha roja indica la posición de Berkeley (California). *Adaptado de <http://www.desertusa.com/north-american-deserts.html>*..... 13
- Figura 3.2.** Flujo de trabajo utilizado para la transformación de las secuencias y la asignación taxonómica. Las flechas rojas indican cambios en la configuración por defecto establecida para cada paso, mientras que las azules hacen referencia a los scripts informáticos usados en cada caso (los que no se indican como propios pertenecen al software bioinformático QIIME). Sec. = secuencia; Rep. = representativa. OTU = unidad taxonómica operativa..... 15
- Figura 4.1.** Análisis de Coordenadas Principales (PC1 (25,3%) vs PC2 (17,8%)) realizado con los perfiles taxonómicos de las muestras incluidas en el análisis comparativo de placas solares de Valencia (máximo nivel de asignación taxonómica: género). Las OTUs que no representaban a más de 25 secuencias fueron eliminadas. Se observan tres grandes grupos constituidos por: 1. Muestras del estudio de Israel (Polvo y Aire Control); 2. Muestras del estudio de Cerdeña (Polvo y Aire Control); 3. El resto de muestras del análisis (aire sin polvo, polvo desértico Valencia, suelo desértico y placas solares)..... 18
- Figura 4.2.** Comparación de los perfiles taxonómicos una vez eliminadas las muestras de polvo/aire control de Israel y Cerdeña (máximo nivel de asignación taxonómica: género). Las OTUs que no representaban a más de 25 secuencias fueron eliminadas. A) Análisis de Coordenadas Principales (PC1 (36,1%) vs PC2 (23,6%)). B) Neighbor-Net. Se observan cuatro grupos constituidos por: 1. Suelo del Sáhara (Chad); 2. Placas solares de Valencia; 3. Aire urbano (Italia); 4. Polvo del Sáhara recogido en Valencia..... 19

- Figura 4.3.** Comparación de los perfiles taxonómicos (máximo nivel de asignación taxonómica: género). A) Análisis de Coordenadas Principales (PC1 (44,7%) vs PC2 (30,0%)). B) Neighbor-Net. Las muestras de Israel y Cerdeña se excluyeron. Solo se tienen en cuenta tres muestras de aire urbano. Se observan cuatro grupos constituidos por: 1. Suelo del Sáhara (Chad); 2. Placas solares de Valencia; 3. Aire urbano (Italia); 4. Polvo del Sáhara recogido en Valencia..... 20
- Figura 4.4.** Contribución de las distintas fuentes en la formación del microbioma de placas solares de Valencia según el software SOURCETRACKER. Solo se muestran los porcentajes de las influencias mayoritarias..... 20
- Figura 4.5.** *Heatmap* y dendrograma basado en el análisis comparativo de los perfiles funcionales inferidos de los distintos tipos de ecosistemas estudiados. Se resaltan las funciones COG en las que más diferencias se encontraron entre los distintos tipos de muestra..... 21
- Figura 4.6.** Análisis de Coordenadas Principales (PC1 (41,6%) vs PC2 (18,6%)) preliminar realizado a partir de perfiles taxonómicos (máximo nivel de asignación taxonómica: género). Se observan cuatro grupos constituidos por: 1. Desiertos de EE.UU.; 2. Aire Control (Colorado); 3. Polvo desértico (Corea); 4. Aire Control (Corea)..... 23
- Figura 4.7.** Comparación de los perfiles taxonómicos incluyendo las muestras de placas solares de Berkeley (máximo nivel de asignación taxonómica: género). **A)** Análisis de Coordenadas Principales (PC1 (33,5%) vs PC2 (21,1%)). **B)** Neighbor-Net. **C)** Análisis de Coordenadas Principales (PC1 (33,5%) vs PC3 (14,6%)). Se observan 3 grupos que se corresponden con los distintos biomas estudiados: **1.** Suelo desiertos (EE.UU). **2.** Placas solares (Berkeley); **3.** Aire (**3.1** Aire control (EE.UU.); **3.2** Aire control (Corea); **3.3** Polvo desértico (Corea))..... 24
- Figura 4.8.** Contribución de las distintas fuentes en la formación del microbioma de placas solares de Berkeley según el software SOURCETRACKER. Solo se muestran los porcentajes de las influencias mayoritarias..... 25
- Figura 4.9.** Análisis de Coordenadas Principales (PC1 (29,0%) vs PC2 (20,1%)) de todas las muestras incluidas en el estudio comparativo global (máximo nivel de asignación taxonómica: género). Se observan 3 grupos que se corresponden con los distintos biomas estudiados: 1. Suelo desiertos (EE.UU y Sáhara). 2. Placas solares (Valencia y Berkeley); 3. Aire (Aire control (EE.UU.); Aire urbano control (Italia)) y polvo desértico (Valencia)..... 26
- Figura 4.10.** Comparación de la abundancia relativa media de los cuatro géneros más sobrerrepresentados en placas solares frente a los otros ecosistemas (polvo desértico, suelo desértico y aire sin polvo)..... 27
- Figura 7.1.** Perfil taxonómico (porcentaje de secuencias asignables a bacterias identificables en bases de datos) procedente de muestras de Cerdeña en días libres de polvo (2, 4) y días de tormenta de polvo (1,3). *Extraído de Roselli et al., 2015*..... 38
- Figura 7.2.** Análisis de Coordenadas Principales (PC1 (33,6%) vs PC2 (23,1%)) realizado a partir de perfiles taxonómicos (máximo nivel de asignación taxonómica: género). Se incorporó al análisis una muestra de placas solares de Valencia secuenciada mediante una estrategia metagenómica. Se observan cuatro grupos constituidos por: 1. Suelo del Sáhara (Chad); 2. Placas solares de Valencia; 3. Aire urbano; 4. Polvo del Sáhara recogido en Valencia..... 38

ÍNDICE DE TABLAS

Tabla 3.1. Set de datos metagenómicos de las muestras secuenciadas en estudios anteriores del grupo. P.S. = Placas solares; WGS = <i>whole genome shotgun</i> ; 16S = Secuenciación del gen 16S rRNA.....	11
Tabla 3.2. Set de datos metagenómicos de las muestras seleccionadas en el estudio comparativo de Valencia. Se incluyen las secuencias propias que se usaron en el análisis. Todos los estudios incluidos seguían una estrategia experimental basada en la secuenciación del gen 16S rRNA.....	12
Tabla 3.3. Set de datos metagenómicos de las muestras seleccionadas en el estudio comparativo de Berkeley. Se incluyen las secuencias propias que se usaron en el análisis. WGS = <i>whole genome sequencing</i> ; 16S = Secuenciación del gen 16S rRNA.....	13
Tabla 7.1. Set de datos metagenómicos de las muestras seleccionadas en la búsqueda bibliográfica preliminar. WGS = <i>whole genome shotgun</i> ; 16S = Secuenciación del gen 16S rRNA.....	36
Tabla 7.2. Muestra de los 15 taxones más abundantes en placas solares de Valencia y su abundancia relativa en cada muestra. Vlc = Valencia; Bk = Berkeley.....	37
Tabla 7.3. Muestra de los 15 taxones más abundantes en placas solares de Berkeley y su abundancia relativa en cada muestra.....	37

1. INTRODUCCIÓN

1.1 La era del microbioma

La biodiversidad, entendida como la variedad estructural y funcional de los seres vivos a distintos niveles, es un hecho inherente a la vida en la Tierra. Los millones de años de evolución desde la aparición de la primera molécula orgánica en el planeta han permitido la aparición de innumerables organismos adaptados a prácticamente cualquier hábitat existente.

Asimismo, el afán por la exploración de dicha diversidad está presente en la idiosincrasia del ser humano desde la formación de nuestra especie. No obstante, las motivaciones han cambiado. Mientras que los primeros humanos pretendían descubrir nuevas formas de vida que usar como recurso para su supervivencia, los científicos actuales basan sus investigaciones en obtener conocimiento molecular, celular o supracelular de los procesos que nos rodean. Sea como fuere, existe un objetivo común: usar la naturaleza como fuente de aplicaciones prácticas (Vilanova, 2016).

En este sentido, los microorganismos juegan un importante papel, ya que se trata de los seres vivos más extendidos y con mayor versatilidad metabólica y funcional. Desde que Antoni van Leewenhoek describió la primera bacteria en 1683 (Brock, 1961), la microbiología ha avanzado a pasos agigantados, sobre todo a partir de la introducción de las técnicas de cultivo por Robert Koch (Carter, 1987). El descubrimiento de las bacterias patógenas causantes de epidemias, el aislamiento de microorganismos productores de antibióticos o el desarrollo de las vacunas son solo algunas muestras de ello.

Sin embargo, en lo referente a la exploración de la biodiversidad microscópica, la verdadera revolución surge del desarrollo de la metagenómica y la bioinformática.

La **metagenómica** es una disciplina englobada dentro de la genómica y que tiene como objeto estudiar el conjunto de genomas o genes concretos (el gen ribosomal 16S, por ejemplo) de una muestra compleja compuesta por distintas especies microbianas. De este modo, se puede inferir información funcional y taxonómica sin necesidad de realizar cultivos previos (Handelsman *et al.*, 2002).

La **bioinformática**, por su parte, desempeña un papel fundamental en los estudios metagenómicos, no solo debido a la gran cantidad de datos con los que se trabaja, sino por la gran complejidad existente en el procesamiento y análisis de la información (Wooley y Ye, 2010).

Solo a través de estas disciplinas ha sido posible la construcción de un nuevo concepto: **el concepto de microbioma**. Este término se emplea para designar genéricamente al “conjunto de microorganismos autóctonos asociados a un huésped” (Salvucci, 2016) o cualquier otro tipo de ambiente. Sin embargo, tiene unas connotaciones mucho mayores, que implican intentar llegar a obtener una imagen fidedigna de la enorme diversidad microbiana presente en cualquier ecosistema.

En un mundo donde se estima que el 99% de los microorganismos no son cultivables (Wooley y Ye, 2010), la exploración de los distintos microbiomas abre una nueva era en el descubrimiento y caracterización de especies con interés humano y biotecnológico, que permitirá responder grandes preguntas: ¿cuál es la diversidad microbiana real de los diferentes ambientes?, ¿qué interacciones se producen entre las distintas especies

microbianas y/o entre estas y el organismo huésped?, ¿cómo evolucionan las comunidades microbianas?

1.2 Técnicas para el análisis masivo de comunidades microbianas

1.2.1 Metagenómica y next generation sequencing (NGS)

La secuenciación del ADN es uno de los pilares fundamentales de la metagenómica. En este sentido, las técnicas de *next generation sequencing* (NGS) han supuesto una gran revolución. Pese al escepticismo inicial, el bajo precio de secuenciación y la capacidad de generar una gran cantidad de información con menos esfuerzo han establecido la tecnología NGS como una herramienta indispensable dentro de este campo.

Existen ciertos inconvenientes en el uso de este tipo de estrategias, tales como la peor calidad y menor longitud de las secuencias obtenidas frente al método Sanger, lo que podría a su vez repercutir en una peor asignación taxonómica debido a los errores de lectura, así como en un ensamblaje y anotación más dificultosa debido al pequeño tamaño de secuencia (Temperton y Giovannoni, 2012; Scholz *et al.*, 2012).

No obstante, esto no ha sido un gran impedimento. Tal y como se puede observar en la Figura 1.1, la caída del precio por megabase (Mb) de ADN secuenciado gracias al desarrollo de las NGS se corresponde claramente con el incremento en el número de proyectos relacionados con la metagenómica. Estos se han beneficiado a su vez del gran esfuerzo invertido en la mejora y optimización de las herramientas bioinformáticas (Oulas *et al.*, 2015).

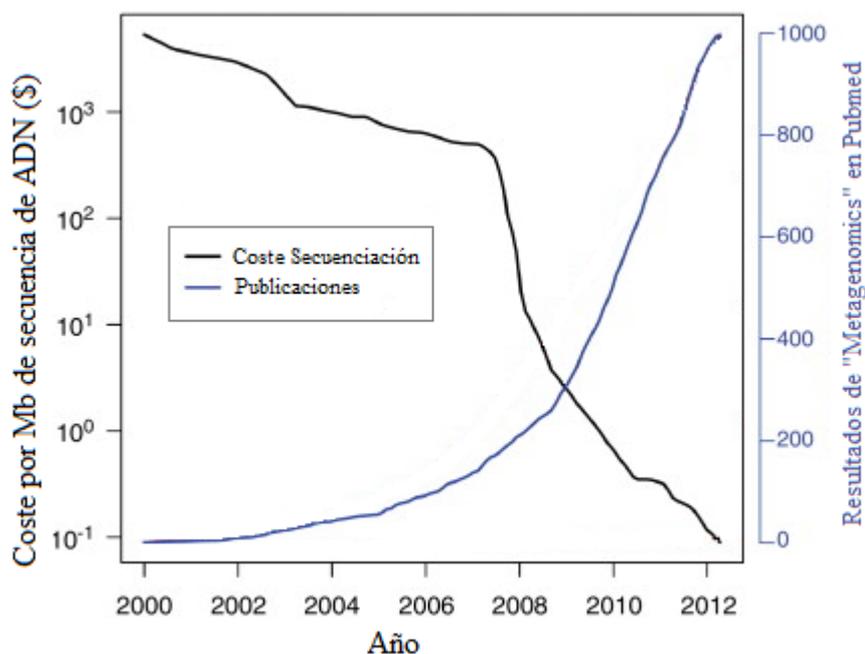


Figura 1.1. Coste de la secuenciación y su impacto en el campo de la metagenómica. Eje y (negro): coste por Mb secuenciada en escala logarítmica. Eje y (azul): el número total de publicaciones en PubMed encontradas bajo el término de búsqueda "Metagenomics". *Adaptado de Temperton y Giovannoni, 2012*

Como consecuencia, han salido a la luz estudios de todo tipo que han permitido entender en mayor profundidad la vida microbiana y su relación con el entorno. Uno de ellos es el Proyecto Microbioma Humano (Human Microbiome Project Consortium, 2012a), el cual ha posibilitado conocer la estructura, función y diversidad de la microbiota normal de individuos

sanos, gracias a lo cual se espera poder identificar de una manera más precisa variaciones responsables de patologías (Human Microbiome Project Consortium, 2012b).

1.2.2 Estrategias: secuenciación dirigida y metagenómica

Pese a la variedad de técnicas independientes de cultivo existentes, las estrategias basadas en secuenciación son las más extendidas y se pueden englobar en dos variantes: secuenciación dirigida a un gen concreto (normalmente un gen marcador de taxonomía que es amplificado por PCR) o la secuenciación de genomas completos (*Whole Genome Sequencing* o WGS), tal y como se puede observar en la Figura 1.2. Esta última aproximación suele recibir el nombre de “metagenómica” (o *metagenomics*) (Bragg & Tyson, 2014).

La **secuenciación basada en genes marcadores** es usada desde mediados de la década de los noventa para determinar la diversidad microbiana presente en muestras ambientales complejas. En concreto, se suele optar por secuenciar las regiones hipervariables del gen que codifica para la subunidad 16S del ARN ribosómico (16S rRNA) (Yoo *et al.*, 2017). El gen candidato en el caso de los microorganismos eucariotas suele ser el 18S rRNA o la región ITS (Kim *et al.*, 2017a).

Las principales ventajas de este método incluyen una mayor rapidez analítica, una mayor facilidad computacional debido al menor número de secuencias que se obtienen, y una mayor eficiencia en la reacción gracias a la posibilidad de combinar muestras usando *barcodes*. En este aspecto, el crecimiento de bases de datos como *SILVA* (Quast *et al.*, 2012) o *Greengenes* (Desantis *et al.*, 2006) permite el estudio de la diversidad y abundancia bacteriana de manera rápida y con una alta resolución taxonómica (Yoo *et al.*, 2017).

Sin embargo, continúan existiendo errores asociados a la necesidad de amplificar la región a estudiar mediante PCR. Algunos de ellos son la formación de quimeras en los amplicones generados, el sesgo de los primers utilizados hacia un determinado grupo taxonómico, o la pérdida de información por errores de secuenciación en los *barcodes* (Yoo *et al.*, 2017).

Por su parte, la **secuenciación de genomas completos** o **metagenómica** se basa en la fragmentación aleatoria del ADN de una muestra ambiental y la secuenciación de estos fragmentos mediante técnicas de NGS. Posteriormente, las secuencias que solapan se van ensamblando en fragmentos mayores y se van rellenando los huecos (*gaps*) hasta obtener un metagenoma completo (Yoo *et al.*, 2017).

A diferencia de la secuenciación de genes concretos, con la metagenómica se obtiene información funcional y metabólica más allá de la información taxonómica, por lo que puede ayudar a comprender el contexto ecológico en el que se desarrollan los microorganismos (Hugenholtz & Tyson, 2008).

Las limitaciones de esta técnica incluyen la variabilidad de resultados con respecto a cambios en el protocolo de extracción de ADN o de secuenciación (Hugenholtz & Tyson, 2008), así como la mayor complejidad computacional y la dificultad de realizar anotaciones precisas debido a la carencia de bases de datos y referencias para comparar los resultados (Thomas *et al.*, 2012). Además, la gran cantidad de datos generados mediante esta aproximación aumentan el tiempo del análisis, sobre todo si se trata de experimentos de metagenómica comparativa donde se evalúan un número elevado de muestras.

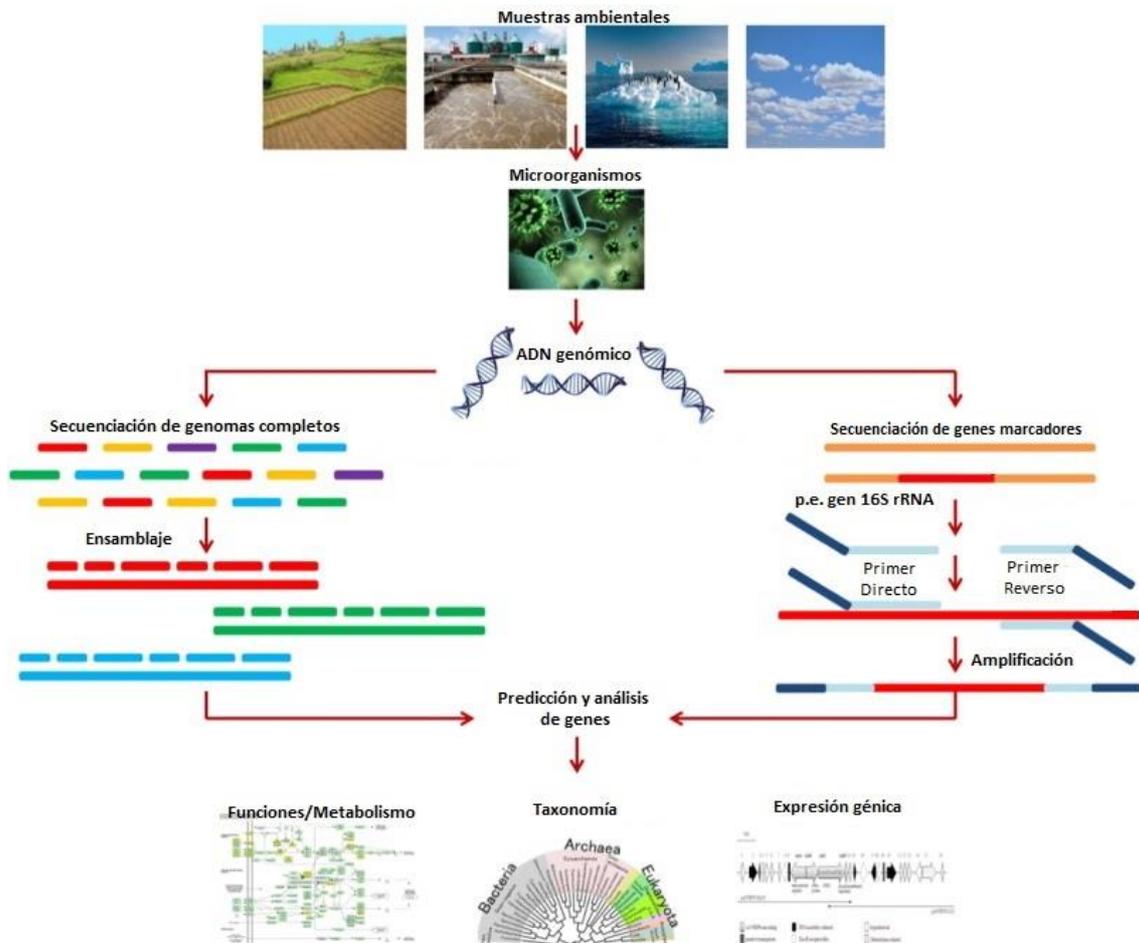


Figura 1.2. Esquema general de las dos principales estrategias empleadas en los estudios metagenómicos: a) Secuenciación de genomas completos b) Secuenciación basada en genes marcadores. *Adaptado de Yoo et al., 2017.*

1.3 El microbioma ambiental

Desde los orígenes de la metagenómica, uno de los principales objetivos ha sido siempre conocer y comprender la ecología microbiana y entender qué adaptaciones adquieren los microorganismos que pueblan los distintos ambientes. De este modo, es posible obtener una idea más precisa del alcance que pueden tener las distintas aplicaciones derivadas.

Al principio, se comenzó explorando hábitats que albergaban una gran biodiversidad, tales como suelos (Daniel, 2005), océanos (Sogin *et al.*, 2006) o el intestino humano (Turnbaugh *et al.*, 2007). De hecho, el primer análisis metagenómico, llevado a cabo por Venter *et al.* (2004), se centró en estudiar el microbioma presente en el mar de los Sargazos, encontrando al menos 1800 especies microbianas y más de 1,2 millones de genes desconocidos hasta el momento, lo que popularizó esta tendencia entre el resto de investigadores.

Con el avance de la disciplina, ha sido posible abordar estudios cada vez más grandes y complejos. Uno de los más interesantes es el proyecto del microbioma terrestre (*The Earth Microbiome Project*), a través del cual se han estudiado 200.000 muestras provenientes de más de 40 biomas diferentes. La intención final es crear un catálogo global de la diversidad de los microorganismos no cultivables que habitan a lo largo de todo el planeta (Gilbert *et al.*, 2014).

No obstante, en los últimos años los mayores esfuerzos se han centrado en la investigación de dos tipos de ambientes con características muy peculiares: los ambientes extremos y los artificiales.

Los primeros tienen un alto valor biotecnológico, ya que disponen de proteínas y enzimas adaptadas a condiciones fisicoquímicas muy estresantes que pueden ser aplicables a la industria (Ferrer *et al.*, 2007). En este sentido, existen diversos estudios que han puesto de manifiesto la ubiquidad de los microorganismos y su capacidad de aclimatarse a cualquier tipo de entorno. Por ejemplo, Konstantinidis *et al.* (2009) estudiaron la comunidad microbiana presente en las profundidades del océano Pacífico, un ecosistema caracterizado por presentar altas presiones y bajas temperaturas, demostrando que la adaptación de los microorganismos a este entorno se debe a sutiles cambios metabólicos y genómicos, en lugar de a grandes alteraciones génicas o estructurales. Otros ambientes extremos como desiertos áridos (Neilson *et al.*, 2012), aguas termales submarinas (López-López *et al.*, 2013), o residuos radioactivos (Williamson *et al.*, 2014) también han demostrado ser colonizables por microorganismos especializados.

En cambio, los ambientes artificiales constituyen un hábitat nuevo a los que los microorganismos han tenido que adaptarse, por lo que *a priori* presentan un alto valor ecológico, así como implicaciones para la salud humana cuando se trata de ambientes artificiales habitados o transitados por personas como viviendas (Barberán *et al.*, 2015) o transportes públicos (Triadó-Margarit *et al.*, 2016). Entre los últimos estudios metagenómicos realizados en ambientes artificiales, destacan las máquinas de café (Vilanova *et al.*, 2015) y otros electrodomésticos (Savage *et al.*, 2016).

1.4 El caso particular de placas solares

Las placas solares combinan dos características que las convierten en un biotopo muy particular: son estructuras artificiales sometidas a condiciones extremas. Su verdadera función es convertir la luz solar en energía o calor, por lo que se engloban dentro de las energías renovables. A día de hoy, estos paneles ocupan alrededor de unos 4000 kilómetros cuadrados de superficie terrestre y se espera que su popularidad siga aumentando hasta convertirse en la principal fuente de energía mundial en 2050 (Dorado-Morales *et al.*, 2016).

1.4.1 Las placas solares como biotopos

Las singulares características fisicoquímicas de los paneles solares incrementan el interés en el estudio de su rol como biotopos. Las superficies de estas estructuras consisten normalmente en un fino y plano vidrio caracterizado por maximizar la exposición a la luz solar y por presentar una capacidad mínima de retención de agua. De este modo, los microorganismos susceptibles a desarrollarse sobre ella deberían ser capaces de soportar condiciones extremas de desecación, calor y radiación, por lo que serían potencialmente aplicables en la industria biotecnológica (Dorado-Morales *et al.*, 2016).

También existe un gran interés en conocer la influencia de estas comunidades microbianas en el rendimiento energético, ya que Shirakawa *et al.* (2015) determinaron que la colonización microbiana afectaba a la eficiencia de los paneles fotovoltaicos, mostrando disminuciones de hasta un 11% en la potencia generada.

1.4.2 Estudios previos

En un estudio previo, nuestro grupo caracterizó el microbioma asociado a las placas solares en la ciudad de Valencia (España) mediante estrategias de cultivo, secuenciación del 16S rRNA y secuenciación metagenómica. La biocenosis presente demostró ser más similar a la que se puede encontrar en ambientes naturales extremos, como desiertos, que a otras comunidades microbianas presentes en cualquier otro ecosistema urbano o asociado a humanos (Figura 1.3).

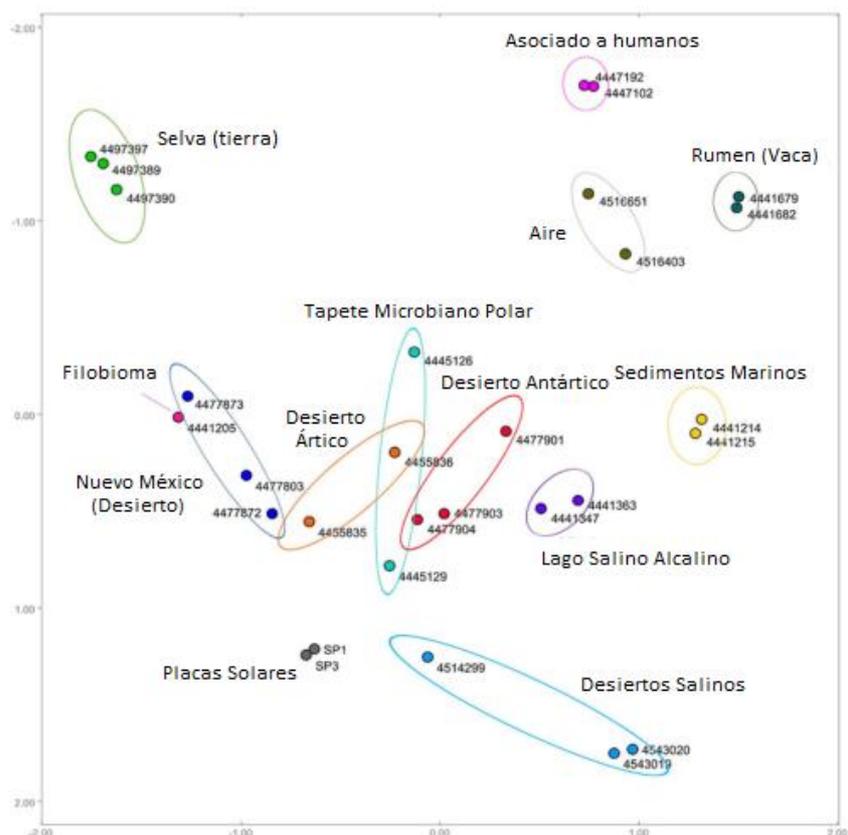


Figura 1.3. Análisis de Coordenadas Principales (PCoA) a partir del perfil taxonómico extraído de un conjunto de metagenomas de diferentes ecosistemas. *Adaptado de Dorado-Morales et al., 2016.*

Además, se encontraron *Deinococcus*, *Hymenobacter* y *Sphingomonas*, como algunos de los géneros mayoritarios. Interesantemente, algunas especies englobadas en estos habían sido descritas con anterioridad como resistentes a la desecación, así como a las altas temperaturas y dosis de radiación. De hecho, una revisión sistemática del lugar donde se aislaron por primera vez las 50 especies bacterianas más abundantes en placas solares revelaron la adaptación de estas a condiciones extremas, ya que muchas de ellas fueron halladas en desiertos, tanto fríos como cálidos, en aguas termales o en zonas de alta salinidad (Dorado-Morales *et al.*, 2016).

En lo referente al metabolismo, muchas de las especies identificadas eran productoras de pigmentos, principalmente carotenoides, así como de esfingolípidos (Dorado-Morales *et al.*, 2016). Los primeros juegan un importante papel en la tolerancia de las bacterias frente a la radiación, mientras que los segundos han sido descritos como mediadores de la adhesión a superficies de sílice y poliamida (Tian & Hua, 2010; Gutman *et al.*, 2014).

Recientemente, se ha secuenciado el microbioma de las placas solares de la ciudad de Berkeley (California) mediante una estrategia metagenómica. Los resultados de este estudio están siendo analizados actualmente en el laboratorio en paralelo al desarrollo de este trabajo. Con esto se pretende observar si existen nexos comunes entre ambas comunidades microbianas y cuantificar hasta qué punto las condiciones geográficas determinan la selección funcional y taxonómica del microbioma.

No obstante, desde el punto de vista ecológico existen otra serie de preguntas biológicas que permanecen sin respuesta: **¿cuál es el origen del microbioma de estas estructuras?** ¿se trata de una simple deposición de microorganismos o estamos ante una comunidad microbiana compleja acondicionada a las distintas presiones selectivas? ¿cuáles son los elementos transportadores de las distintas especies microbianas? ¿la colonización de las placas solares ocurre siempre de la misma manera?, es decir, ¿se trata de un proceso universal?, ¿existen especies marcadoras que demuestren la universalidad del proceso?

1.4.3 Origen del microbioma de placas solares: hipótesis

Las investigaciones previas de Dorado-Morales *et al.* (2016) parecen descartar a los humanos como fuente de la comunidad microbiana presente en las placas solares. Este hecho cobra todavía más sentido si tenemos en cuenta que estas estructuras se suelen instalar normalmente en lugares con acceso restringido para las personas, tales como azoteas de edificios o granjas solares. De este modo, la superficie donde habitan los microorganismos es raramente manipulada por los usuarios o técnicos.

Habiendo descartado en primera instancia a los humanos como vehículo transportador de las distintas especies microbianas, la opción más evidente es que sea el aire quien ejerza esta función. Con respecto a esto, existen varios estudios que demuestran que incluso en instalaciones con un elevado tráfico de personas como oficinas o universidades, la composición del aire exterior determina en gran medida el microbioma de dichas instalaciones (Meadow *et al.*, 2014; Adams *et al.*, 2015).

Sin embargo, este hecho no explicaría la presencia de una comunidad microbiana tan parecida a otros ambientes extremos naturales. Por tanto, la **hipótesis inicial** que se plantea es que los microorganismos que se encuentran en las placas solares tienen un origen en el propio desierto y que son las tormentas de polvo las que ejercen un papel vehicular en el transporte de estos hasta la superficie de los paneles fotovoltaicos. Así, la **hipótesis nula** sería que este microbioma sea una simple deposición de los microorganismos presentes en el aire.

Existen varias razones que apoyarían *a priori* la hipótesis que se propone. Una de las más obvias es el hecho de que ambos ecosistemas están sometidos a condiciones extremas muy parecidas. Con respecto a esto, al igual que pasa en las placas solares, los desiertos se caracterizan por presentar altas temperaturas y grandes fluctuaciones entre el día y la noche y elevados niveles de radiación ultravioleta (Lester *et al.*, 2007).

Por otro lado, es común que en los desiertos se formen tormentas de polvo capaces de migrar miles de kilómetros. En este sentido, se calcula que existen unas siete áreas que ejercen un papel fundamental como fuentes del polvo que se distribuye globalmente debido al movimiento de las masas de aire. Estas suelen ser zonas áridas o semiáridas tales como el desierto del Sáhara, los desiertos árabes o los desiertos y las cuencas situadas en Asia central (Figura 1.4) (Varga, 2012).

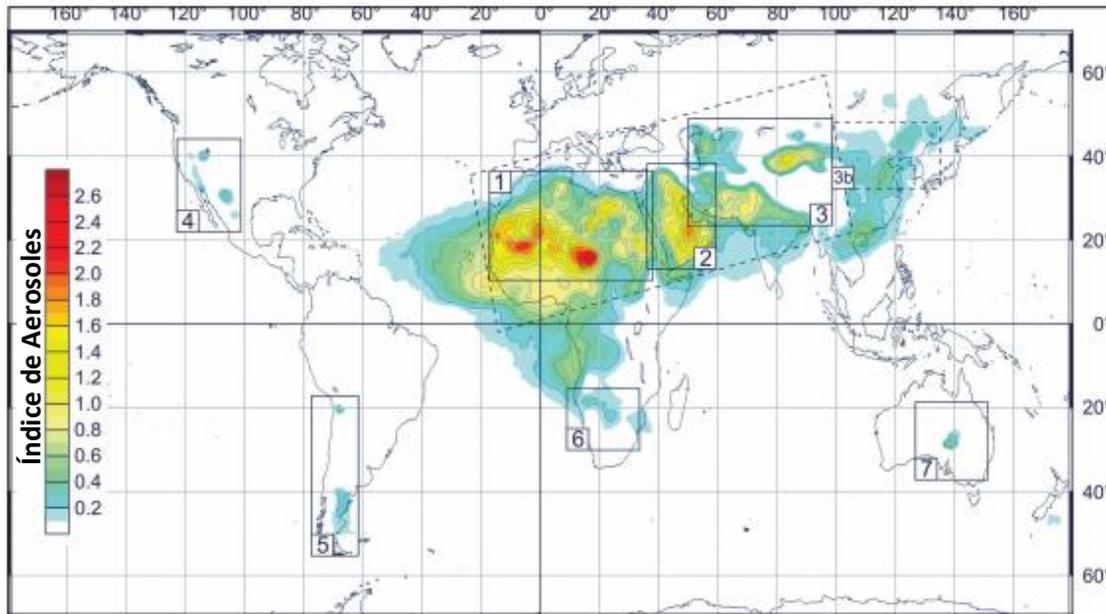


Figura 1.4. Mapa de la distribución global media de las medidas diarias del Índice de Aerosoles capturadas por el satélite TOMS (NASA). Los cuadrados indican las zonas que se proponen como las principales fuentes mundiales de polvo: 1. Sáhara 2. Arabia 3. Asia 3b. Asia del Este 4. América del Norte 5. América del Sur 6. Sudáfrica 7. Australia. *Adaptado de Varga, 2012.*

Por poner un ejemplo, se estima que el viento es responsable de transportar por todo el mundo más de 182 millones de toneladas de polvo proveniente del Sáhara cada año, pudiendo llegar a recorrer distancias de más de 5000 kilómetros (National Aeronautics And Space Administration, 2015). Además, no es raro que estas tormentas lleguen cada cierto tiempo a núcleos urbanos o zonas susceptibles de albergar placas solares (InfoLibre, 2017).

Asimismo, las ciudades en las que se ha estudiado el microbioma de placas solares, esto es, Valencia y Berkeley, coinciden con rutas de transporte de polvo desértico proveniente del Sáhara en el primer caso, y de los desiertos de Asia Central y de América del Norte en el segundo.

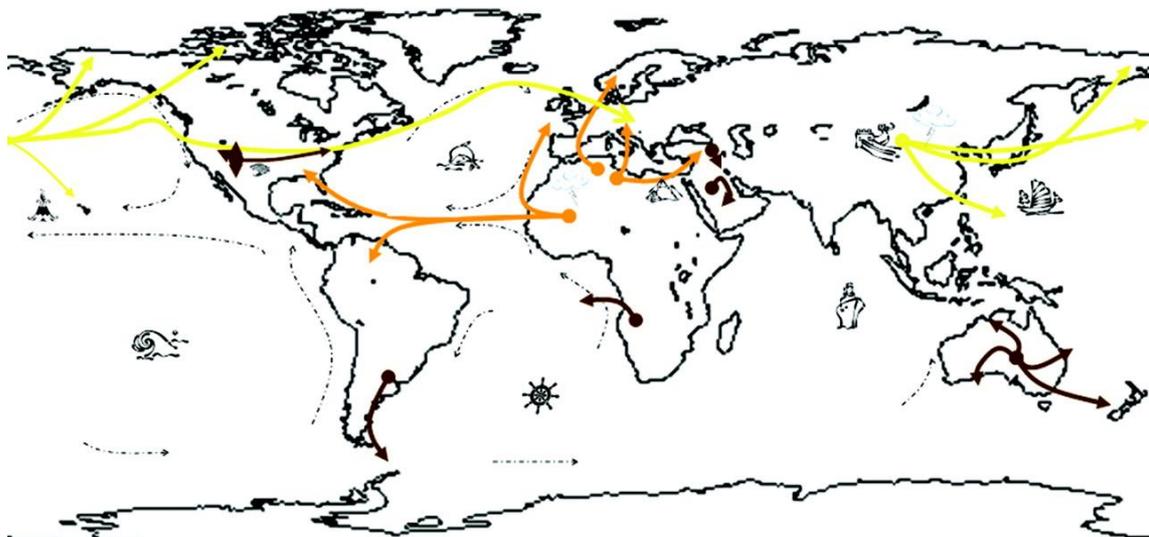


Figura 1.5. Fuentes principales del polvo desértico y su distribución atmosférica global. Las líneas naranjas muestran las rutas del polvo africano, las amarillas del proveniente de los desiertos de Asia Central, las marrones muestran las rutas del polvo proveniente de otros desiertos, y las líneas negras punteadas representan los movimientos del aire. *Adaptado de Griffin, 2007.*

Uno de los primeros científicos que describió el proceso de transporte de las partículas del desierto mediante el aire fue Charles Darwin, quien mientras viajaba por el Atlántico a bordo del Beagle recogió muestras de polvo que enviaría a su amigo el Profesor Christian G. Ehrenberg (Darwin, 1846). Más de siglo y medio después, Gorbushina *et al.* (2007) fueron capaces de observar vida microscópica en aquellas muestras, demostrando que los microorganismos son capaces de adherirse a las partículas de polvo y sobrevivir fácilmente al transporte incluso pasados cientos de años.

Otros estudios recientes han mostrado la diversidad microbiana presente en las muestras de aire con polvo del desierto del Sáhara (Favet *et al.*, 2012) o de los desiertos del Gobi y Takla Maklan (Asia Central) (Cha *et al.*, 2017). Asimismo, existen investigaciones que revelan cómo la comunidad microbiana asociada a estas partículas hacen variar la composición del microbioma autóctono de la nieve de Los Alpes (Meola *et al.*, 2015) y de la interfase agua-aire de algunos lagos en los Pirineos (Barberán *et al.*, 2014), entre otros.

Con todo esto, existen indicios suficientes que parecen indicar que el microbioma de las placas solares podría estar influenciado por las partículas provenientes de las tormentas de polvo desértico. Por lo tanto, en el siguiente trabajo se propone una aproximación de metagenómica comparativa para explorar en qué grado estos ecosistemas están ecológicamente conectados y resolver así el misterio sobre el origen del microbioma de estas estructuras.

2. OBJETIVOS

El objetivo general de este trabajo es discernir el origen del microbioma presente en la superficie de las placas solares y evaluar, en la medida de lo posible, el grado de universalidad en el proceso de colonización de estas estructuras.

Para la consecución de este objetivo, se realizarán distintos análisis metagenómicos combinando muestras secuenciadas por el grupo con secuencias publicadas en otros estudios. Así, los objetivos específicos planteados son:

- Comparar los perfiles taxonómicos y funcionales de placas solares de la ciudad de Valencia con distintos ecosistemas susceptibles de ser el germen del microbioma de estos paneles fotovoltaicos: aire de la región mediterránea sin polvo desértico, aire con polvo proveniente del desierto del Sáhara y suelo desértico de este mismo.
- Comparar los perfiles taxonómicos de placas solares de la ciudad de Berkeley con distintos ecosistemas susceptibles de ser el germen del microbioma de estos paneles fotovoltaicos: aire de Colorado, aire con polvo y control proveniente de los desiertos de Asia Central y muestreado en Corea Sur, y suelo de los desiertos estadounidenses de Chihuahua, Mojave y la Gran Cuenca
- Determinar la posible universalidad del proceso de colonización, así como cuantificar el grado de contribución de cada ecosistema germen al microbioma de las placas solares
- Analizar en detalle la biocenosis propia de placas solares en busca de un núcleo común de microorganismos característicos de este ecosistema con respecto a los ecosistemas germen.

3. MATERIALES Y MÉTODOS

3.1 Diseño experimental

Debido al elevado coste y tiempo que se emplearía en poner en marcha un proyecto puramente experimental que permitiera responder a la pregunta que se está planteando, se optó por realizar una estrategia de metagenómica comparativa a partir de estudios previamente publicados.

Para ello, se realizó una revisión sistemática de la bibliografía y las bases de datos metagenómicas en busca de secuencias provenientes de aquellos ecosistemas que se proponen como germen del microbioma de placas solares, esto es, desiertos, polvo desértico, y aire.

Tras la primera búsqueda, se encontraron un total de 31 estudios con muestras que podrían resultar de interés para el presente trabajo (Tabla 7.1, en Anexo). De entre ellos, solo 6 proyectos seguían una estrategia de secuenciación de genomas completos, mientras que el resto se basaban en genes marcadores, y en concreto, en 16S rRNA. Por ello, se decidió seleccionar únicamente este último tipo de estudios debido a la complejidad que resultaría de comparar datos provenientes de ambas aproximaciones.

En este sentido, el trabajo se divide principalmente en **tres ejes** con el fin de estudiar cada uno de los casos particulares, así como la universalidad del proceso de colonización de los paneles:

1. Estudio comparativo del microbioma de placas solares de Valencia (España).
2. Estudio comparativo del microbioma de placas solares de Berkeley (EE. UU.).
3. Estudio comparativo global usando muestras de ambos estudios.

3.2 Selección de muestras

3.2.1 Secuencias propias

El grupo disponía de una serie de secuencias propias correspondientes a los estudios previos a este trabajo. En concreto, se trata de muestras de placas solares de Valencia y de Berkeley. Además, se secuenció una muestra de polvo recogida en la ciudad de Valencia durante una tormenta de polvo del Sáhara. Las características de todas estas secuencias vienen detalladas en la Tabla 3.1.

Tabla 3.1. Set de datos metagenómicos de las muestras secuenciadas en estudios anteriores del grupo.
P.S. = Placas solares; WGS = *whole genome shotgun*; 16S = Secuenciación del gen 16S rRNA.

Muestra	Estrategia experimental	Tecnología de secuenciación	Fecha Muestreo
P. S. Valencia	16S/18S	Ion Torrent	Junio 2014
P.S. Valencia	WGS	Illumina	Junio 2014
Polvo Sáhara	16S	Illumina	Abril 2015
P. S. Berkeley	WGS	Illumina	Agosto 2016

3.2.2 Estudio comparativo de placas solares: Valencia

Se realizó una selección de secuencias provenientes de muestras del desierto del Sáhara, así como de aire urbano de ciudades cercanas a Valencia o con unas características climáticas similares, ya que no existen estudios de este tipo realizados en esta ciudad. Además, se añadieron muestras de aire rico en polvo del Sáhara y aire control (días antes o después de la tormenta de polvo) recogidas en otras ciudades mediterráneas. Todas las características de las secuencias que se seleccionaron para este análisis se describen en la Tabla 3.2. Existen varios estudios que no contenían ninguna información relativa a la descarga de las secuencias desde repositorios públicos. En esos casos se contactó con los autores para solicitar una copia de las secuencias.

Tabla 3.2. Set de datos metagenómicos de las muestras seleccionadas en el estudio comparativo de Valencia. Se incluyen las secuencias propias que se usaron en el análisis. Todos los estudios incluidos seguían una estrategia experimental basada en la secuenciación del gen 16S rRNA.

Tipo de muestra	Localización	Estudio	Tecnología de secuenciación	Incluido en análisis
Placas Solares	Valencia	Dorado-Morales <i>et al.</i> , 2016	Ion Torrent	Sí
Polvo Sáhara	Valencia	-	Illumina	Sí
Suelo Sáhara	Chad	Favet <i>et al.</i> , 2012	Illumina	Sí
Suelo Sáhara	Mauritania	Barberán <i>et al.</i> , 2014	454	No ³
Polvo Sáhara / Aire Control	Rejovot (Israel)	Mazar <i>et al.</i> , 2016	Illumina	Sí ²
Polvo Sáhara / Aire Control	Beerseba (Israel)	Katra <i>et al.</i> , 2014	454	No ³
Polvo Sáhara / Aire Control	Cerdeña	Roselli <i>et al.</i> , 2015	Illumina	Sí ²
Aire Urbano	Milán	Bertolini <i>et al.</i> , 2013	Illumina	Sí
Aire Urbano	Milán/Venecia ¹	Gandolfi <i>et al.</i> , 2015	Illumina	Sí

1. Solo se tuvieron en cuenta las secuencias provenientes de Venecia en este caso.
2. En estos casos no se recibió información esencial para el análisis de las secuencias (presencia de barcodes, correspondencia entre barcodes y muestras, pre-tratamiento, etc.)
3. No se recibió respuesta por parte de los autores.

3.2.3 Estudio comparativo de placas solares: Berkeley

Se seleccionaron muestras de todos los desiertos de América del Norte cercanos a Berkeley (ver Figura 3.1), así como de los desiertos de Asia Central (Takla Makan y Gobi), debido a las corrientes de aire que conectan estas dos localizaciones (ver Figura 1.5). Por otra parte, se usaron muestras de aire (sin polvo) del Storm Peak Laboratory (Colorado) como control. Finalmente, como no fue posible conseguir muestras de polvo desértico recogido en la propia ciudad de Berkeley o cerca de esta, se añadió al estudio muestras de polvo procedente de los desiertos de Asia Central, así como aire control (sin polvo) de la misma región (ambas muestras fueron capturadas en Seúl, Corea del Sur). Todas las características de las secuencias

que se seleccionaron para este análisis se describen en la Tabla 3.3, donde se indican también las secuencias que no fue posible obtener.

Tabla 3.3. Set de datos metagenómicos de las muestras seleccionadas en el estudio comparativo de Berkeley. Se incluyen las secuencias propias que se usaron en el análisis. WGS = *whole genome sequencing*; 16S = Secuenciación del gen 16S rRNA

Tipo de muestra	Localización	Estudio	Estrategia experimental	Tecnología de secuenciación	Incluido en análisis
Placas Solares	Berkeley (EE.UU.)	-	WGS	Illumina	Sí
Suelo La Gran Cuenca	EE.UU.	Steven <i>et al.</i> , 2013	16S	454	Sí
Suelo Sonora	EE.UU.	Andrew <i>et al.</i> , 2012	16S	454	No ¹
Suelo Mojave	EE.UU.	Fierer <i>et al.</i> , 2012	16S	Illumina	Sí
Suelo Mojave	EE.UU.	Lauber <i>et al.</i> , 2009	16S	454	Sí
Suelo Chihuahua	EE.UU.	Fierer <i>et al.</i> , 2012	16S	Illumina	Sí
Suelo Takla Makan	China	An <i>et al.</i> , 2013	16S	454	No ²
Suelo Gobi	Mongolia	An <i>et al.</i> , 2013	16S	454	No ²
Aire Control	EE.UU.	Bowers <i>et al.</i> , 2012	16S	454	Sí
Polvo/Control Desierto Asia	Corea del Sur	Cha <i>et al.</i> , 2017	16S	454	Sí

1. Estudio no incluido en el análisis porque no se recibió la información relativa a la relación de cada archivo con el tipo de muestra que se estudiaba en cada caso.
2. No se recibió respuesta por parte de los autores.

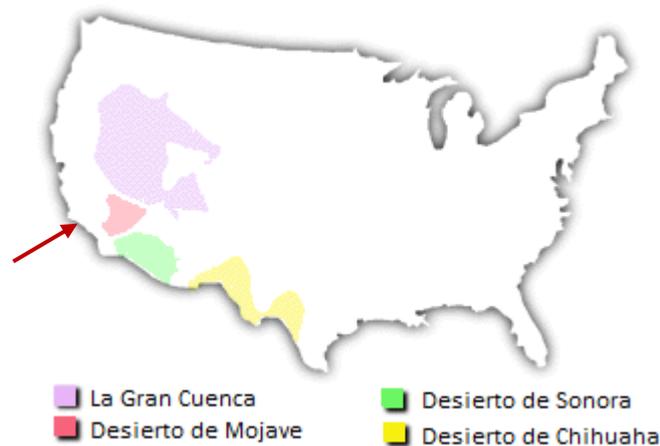


Figura 3.1. Desiertos de EE. UU. Solo se incluye la parte estadounidense de los desiertos de Sonora y Chihuahua. La flecha roja indica la posición de Berkeley (California). Adaptado de <http://www.desertusa.com/north-american-deserts.html>

3.2.4 Estudio comparativo global

Para este estudio se seleccionaron las muestras de los estudios de Valencia y Berkeley que resultaron más informativas

3.3 Análisis bioinformático

3.3.1 Pre-procesamiento de secuencias

Con el fin de evitar sobredimensionar el análisis, tan solo se seleccionaron tres réplicas experimentales en aquellos casos en que los estudios seleccionados incluyeran un elevado número de estas.

En lo referente al tratamiento de las secuencias, aquellas muestras que presentaban cientos de miles de secuencias fueron filtradas aleatoriamente ajustando el número de estas a 100.000 por razones de optimización computacional.

Por otro lado, las secuencias que se recibieron sin ser previamente filtradas por tamaño y calidad de secuencia fueron ajustadas siguiendo las mismas indicaciones que aparecían en el estudio del que fueron extraídas.

3.3.2 Asignación taxonómica

La identificación de las secuencias se realizó en cada uno de los tres casos descritos con anterioridad mediante la herramienta bioinformática QIIME (Caporaso *et al.*, 2010). El flujo de trabajo seguido hasta la obtención de los archivos de salida se indica en la Figura 3.2.

El primer paso fue convertir todas las secuencias descargadas a formato FASTA, puesto que es el formato con el que trabajan los scripts informáticos de QIIME. Para ello, se usaron scripts de elaboración propia capaces de transformar tanto los formatos FASTQ como los QSEQ. A continuación, se unieron todas las secuencias en un único archivo usando el script de QIIME *add_qiime_labels.py*.

El segundo paso consistió en agrupar las secuencias del 16S rRNA en unidades taxonómicas operativas (OTUs, por sus siglas en inglés). En este contexto, el término OTU hace referencia a un grupo de secuencias que son consideradas como una única unidad de diversidad microbiana dentro del análisis debido a su similitud nucleotídica (Koepfel y Wu, 2013). Para realizar este paso se usó el script *pick_otus.py*. El umbral de similitud se fijó en el 97%, ya que es el más aceptado cuando se quiere trabajar a nivel de especie (Stackebrandt y Goebel, 1994). La similitud de secuencia se determinó mediante el método *uclust* (Edgar, 2010). Además, también se habilitó la opción que permitía la coincidencia de la cadena reversa para poder tener en cuenta las muestras que habían sido secuenciadas en sentidos opuestos. El resto de parámetros se mantuvo por defecto.

Posteriormente, se seleccionó una secuencia para representar a cada OTU mediante el script *pick_rep_set.py*. Para ello, se usaron los parámetros por defecto y se seleccionó arbitrariamente la primera secuencia de cada *cluster* como representativa. Seguidamente, se asignó la taxonomía de dichas secuencias mediante el script *assign_taxonomy.py* usando el método *blast* (Altschul *et al.*, 1990). La base de datos que se usó como referencia fue *Greengenes* (Desantis *et al.*, 2006) en su versión más reciente (*Greengenes 13_8*). El umbral de *e-value* máximo para dar por válida una asignación se estableció en su valor por defecto (0,001), al igual que el resto de parámetros. Además, se eliminaron aquellas secuencias que no

recibieron ninguna asignación o que fueron catalogadas como mitocondrias o cloroplastos, ya que en estos casos las secuencias puede provenir de restos de polen y no de microorganismos.

Por último, se usaron los scripts *make_otu_table.py* y *summarize_taxa.py* para obtener *outputs* analizables por el resto de programas que se describirán a continuación. En cualquier caso, todos los archivos de salida que se obtuvieron fueron configurados para presentar la abundancia en números absolutos y la taxonomía hasta el nivel de género.

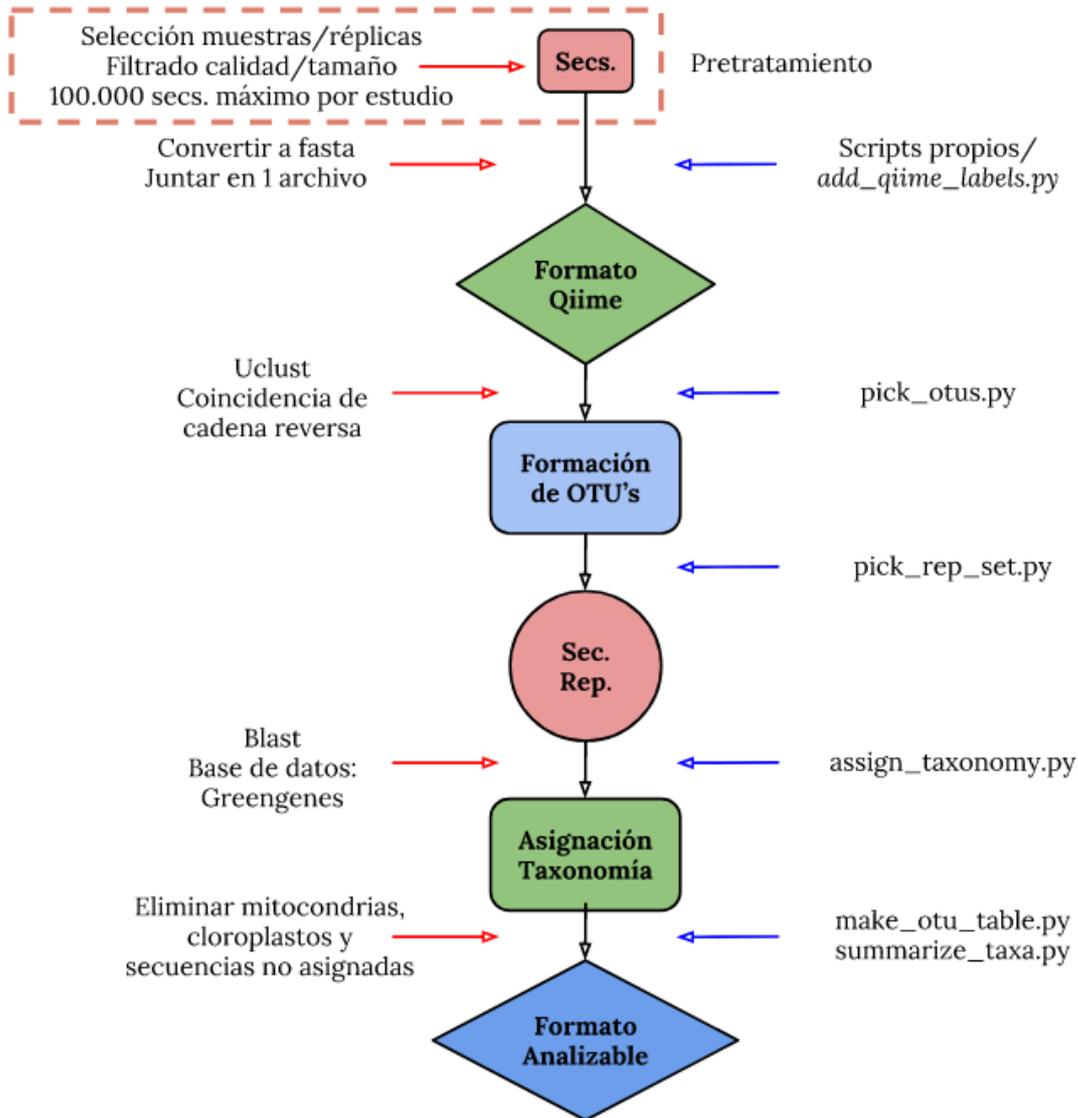


Figura 3.2. Flujo de trabajo utilizado para la transformación de las secuencias y la asignación taxonómica. Las flechas rojas indican cambios en la configuración por defecto establecida para cada paso, mientras que las azules hacen referencia a los scripts informáticos usados en cada caso (los que no se indican como propios pertenecen al software bioinformático QIIME). Sec. = secuencia; Rep. = representativa. OTU = unidad taxonómica operativa.

3.3.3 Comparación de los perfiles taxonómicos

Los diferentes perfiles taxonómicos fueron procesados con MEGAN (Huson y Mitra, 2012). Los datos fueron normalizados y la distancia entre los perfiles fue calculada por el método de Bray-Curtis (Beals, 1984) y representadas en los siguientes formatos:

- PCoA: algoritmo que transforma distancias filogenéticas en distancias físicas en un plano de 2 dimensiones (Grower, 2005)
- Neighbor-Net: árbol consenso que representa todas las componentes en que se descompone las distancias (Bryant y Moulton, 2004).

Por otro lado, se usó el programa estadístico STAMP (*Statistical Analysis of Metagenomic Profiles*) (Parks *et al.*, 2014) para la comparación de la abundancia relativa de los géneros bacterianos de interés presentes en los distintos ecosistemas incluidos en el estudio comparativo global.

3.3.4 Cuantificación de la fuente de origen del microbioma

A fin de tener una cuantificación de los ecosistemas contribuyentes a la formación del microbioma de las placas solares se empleó el método SOURCETRACKER (Knights *et al.*, 2011) tanto en el estudio comparativo de Valencia, como en el de Berkeley. Para ello, se usó como *input* el archivo de salida obtenido tras realizar el flujo de trabajo de QIIME descrito en la Figura 3.2, y se siguieron los mismos pasos que se describen en el tutorial de SOURCETRACKER para QIIME (http://qiime.org/tutorials/source_tracking.html), excepto el de filtrado, ya que de este modo se perdía información de interés.

3.3.5 Inferencia funcional

Se realizó una predicción funcional inferida a partir del perfil taxonómico basado en el gen 16S rRNA de las muestras de aire control, polvo desértico, suelo del Sáhara y placas solares que se incluyeron en el análisis comparativo de Valencia (ver Tabla 3.2). Con este propósito se usó PICRUST (*Phylogenetic Investigation Of Communities By Reconstruction Of Unobserved States*), un software bioinformático que predice el contenido funcional de una determinada comunidad microbiana mediante el uso de la información genómica de OTUs de referencia que coinciden con las OTUs secuenciadas presentes en las muestras (Langille *et al.*, 2013).

Para obtener esta información se siguieron los pasos que se indican en el tutorial de PICRUST de GITHUB (https://picrust.github.io/picrust/tutorials/metagenome_prediction.html). Como *input* se usó el archivo de salida del script *pick_closed_reference_otus.py* de QIIME. Este script usa el método *closed reference* para seleccionar las OTUs, el cual consiste en comparar las secuencias contra una referencia preestablecida y descartar aquellos *reads* que no hayan dado resultado (Rideout *et al.*, 2014). Se usó un umbral del 97% de similitud de secuencia para seleccionar las OTUs y se permitió la coincidencia de la cadena reversa. El resto de parámetros se mantuvieron por defecto. La asignación taxonómica se hizo frente a la base de datos Greengenes en su versión anterior (13_5), ya que la versión actual no está disponible para su uso en PICRUST.

Las anotaciones funcionales se obtuvieron a través de la base de datos COG (*Cluster of Orthologous Groups of Proteins*) y aportaban información sobre proteínas concretas que se inferían a partir de las distintas OTUs. No obstante, como el interés del estudio era observar cambios en grandes grupos de funciones (p.e. síntesis de metabolitos secundarios o motilidad celular), las anotaciones obtenidas fueron convertidas y agrupadas en categorías funcionales más generales mediante un script de elaboración propia tomando como referencia la información COG actualizada en 2014 (<ftp://ftp.ncbi.nih.gov/pub/COG/COG2014/data>).

Finalmente, esta información fue procesada mediante STAMP de manera análoga a la descrita anteriormente, pero teniendo en cuenta que ahora los datos hacen referencia a la

función y no a la taxonomía. En esta ocasión, se decidió representar los resultados mediante un dendrograma y un mapa de calor (o *heatmap*).

Cabe destacar que en el caso de Berkeley no fue posible realizar este análisis. Pese a que se intentó comparar el output de PICRUSt directamente con los datos metagenómicos de las muestras de placas solares, resultó inviable obtener un formato que permitiera el análisis de una forma computacionalmente viable.

3.3.6 Adaptación de datos de WGS

Las secuencias que se disponían provenientes de las placas solares de Berkeley (California) correspondían a una estrategia metagenómica y no de genes marcadores (16S rRNA) como el resto de muestras. Por ello, estas muestras tuvieron que ser transformadas en un formato comparable al resto. Con este objetivo se utilizó el algoritmo LMAT (Ames *et al.*, 2013) para obtener asignaciones taxonómicas para cada lectura de manera rápida. Posteriormente, se elaboraron scripts que permitieron unir el output de dicha herramienta con el obtenido por el flujo de trabajo de QIIME.

Como control positivo de que esta aproximación podía dar lugar a resultados comparables a los que se obtendrían si se hubieran secuenciado las placas solares de Berkeley siguiendo una estrategia de genes marcadores, se realizó un análisis adicional con el grupo de muestras de Valencia en el que se incorporó una muestra de placas solares que había sido secuenciada mediante una estrategia metagenómica y transformada siguiendo los pasos anteriormente descritos.

4. RESULTADOS Y DISCUSIÓN

4.1 Estudio comparativo de placas solares: Valencia

En este estudio se incluyeron secuencias correspondientes a muestras de aire urbano de ciudades italianas, del desierto del Sáhara y de polvo desértico procedente del mismo y capturado en Valencia, además de las muestras de placas solares. Cabe destacar que, por limitaciones experimentales, sólo pudo disponerse de una muestra de polvo desértico.

El número total de secuencias que se incluyeron en el análisis teniendo en cuenta el conjunto de todas las muestras consideradas fue 1.816.773.

4.1.1 Comparación de los perfiles taxonómicos

De manera preliminar, se optó por realizar un análisis en el que se filtraran aquellas OTUs que no fueran mayoritarias, es decir, que no representaran a un número de secuencias considerable. El umbral se estableció en 25 secuencias por OTU (cabe recordar que en nuestros análisis, cada OTU corresponde con una especie microbiana, como se explica en el apartado 3.3.2).

En los resultados obtenidos de la comparación de los perfiles taxonómicos se pudo observar una gran diferencia entre las muestras de los estudios de tipo polvo desértico y aire control de Israel y Cerdeña respecto al resto de muestras (Figura 4.1).

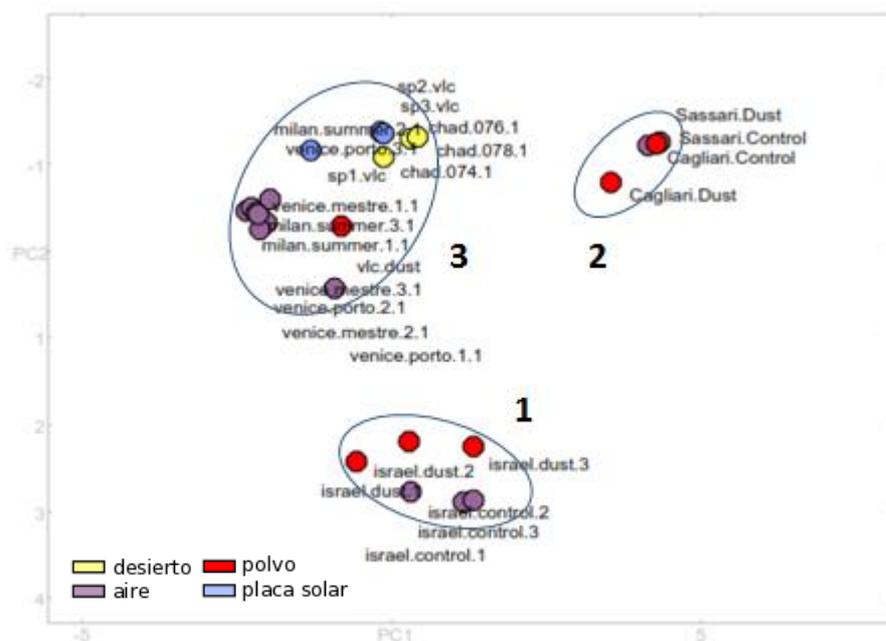


Figura 4.1. Análisis de Coordenadas Principales (PC1 (25,3%) vs PC2 (17,8%)) realizado con los perfiles taxonómicos de las muestras incluidas en el análisis comparativo de placas solares de Valencia (máximo nivel de asignación taxonómica: género). Las OTUs que no representaban a más de 25 secuencias fueron eliminadas. Se observan tres grandes grupos constituidos por: 1. Muestras del estudio de Israel (Polvo y Aire Control); 2. Muestras del estudio de Cerdeña (Polvo y Aire Control); 3. El resto de muestras del análisis (aire sin polvo, polvo desértico Valencia, suelo desértico y placas solares)

El hecho de no haber recibido suficiente información acerca del contexto en que estas secuencias fueron tratadas antes de descargarlas pudo haber ocasionado que se cometieran

errores durante el tratamiento bioinformático, que podrían explicar la formación de estos tres grandes grupos. De hecho, se observaron errores recurrentes en la asignación de las secuencias provenientes de dichos estudios, por lo que se decidió eliminarlas del análisis.

Se realizó un nuevo análisis con las muestras restantes y se obtuvieron los resultados que se presentan en la Figura 4.2. Es interesante observar como las muestras correspondientes a los cuatro ecosistemas que se están considerando (aire, polvo desértico, suelo desértico y placas solares) se agrupan formando *clusters* bien diferenciados, lo que sugiere la individualidad de los mismos. No obstante, cabe remarcar el hecho de que las muestras que supuestamente tienen un origen en el desierto (polvo desértico y placas solares, además del propio suelo desértico) se encuentran en el mismo lado de la componente principal 1. Esta disposición podría indicar que, pese a tratarse de comunidades microbianas diferentes, parece existir un nexo de unión entre todas ellas.

Por otro lado, la muestra correspondiente al polvo desértico recogido en Valencia se encuentra también próxima a las muestras de aire urbano. Esta similitud concuerda con lo descrito por Roselli *et al.* (2015) en su estudio de Cerdeña, donde se determinó que existía un gran núcleo de especies microbianas que se compartían entre las muestras de aire en días libres de polvo y en días de tormenta de polvo (Figura 7.1, en Anexo).

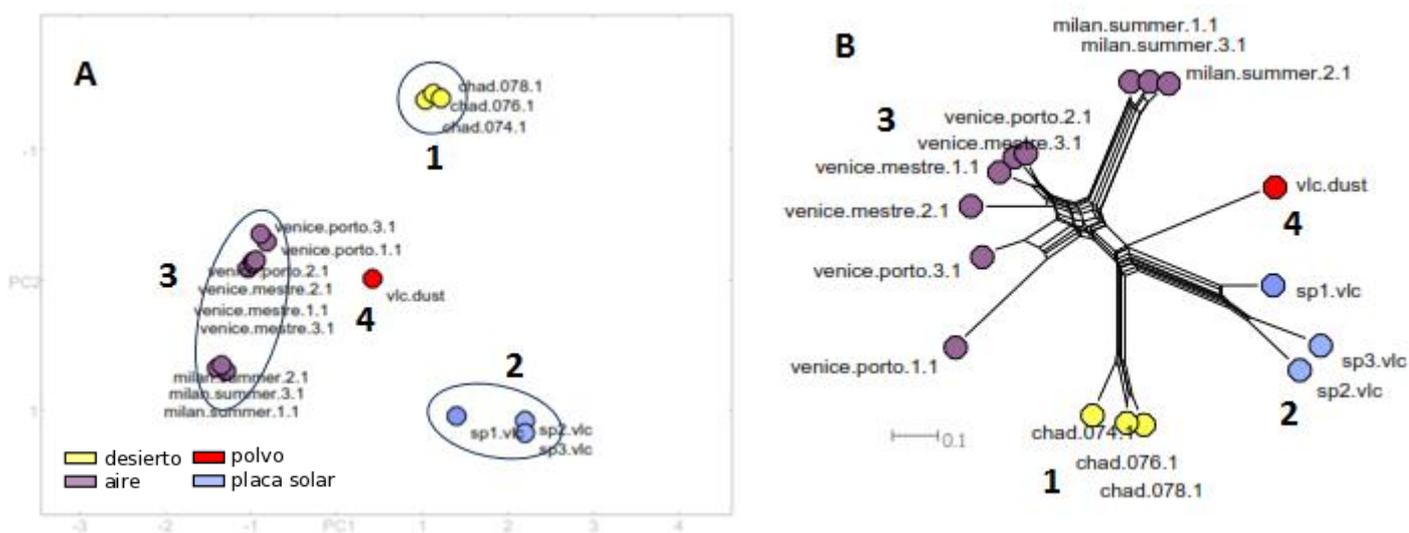


Figura 4.2. Comparación de los perfiles taxonómicos una vez eliminadas las muestras de polvo/aire control de Israel y Cerdeña (máximo nivel de asignación taxonómica: género). Las OTUs que no representaban a más de 25 secuencias fueron eliminadas. **A)** Análisis de Coordenadas Principales. (PC1 (36,1%) vs PC2 (23,6%)). **B)** Neighbor-Net. Se observan cuatro grupos constituidos por: **1.** Suelo del Sáhara (Chad); **2.** Placas solares de Valencia **3.** Aire urbano (Italia); **4.** Polvo del Sáhara recogido en Valencia;

Habiendo obtenido estos resultados, se decidió realizar un nuevo análisis teniendo en cuenta todas las OTUs, es decir, sin establecer ningún umbral de abundancia. Como las muestras de aire habían demostrado compartir un alto grado de similitud, se decidió escoger solo tres réplicas para optimizar la parte computacional. Las muestras de Israel y Cerdeña no se tuvieron en cuenta por los motivos antes mencionados.

Tal y como se observa en la Figura 4.3, los resultados que se obtuvieron en este caso presentan variaciones respecto al análisis anterior. Sin embargo, se puede realizar una interpretación muy similar: se trata de ecosistemas independientes que parecen estar conectados atendiendo a cómo la muestra de polvo desértico se dispone prácticamente en medio de las muestras de placas solares y suelo del Sáhara. Además, se continúa apreciando la similitud existente entre esta muestra de polvo y las muestras de aire.

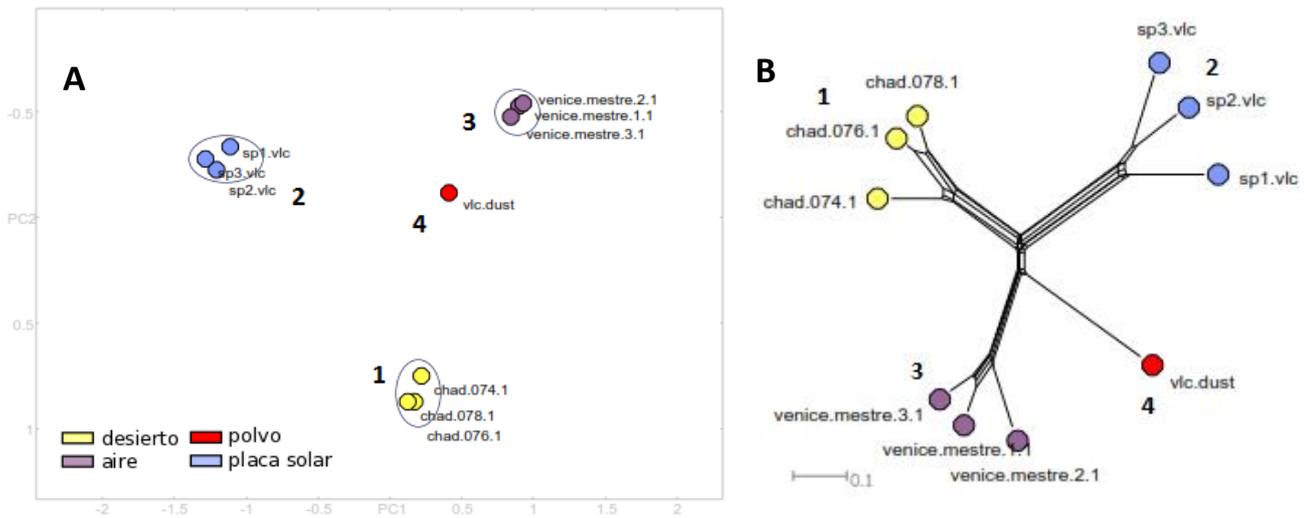


Figura 4.3. Comparación de los perfiles taxonómicos (máximo nivel de asignación taxonómica: género). **A)** Análisis de Coordenadas Principales (PC1 (44,7%) vs PC2 (30,0%)). **B)** Neighbor-Net. Las muestras de Israel y Cerdeña se excluyeron. Solo se tienen en cuenta tres muestras de aire urbano. Se observan cuatro grupos constituidos por: 1. Suelo del Sáhara (Chad); 2. Placas solares de Valencia; 3. Aire urbano (Italia); 4. Polvo del Sáhara recogido en Valencia.

4.1.2 Cuantificación de la fuente de origen del microbioma

Para cuantificar las fuentes de origen del microbioma de placas solares, se utilizó el software SOURCETRACKER, que permite calcular qué proporción de un determinado microbioma proviene de otros ecosistemas “fuente” (Figura 4.4).

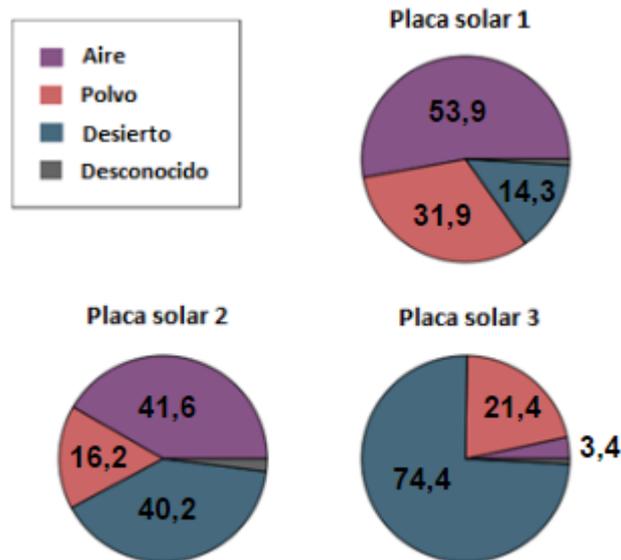


Figura 4.4. Contribución de las distintas fuentes a la formación del microbioma de placas solares de Valencia según el software SOURCETRACKER. Solo se muestran los porcentajes de las influencias mayoritarias

Los resultados obtenidos incidieron en la conexión existente entre el microbioma de placas solares y el resto de ecosistemas estudiados. En este sentido, es interesante observar como el aire urbano ejerció una influencia mayor de la que se pensaba en un primer momento como fuente de la comunidad microbiana de la superficie de los paneles. Esto se puede explicar atendiendo a la gran diversidad microbiana detectada en este ecosistema.

No obstante, el suelo desértico presentó una influencia elevada en todas las placas solares. De hecho, en la placa número 3 se detectó un influjo mayoritario de microorganismos procedentes del Sáhara, mientras que el aire ejerció un peso ínfimo. Esta muestra se correspondía con aquella en la que se detectó una mayor cantidad de *Deinococcus*, un género bacteriano típico de un gran número de desiertos (Yu *et al.*, 2015; Hussain *et al.*, 2016) (Tabla 7.2, en Anexo). Por su parte, el polvo desértico mostró un peso relativo similar como fuente de origen del microbioma de todas las muestras de placas solares, corroborándose así su importancia como vehículo transportador de diferentes microorganismos hasta las placas solares.

4.1.3 Comparación de los perfiles funcionales inferidos

Con el fin de examinar los genes que se seleccionan en cada ecosistema, se realizó una inferencia de los perfiles funcionales a partir de los perfiles taxonómicos usando la herramienta PICRUSt. Los resultados derivados de esta predicción demostraron que existía una estrecha similitud entre las muestras de suelo desértico, placas solares y polvo desértico más allá de la taxonomía (Figura 4.5).

En lo referente a las funciones COG concretas, estos ecosistemas presentaban un aumento a nivel genómico de secuencias dedicadas a motilidad celular, la transducción de señales y los mecanismos de defensa respecto a las muestras de aire. En líneas generales, los microorganismos presentes en el aire los días libres de polvo mostraban un aumento en las funciones dedicadas al metabolismo (transporte y metabolismo de aminoácidos, producción y conversión de energía, etc) y la transcripción, lo que coincide en gran medida con los resultados que se obtuvieron en el anterior estudio de Dorado-Morales *et al.* (2016), a pesar de que las muestras de aire que se usaron en este caso provenían de otra ciudad diferente (Pekín).

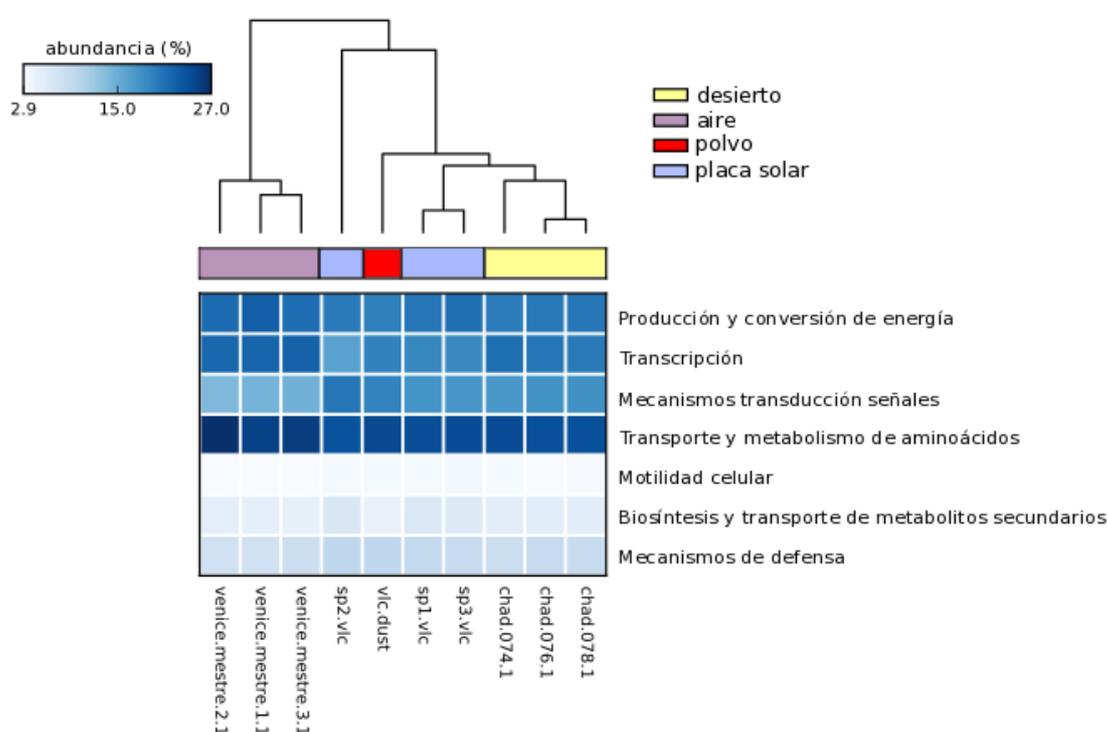


Figura 4.5. Heatmap y dendrograma basado en el análisis comparativo de los perfiles funcionales inferidos de los distintos tipos de ecosistemas estudiados. Se resaltan las funciones COG en las que más diferencias se encontraron entre los distintos tipos de muestra.

Por otro lado, también se observó que en las placas solares existía un mayor número de genes dedicados a la biosíntesis y transporte de metabolitos secundarios. Esto se correspondería con la gran cantidad de microorganismos productores de carotenoides y esfingolípidos presentes en el microbioma de placas solares (Dorado-Morales et al., 2016), lo que confirmaría su gran potencial como fuente de nuevas cepas microbianas super-productoras de estos compuestos, y por tanto, susceptibles de tener aplicación industrial.

En cualquier caso, estos datos deben tratarse como resultado de una inferencia funcional hecha en base a información taxonómica, y precisarían de estudios complementarios para ser validados.

4.2 Estudio comparativo de placas solares: Berkeley

En este estudio se consideraron muestras de aire sin polvo de Colorado, polvo desértico procedente de los desiertos de Asia Central (Takla Makan y Gobi) y aire control tomados en Corea del Sur, así como suelo de tres desiertos distintos de EE.UU. (Mojave, Gran Cuenca y Chihuahua). Todas las secuencias fueron unidas y comparadas posteriormente con las derivadas del estudio de placas solares de Berkeley. Cabe destacar que no fue posible obtener las secuencias de las muestras correspondientes a suelos de los desiertos asiáticos.

El número final de muestras que se consideraron en el análisis fue de 22, incluyendo las muestras de placas solares de Berkeley que se analizaron mediante secuenciación metagenómica, las cuales fueron sometidas a un tratamiento diferente. En total, se analizaron 405.201 secuencias de 16S rRNA, y más de 36 millones de secuencias metagenómicas.

4.2.1 Comparación de los perfiles taxonómicos

A fin de obtener una idea previa de las similitudes existentes entre los distintos microbiomas estudiados, se realizó un análisis preliminar en el que se tuvieron en cuenta todas las muestras excepto las de placas solares (Figura 4.6). Los resultados mostraron que existía una clara separación entre muestras de suelo desértico y las de aire o polvo. Asimismo, se pudo observar que todos los desiertos de Estados Unidos tienen un grado de similitud muy alto atendiendo a su composición microbiana y sin importar la tecnología utilizada para la secuenciación de las muestras (Illumina o 454; Tabla 3.3). En lo referente a las muestras de aire, se aprecia una división entre las muestras de origen asiático y las de origen estadounidense. Dentro de estas últimas, es interesante el hecho de que el perfil taxonómico del polvo desértico se encuentre más cercano a las muestras de desiertos, pese a tener origen en un continente diferente (Asia).

Al incorporar las muestras de placas solares de Berkeley los resultados que se obtuvieron fueron ligeramente diferentes (Figura 4.7 A y B). En este caso, se formaron tres grupos compuestos por cada uno de los tres grandes ecosistemas estudiados: suelo, aire y placas solares.

El grupo más complejo es el correspondiente al aire, puesto que en él se incluyen muestras de aire control (EE.UU. y Corea del Sur) y de polvo desértico, sin apreciar una distinción tan clara como la hallada en el análisis preliminar. Este hecho hace pensar que debe existir un núcleo común de microorganismos aéreos que se mantiene estable a lo largo de la Tierra y que no varía demasiado pese a sufrir episodios de tormentas de polvo. De este modo, se explicaría que al añadir al análisis muestras con perfiles taxonómicos muy similares (placas solares de Berkeley) se fuerce esta agrupación.

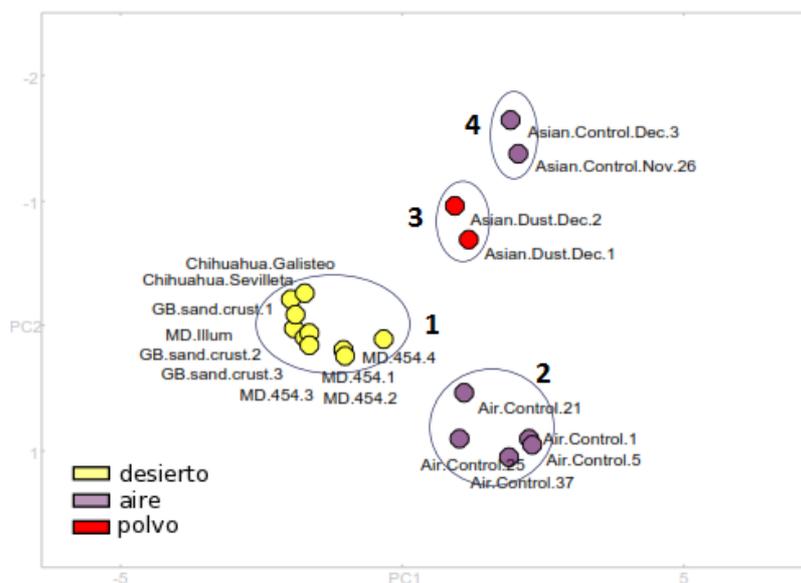


Figura 4.6. Análisis de Coordenadas Principales (PC1 (41,6%) vs PC2 (18,6%)) preliminar realizado a partir de perfiles taxonómicos (máximo nivel de asignación taxonómica: género). Se observan cuatro grupos constituidos por: **1.** Desiertos de EE.UU.; **2.** Aire Control (Colorado); **3.** Polvo desértico (Corea); **4.** Aire Control (Corea)

Por otro lado, se puede observar que existe una distancia relativamente grande entre los tres grupos. Esta percepción podría estar causada por el hecho de que no existe ningún estado intermedio entre los diferentes ecosistemas, al contrario de lo que pasaba en el anterior estudio comparativo donde este papel era ejercido por el polvo desértico recogido en Valencia (ver Figura 4.3). En este sentido, hubiera sido muy interesante disponer de muestras de polvo recogidas en la propia ciudad de Berkeley para ver cómo se comportaban en el análisis.

No obstante, cuando se representaron las componentes 1 y 3 del PCoA, en lugar de las componentes 1 y 2 (la variabilidad explicada disminuyó tan solo un 5%), se obtuvo una distribución que establecía una clara diferencia entre las muestras de aire procedentes de EE.UU y las tomadas en Corea (Figura 4.7 C). Dentro de estas últimas, también existía una separación entre las muestras de aire control y las de polvo desértico. Resulta especialmente interesante la posición de las placas solares, ya que se encuentran a una distancia menor de los desiertos estadounidenses y del polvo asiático que de los dos grupos de aire control. Este hecho incidiría sobre la influencia que ejercen los desiertos en la formación de la comunidad microbiana de la superficie de los paneles fotovoltaicos.

Cabe destacar que la fiabilidad del proceso de transformación de los datos metagenómicos procedentes de las placas solares de Berkeley se corroboró mediante un experimento complementario. Para ello, se usó como modelo el estudio comparativo de Valencia, se introdujo en él una muestra de placas solares de esta misma ciudad secuenciada mediante una estrategia metagenómica, y se comprobó que la muestra transformada se agrupaba correctamente en el análisis (Figura 7.2, en Anexo).

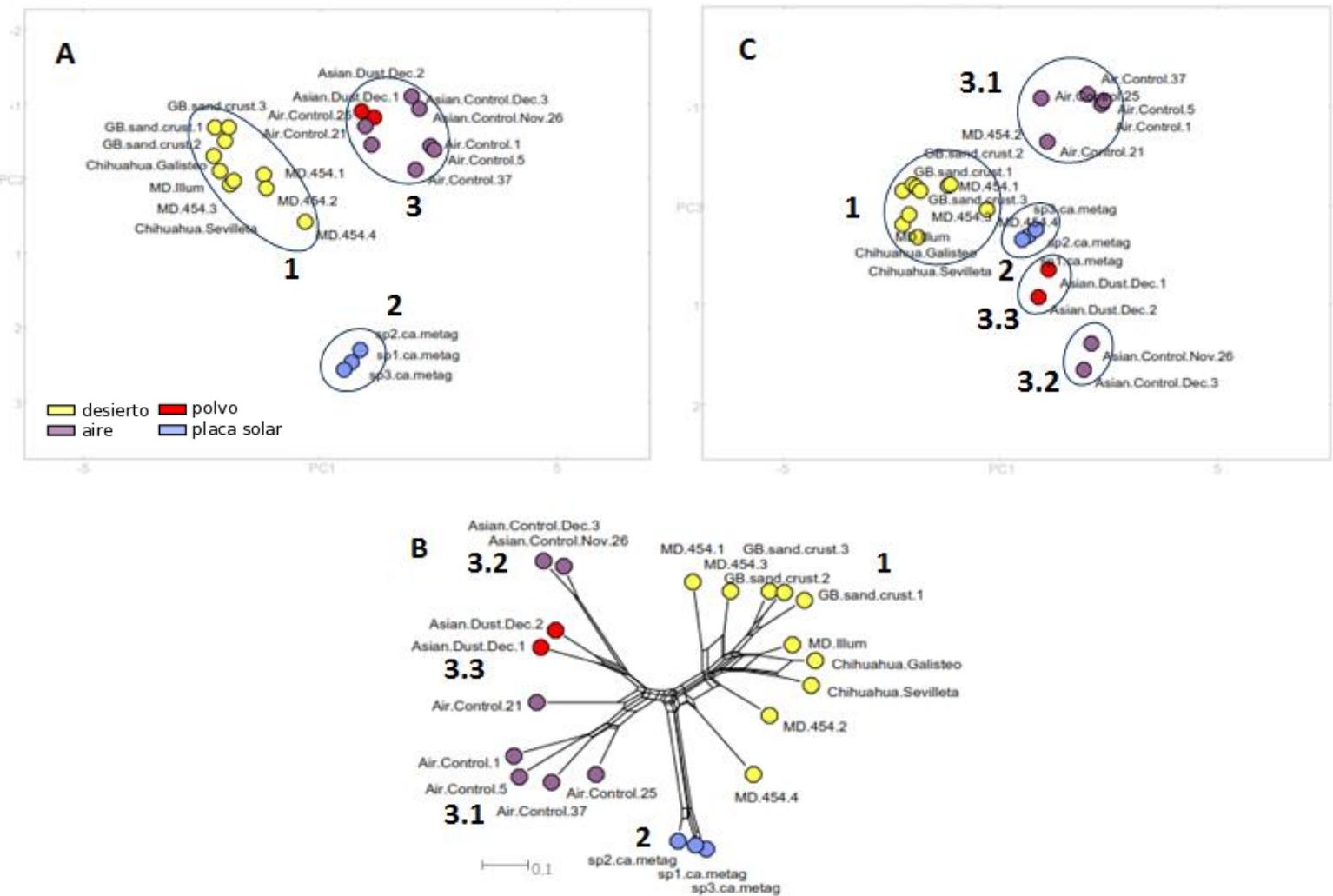


Figura 4.7. Comparación de los perfiles taxonómicos incluyendo las muestras de placas solares de Berkeley (máximo nivel de asignación taxonómica: género). **A)** Análisis de Coordenadas Principales (PC1 (33,5%) vs PC2 (21,1%)). **B)** Neighbor-Net. **C)** Análisis de Coordenadas Principales (PC1 (33,5%) vs PC3 (14,6%)). Se observan 3 grupos que se corresponden con los distintos biomas estudiados: **1.** Suelo desiertos (EE.UU). **2.** Placas solares (Berkeley); **3.** Aire (**3.1** Aire control (EE.UU.); **3.2** Aire control (Corea); **3.3** Polvo desértico (Corea))

4.2.2 Cuantificación de la fuente de origen del microbioma

En esta ocasión, los resultados del análisis con SOURCETRACKER presentan un alto grado de similitud con los obtenidos en el caso de Valencia. Tal y como se puede apreciar en la Figura 4.8, el aire parece tener un gran peso como fuente de origen del microbioma de placas solares. No obstante, el suelo desértico también aporta un contenido elevado de microorganismos a cada muestra, siendo la fuente mayoritaria en la placa solar 3.

Cabe destacar el poco peso de las muestras procedentes de Asia, tanto aire control como polvo desértico. Al mismo tiempo se puede observar como el porcentaje de origen desconocido aumenta en todos los casos de forma análoga. Tomando como base los resultados de Valencia y lo descrito en los apartados anteriores, todo parece indicar que esta aportación desconocida podría corresponderse con el polvo desértico de Berkeley, ya que casualmente representa un porcentaje similar al que presentaba el polvo del Sáhara en el anterior análisis.

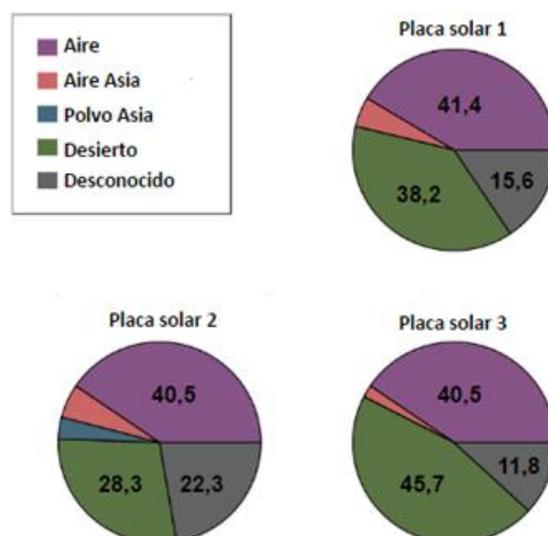


Figura 4.8. Contribución de las distintas fuentes en la formación del microbioma de placas solares de Berkeley según el software SOURCETRACKER. Solo se muestran los porcentajes de las influencias mayoritarias.

4.3 Estudio comparativo global

El número total de secuencias total que se habían seleccionado para realizar los dos experimentos anteriormente descritos ascendía a 2.221.974 correspondientes a 16S rRNA y más de 36 millones secuencias metagenómicas. A fin de realizar un análisis que resultara computacionalmente viable, hubo que realizar una criba de muestras que fueran prescindibles por presentar alta redundancia respecto a otras. En primer lugar, se decidió eliminar las muestras de aire control y polvo desértico de los estudios de Israel y Cerdeña por los motivos comentados en el apartado 4.1.1. También se redujo el número de muestras de aire urbano de ciudades italianas a únicamente 3. Por último, se eliminó el estudio de Corea del Sur (aire control y polvo desértico) al considerarse que su influencia en la formación del microbioma de las placas solares que se están estudiando no era muy elevada. De este modo, se consiguió reducir el número de secuencias de 16S rRNA a 892.973.

La comparación de los perfiles taxonómicos reforzó los resultados obtenidos en los anteriores análisis comparativos. En la Figura 4.9 se observa como los perfiles taxonómicos se agruparon claramente en tres grupos: uno constituido por todas las muestras de suelo desértico, otro formado por las placas solares, y un último grupo en el que se engloban todas las muestras de aire control junto a la muestra de polvo desértico de Valencia. Esto último fortalece la idea presentada en el apartado 4.2.1 de que existe un *core* de microorganismos aéreos que se mantiene constante a lo largo del tiempo y del espacio pese a las variaciones climáticas y las fluctuaciones en el resto de comunidad microbiana.

Por último, también resulta resaltable la fuerte influencia que ejercen las distintas presiones selectivas en el establecimiento de los ecosistemas microbianos. En este sentido, se puede ver cómo ecosistemas que se encuentran separados por miles de kilómetros acaban albergando una comunidad microbiana similar siempre que compartan unas características fisicoquímicas semejantes. El paradigma de todo esto se ve reflejado claramente en las placas solares, ya que independientemente de las condiciones climáticas y geográficas del lugar donde se instalan, el microbioma que se forma acaba teniendo una composición similar.

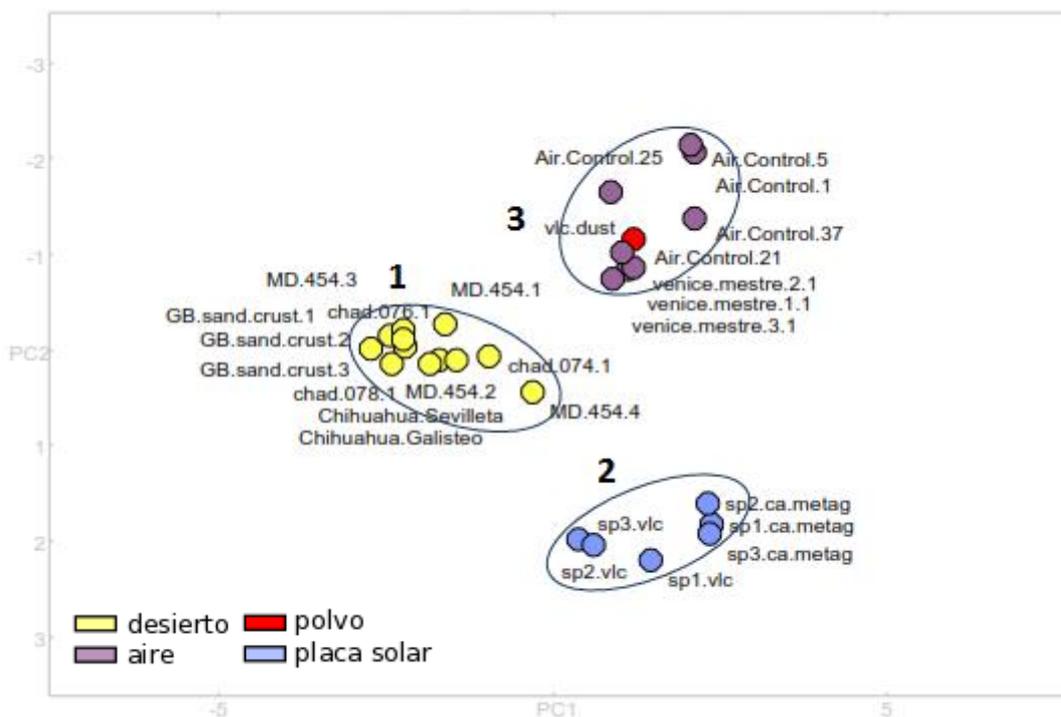


Figura 4.9. Análisis de Coordenadas Principales (PC1 (29,0%) vs PC2 (20,1%)) de todas las muestras incluidas en el estudio comparativo global (máximo nivel de asignación taxonómica: género). Se observan 3 grupos que se corresponden con los distintos biomas estudiados: **1.** Suelo desiertos (EE.UU y Sáhara). **2.** Placas solares (Valencia y Berkeley); **3.** Aire (Aire control (EE.UU.); Aire urbano control (Italia) y polvo desértico (Valencia))

4.3.1 La biocenosis de placas solares

Con el objetivo de comprender y conocer el ecosistema típico de placas solares, se realizó una revisión más profunda de los perfiles taxonómicos de las muestras incluidas en este análisis. En este aspecto, se observó que, a nivel cualitativo, la composición microbiana de estas estructuras era muy similar en ambas localizaciones (Tabla 7.2 y Tabla 7.3, en Anexo), lo que explicaría los resultados presentados.

Al comparar el microbioma de placas solares con el del resto de muestras, se encontró que había cuatro géneros bacterianos que se encontraban sobrerrepresentados: *Hymenobacter*, *Deinococcus*, *Rubellimicrobium* y *Kaistobacter* (Figura 4.10). Los tres primeros engloban especies que han sido descritas anteriormente como productoras de pigmentos (Lee *et al.*, 2016; Kim *et al.*, 2017b; Denner *et al.*, 2006), mientras que *Kaistobacter* está poco caracterizado. Tomando estos resultados en conjunto con los obtenidos en el apartado de inferencia funcional (ver apartado 4.1.3), se pone de manifiesto la importancia de los paneles fotovoltaicos como reservorio de bacterias con elevado interés biotecnológico e industrial.

El género *Hymenobacter* representó de media el 19,5% de la composición microbiana de las placas solares, siendo el único taxón sobrerrepresentado de una forma estadísticamente significativa si se compara con microbiomas del resto de ecosistemas estudiados (p -valor = $9,26e^{-3}$). Se han descrito anteriormente especies pertenecientes a este género que poseen complejos mecanismos antioxidantes y de reparación del ADN (Oh *et al.*, 2016; Kim *et al.*, 2016) que les confieren una gran resistencia a la radiación ultravioleta (UV) (Marizcurrena *et al.*, 2017; Paulino-Lima *et al.*, 2016). Este hecho explicaría la enorme adaptación que presentan estas bacterias al biotopo que conforma la superficie de los paneles solares. Interesantemente, existen otros estudios que se centran en la influencia que tienen los pigmentos producidos por

Hymenobacter en la producción de energía de las células fotovoltaicas (Órdenes-aenishanslins *et al.*, 2016).

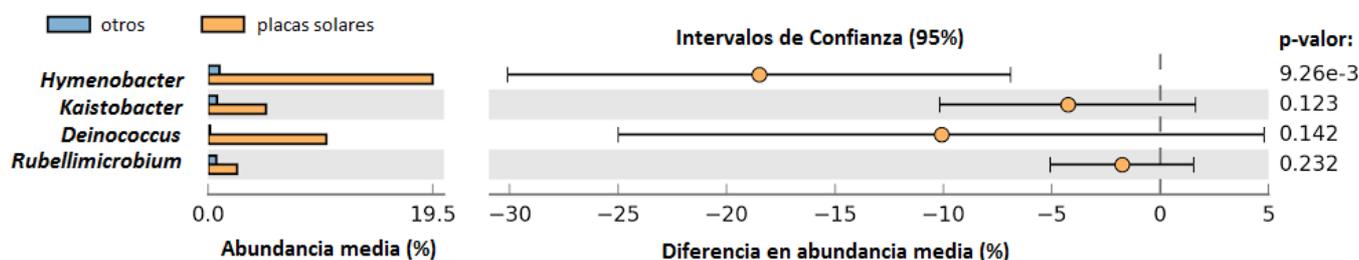


Figura 4.10. Comparación de la abundancia relativa media de los cuatro géneros más sobrerrepresentados en placas solares frente a los otros ecosistemas (polvo desértico, suelo desértico y aire sin polvo).

Rubellimicrobium y *Deinococcus* se caracterizan por ser extremófilos (Köberl *et al.*, 2011). En este sentido, es bien conocido que existen bacterias dentro del género *Deinococcus* adaptadas a condiciones extremas de desecación y exposición a la radiación UV muy similares a las que se encuentran en la superficie de los paneles solares, llegando a desarrollarse incluso bajo condiciones que imitan a las del espacio exterior o el planeta Marte (de la Tour *et al.*, 2017; Ranawat y Rawat, 2017; Frösler *et al.*, 2017).

4.4 Discusión general

4.4.1 Limitaciones del análisis

Hay que tener la realización de un metaanálisis como el descrito en este trabajo conlleva algunas limitaciones que deben ser consideradas.

Una de las más importantes ha sido el no haber podido conseguir las secuencias de todas las muestras que se habían seleccionado en un principio, especialmente las correspondientes a los desiertos de Asia Central. En este aspecto, también cabe resaltar la falta de muestras de polvo desértico recogido en la propia ciudad de Berkeley, lo que hubiera aportado mucha más información sobre el papel de este como vehículo transportador de los microorganismos hasta la superficie de los paneles fotovoltaicos. Además, existen limitaciones derivadas de las diferentes metodologías empleadas en cada estudio, lo que supone una heterogeneidad en términos de estrategias de muestreo, tratamiento de las muestras, tecnologías empleadas en la secuenciación y la falta de uniformidad en el pretratamiento de las secuencias (filtrado de calidad, tamaño, etc). Además, las muestras de placas solares de Berkeley fueron secuenciadas siguiendo una estrategia metagenómica, por lo que tuvieron que ser tratadas independientemente para obtener un archivo de salida comparable con el *output* del resto de muestras. Pese a que los datos se normalizaran, la enorme diferencia en la profundidad de secuenciación pudo desembocar en sesgos y discrepancias en la asignación taxonómica. Todo ello, unido a la falta en muchos casos del contexto en que las muestras fueron secuenciadas, pudo haber originado distorsiones en los resultados.

4.4.2. El origen del microbioma de placas solares

Analizando conjuntamente los resultados obtenidos a lo largo del trabajo se hace evidente que las placas solares albergan un microbioma único, y con una composición muy similar entre localizaciones tan distantes como Valencia y Berkeley. A la luz de nuestros análisis, el microbioma de las placas solares no es el resultado de una simple deposición de los microorganismos presentes en el aire, sino de una interacción de varias fuentes de

microorganismos y de las presiones selectivas a las que las placas solares están sometidas como biotopos.

Nuestros resultados indican que pese al importante papel que ejerce el aire como fuente de microorganismos, existe una alta contribución de otros ecosistemas al microbioma de placas solares. Estos ecosistemas son los desiertos, y el polvo de éstos que es transportado por las corrientes de viento. El transporte de microorganismos resistentes a desecación, estrés oxidativo, y sobre todo, luz UV, hasta la superficie de las placas, hace que estos sean seleccionados en el nuevo ecosistema, aumentando su abundancia relativa y dando lugar a una comunidad adaptada a las condiciones extremas y más similar, en último término, a las comunidades presentes en los desiertos. De acuerdo con los análisis realizados, este proceso se produce en las dos localizaciones estudiadas, lo que sugeriría que el proceso de colonización de estas estructuras artificiales tiene un carácter universal.

Finalmente, las particularidades de las especies identificadas en este trabajo como características de las placas solares con respecto a los otros ecosistemas considerados (géneros *Hymenobacter*, *Deinococcus*, *Rubellomicrobium* y *Kaistobacter*) resaltan la importancia de las placas solares como fuentes de microorganismos de gran interés biotecnológico por su elevada resistencia a condiciones de estrés, así como a su capacidad de producir compuestos de alto valor añadido como carotenoides u otras sustancias antioxidantes.

4.4.3 Plan de trabajo futuro

Con el fin de entender en mayor grado el proceso de colonización de las placas solares, se está llevando a cabo un experimento enfocado a estudiar los cambios en las comunidades microbianas en los meses sucesivos a la instalación de unos paneles fotovoltaicos. De este modo, se espera identificar las especies pioneras y obtener información sobre cómo los eventos climáticos, tales como las lluvias o las tormentas de polvo, afectan al dinamismo de este microbioma.

Además, se están secuenciando muestras de placas solares de la Antártida y el Ártico, con el fin de realizar un nuevo análisis comparativo que permita confirmar la existencia de un *core* de especies todavía más robusto que el planteado en este trabajo, así como establecer el grado de universalidad de la colonización de placas solares a nivel global.

5. CONCLUSIONES

Teniendo en cuenta los resultados discutidos anteriormente, las conclusiones más importantes que se derivan de este trabajo son:

- Los perfiles taxonómicos de las comunidades microbianas de las placas solares de Valencia son más próximos a los de las comunidades presentes en el desierto del Sáhara o del polvo del Sáhara transportado por el aire, que a los perfiles de las comunidades presentes en el aire de regiones de clima mediterráneo no desértico.
- Las similitudes observadas entre las placas solares de Valencia y los ambientes relacionados con el desierto del Sáhara (suelo y polvo desértico) se producen tanto a nivel taxonómico como funcional.
- Los perfiles taxonómicos de las comunidades microbianas de las placas solares de Berkeley son más próximos a los de las muestras de los distintos desiertos de EE.UU., que a las procedentes de aire libre de polvo (procedente de Norteamérica o Asia) o aire con polvo desértico recogido en Asia.
- En los análisis de las placas solares de ambas localizaciones (Valencia y Berkeley), el microbioma demuestra no ser consecuencia de una simple deposición de los microorganismos presentes en el aire. Si bien el aire ejerce un papel importante como fuente de microorganismos pioneros del microbioma de placas solares, existen conexiones a nivel taxonómico que vinculan la biocenosis de las placas solares con las presentes en los suelos de distintos desiertos, así como con el polvo de los mismos transportado a través del aire.
- Existe un núcleo de microorganismos representativos del ecosistema de placas solares. Entre ellos destacan especies de los géneros *Hymenobacter*, y en menor medida, *Deinococcus*, *Kaistobacter* y *Rubellimicrobium*, caracterizados por ser productores de pigmentos y altamente resistentes a la desecación y la radiación ultravioleta. Estos géneros son detectables en los otros ecosistemas estudiados, pero su frecuencia es significativamente más alta en placas solares.
- Las biocenosis de las placas solares de Valencia y Berkeley han demostrado ser sorprendentemente similares. La elevada similitud taxonómica detectada en placas solares de localizaciones tan distantes, enfatiza la universalidad del proceso de colonización así como el papel de presiones selectivas comunes (radiación ultravioleta, desecación, etc) en la formación de este microbioma, muy poco estudiado hasta la fecha.

6. REFERENCIAS BIBLIOGRÁFICAS

ADAMS, R. I.; BHANGAR, S.; PASUT, W.; ARENS, E. A.; TAYLOR, J. W.; LINDOW, S. E.; NAZAROFF, W. W. & BRUNS, T. D. (2015). Chamber bioaerosol study: outdoor air and human occupants as sources of indoor airborne microbes. *PLoS One*, 10(5), e0128022.

ALTSCHUL, S. F.; GISH, W.; MILLER, W.; MYERS, E. W. & LIPMAN, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, 215(3), 403-410.

AMES, S. K.; HYSOM, D. A.; GARDNER, S. N.; LLOYD, G. S.; GOKHALE, M. B. & ALLEN, J. E. (2013). Scalable metagenomic taxonomy classification using a reference genome database. *Bioinformatics*, 29(18), 2253-2260.

AN, S.; COUTEAU, C.; LUO, F.; NEVEU, J. & DUBOW, M. S. (2013). Bacterial diversity of surface sand samples from the Gobi and Taklamaken deserts. *Microbial ecology*, 66(4), 850-860.

ANDREW, D. R.; FITAK, R. R.; MUNGUÍA-VEGA, A.; RACOLTA, A.; MARTINSON, V. G. & DONTSOVA, K. (2012). Abiotic factors shape microbial diversity in Sonoran Desert soils. *Applied and environmental microbiology*, 78(21), 7527-7537.

BARBERÁN, A., HENLEY, J., FIERER, N., & CASAMAYOR, E. O. (2014). Structure, inter-annual recurrence, and global-scale connectivity of airborne microbial communities. *Science of the Total Environment*, 487, 187-195.

BARBERÁN, A.; LADAU, J.; LEFF, J. W.; POLLARD, K. S.; MENNINGER, H. L.; DUNN, R. R. & FIERER, N. (2015). Continental-scale distributions of dust-associated bacteria and fungi. *Proceedings of the National Academy of Sciences*, 112(18), 5756-5761.

BEALS, E. W. (1984). Bray-Curtis ordination: an effective strategy for analysis of multivariate ecological data. *Advances in Ecological Research*, 14, 1-55.

BERTOLINI, V., GANDOLFI, I., AMBROSINI, R., BESTETTI, G., INNOCENTE, E., RAMPAZZO, G., & FRANZETTI, A. (2013). Temporal variability and effect of environmental variables on airborne bacterial communities in an urban area of Northern Italy. *Applied microbiology and biotechnology*, 97(14), 6561-6570.

BOWERS, R. M.; MCCUBBIN, I. B.; HALLAR, A. G. & FIERER, N. (2012). Seasonal variability in airborne bacterial communities at a high-elevation site. *Atmospheric Environment*, 50, 41-49.

BRAGG, L. & TYSON, G. W. (2014). Metagenomics using next-generation sequencing. *Environmental Microbiology: Methods and Protocols*, 183-201.

BROCK, T. D. (1961). Milestones in Microbiology. *Academic Medicine*, 36(7), 847.

BRYANT, D. & MOULTON, V. (2004). Neighbor-net: an agglomerative method for the construction of phylogenetic networks. *Molecular biology and evolution*, 21(2), 255-265.

CAPORASO, J. G.; KUCZYNSKI, J.; STOMBAUGH, J.; BITTINGER, K.; BUSHMAN, F. D.; COSTELLO, E. K.; FIERER, N.; PEÑA, A. G.; GOODRICH, J. K.; GORDON, J. I.; HUTTLEY, G. A.; KELLEY, S. T.; KNIGHTS, D.; KOENIG, J. E.; LEY, R. E.; LOZUPONE, C. A.; MCDONALD, D.; MUEGGE, B. D.; PIRRUNG, M.; REEDER, J.; SEVINSKY, J. R.; TURNBAUGH, P. J.; WALTERS, W. A.; WIDMANN, J.; YATSUNENKO, T.; ZANEVELD, J. & KNIGHT, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, 7(5), 335-336.

CARTER, K. C. (1987). Essays of Robert Koch. *Carter KC (transl.). New York: Greenwood Press.*

CHA, S.; SRINIVASAN, S.; JANG, J. H.; LEE, D.; LIM, S.; KIM, K. S.; JHEONG, W.; LEE, D.; PARK, E.; CHUNG, H.; CHOE, J.; KIM, M. K. & SEO, T. (2017). Metagenomic Analysis of Airborne Bacterial

- Community and Diversity in Seoul, Korea, during December 2014, Asian Dust Event. *PLOS ONE*, 12(1).
- DANIEL, R. (2005). The metagenomics of soil. *Nature reviews. Microbiology*, 3(6), 470.
- DARWIN, C.R. (1846). An account of the fine dust which often falls on vessels in the Atlantic Ocean. *Quart J Geol Soc London II*, 26–30.
- DE LA TOUR, C. B.; MATHIEU, M.; MEYER, L.; DUPAIGNE, P.; PASSOT, F.; SERVANT, P.; SOMMER, S.; LE CAM, E. & CONFALONIERI, F. (2017). In vivo and in vitro characterization of DdrC, a DNA damage response protein in *Deinococcus radiodurans* bacterium. *PloS one*, 12(5), e0177751.
- DENNER, E. B.; KOLARI, M.; HOORNSTRA, D.; TSITKO, I.; KÄMPFER, P.; BUSSE, H. J. & SALKINOJA-SALONEN, M. (2006). *Rubellimicrobium thermophilum* gen. nov., sp. nov., a red-pigmented, moderately thermophilic bacterium isolated from coloured slime deposits in paper machines. *International journal of systematic and evolutionary microbiology*, 56(6), 1355-1362.
- DESANTIS, T. Z.; HUGENHOLTZ, P.; LARSEN, N.; ROJAS, M.; BRODIE, E. L.; KELLER, K.; HUBER, T.; DAVELI, D.; HU, P. & 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.
- DORADO-MORALES, P.; VILANOVA, C.; PERETO, J.; CODOÑER, F. M.; RAMON, D. & PORCAR, M. (2016). A highly diverse, desert-like microbial biocenosis on solar panels in a Mediterranean city. *Scientific reports*, 6.
- EDGAR, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics*, 26(19), 2460-2461.
- FAVET, J.; LAPANJE, A.; GIONGO, A.; KENNEDY, S.; AUNG, Y.; & CATTANEO, A.; DAVIS-RICHARDSON, A. G., BROWN, C. T., KORT, R., BRUMSACK, H.; SCHNETGER, B.; CHAPPELL, A.; KROIJENGA, J.; BECK, A.; SCHWIBBERT, K.; MOHAMED, A. H.; KIRCHNER, T.; DE QUADROS, P. D.; TRIPLETT, E. W.; BROUGHTON, W. J. & GORBUSHINA, A. A. (2012). Microbial hitchhikers on intercontinental dust: catching a lift in Chad. *The ISME Journal*, 7(4), 850-867.
- FERRER, M.; GOLYSHINA, O.; BELOQUI, A., & GOLYSHIN, P. N. (2007). Mining enzymes from extreme environments. *Current opinion in microbiology*, 10(3), 207-214.
- FIERER, N.; LEFF, J. W.; ADAMS, B. J.; NIELSEN, U. N.; BATES, S. T.; LAUBER, C. L.; OWENS, S.; GILBERT, J. A.; WALL, D. H. & CAPORASO, J. G. (2012). Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, 109(52), 21390-21395.
- FRÖSLER, J.; PANITZ, C.; WINGENDER, J.; FLEMMING, H. C. & RETTBERG, P. (2017). Survival of *Deinococcus geothermalis* in Biofilms under Desiccation and Simulated Space and Martian Conditions. *Astrobiology*, 17(5), 431-447.
- GANDOLFI, I., BERTOLINI, V., BESTETTI, G., AMBROSINI, R., INNOCENTE, E., RAMPAZZO, G., PAPANICHI, M. Y FRANZETTI, A. (2015). Spatio-temporal variability of airborne bacterial communities and their correlation with particulate matter chemical composition across two urban areas. *Applied microbiology and biotechnology*, 99(11), 4867-4877.
- GILBERT, J. A.; JANSSON, J. K. & KNIGHT, R. (2014). The Earth Microbiome project: successes and aspirations. *BMC biology*, 12(1), 69.

- GORBUSHINA, A. A.; KORT, R.; SCHULTE, A.; LAZARUS, D.; SCHNETGER, B.; BRUMSACK, H. J.; BROUGHTON, W. J. & FAVET, J. (2007). Life in Darwin's dust: intercontinental transport and survival of microbes in the nineteenth century. *Environmental Microbiology*, 9(12), 2911-2922.
- GOWER, J. C. (2005). Principal coordinates analysis. *Wiley StatsRef: Statistics Reference Online*.
- GRIFFIN, D. W. (2007). Atmospheric movement of microorganisms in clouds of desert dust and implications for human health. *Clinical microbiology reviews*, 20(3), 459-477.
- GUTMAN, J.; KAUFMAN, Y.; KAWAHARA, K.; WALKER, S. L.; FREGER, V. & HERZBERG, M. (2014). Interactions of glycosphingolipids and lipopolysaccharides with silica and polyamide surfaces: adsorption and viscoelastic properties. *Biomacromolecules*, 15(6), 2128-2137.
- HANDELSMAN, J.; LILES, M., MANN, D.; RIESENFELD, C. & GOODMAN, R. M. (2002). Cloning the metagenome: Culture-independent access to the diversity and functions of the uncultivated microbial world. *Methods in Microbiology*, 33, 241-255.
- HUGENHOLTZ, P. & TYSON, G. W. (2008). Microbiology: metagenomics. *Nature*, 455(7212), 481-483.
- HUMAN MICROBIOME PROJECT CONSORTIUM. (2012). A framework for human microbiome research. *Nature*, 486(7402), 215-221.
- HUMAN MICROBIOME PROJECT CONSORTIUM. (2012). Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402), 207-214.
- HUSON, D. H. & MITRA, S. (2012). Introduction to the analysis of environmental sequences: metagenomics with MEGAN. *Evolutionary Genomics: Statistical and Computational Methods*, Volume 2, 415-429
- HUSSAIN, F.; KHAN, I. U.; HABIB, N.; XIAN, W. D.; HOZZEIN, W. N.; ZHANG, Z. D.; ZHI, X.Y. & LI, W. J. (2016). *Deinococcus saudiensis* sp. nov., isolated from desert. *International journal of systematic and evolutionary microbiology*, 66(12), 5106-5111.
- Infolibre (2017). *El polvo del Sáhara sitúa a Madrid como la gran ciudad más contaminada del mundo este jueves por partículas en suspensión*. Editor: Javier Martínez. Visto 17 de Mayo de 2017: http://www.infolibre.es/noticias/politica/2017/02/23/el_polvo_del_sahara_situa_madrid_como_ciudad_mas_contaminada_del_mundo_este_jueves_por_particulas_suspension_61608_1012.html
- KATRA, I.; AROTSKER, L.; KRASNOV, H.; ZARITSKY, A.; KUSHMARO, A. & BEN-DOV, E. (2014). Richness and diversity in dust stormborne biomes at the southeast mediterranean. *Scientific reports*, 4, 5265.
- KIM, D.; HOFSTAEDTER, C. E.; ZHAO, C.; MATTEI, L.; TANES, C.; CLARKE, E.; LAUDER, A.; SHERRILL-MIX, S.; CHEHOUD, C.; KELSEN, J.; CONRAD, M.; COLLMAN, R. G.; BALDASSANO, R.; BUSHMAN, F. D. & BITTINGER, K. (2017). Optimizing methods and dodging pitfalls in microbiome research. *Microbiome*, 5(1), 52.
- KIM, E. B.; KANG, M. S.; JOO, E. S.; JEON, S. H.; JEONG, S. W., LIM, S. Y., JUNG, H.Y., SRINIVASAN, S. & KIM, M. K. (2017). *Deinococcus ruber* sp. nov., a radiation-resistant bacterium isolated from soil. *International journal of systematic and evolutionary microbiology*, 67(1), 72-76.
- KIM, M. K.; JOO, E. S.; LEE, S. Y.; LEE, D. S.; SRINIVASAN, S. & JUNG, H. Y. (2016). Complete genome sequence of *Hymenobacter* sp. DG25B, a novel bacterium with gamma-radiation resistance isolated from soil in South Korea. *Journal of biotechnology*, 217, 98-99.

- KNIGHTS, D.; KUCZYNSKI, J.; CHARLSON, E. S.; ZANEVELD, J.; MOZER, M. C.; COLLMAN, R. G.; BUSHMAN, F. D.; KNIGHT, R. & KELLEY, S. T. (2011). Bayesian community-wide culture-independent microbial source tracking. *Nature methods*, 8(9), 761-763.
- KÖBERL, M.; MÜLLER, H.; RAMADAN, E. M. & BERG, G. (2011). Desert farming benefits from microbial potential in arid soils and promotes diversity and plant health. *PLoS One*, 6(9), e24452.
- KOEPEL, A. F. & WU, M. (2013). Surprisingly extensive mixed phylogenetic and ecological signals among bacterial Operational Taxonomic Units. *Nucleic acids research*, gkt241.
- KONSTANTINIDIS, K. T.; BRAFF, J.; KARL, D. M. & DELONG, E. F. (2009). Comparative metagenomic analysis of a microbial community residing at a depth of 4,000 meters at station ALOHA in the North Pacific subtropical gyre. *Applied and environmental microbiology*, 75(16), 5345-5355.
- LANGILLE, M. G. I.; ZANEVELD, J.; CAPORASO, J. G.; MCDONALD, D.; KNIGHTS, D.; REYES, J. A., CLEMENTE, J. C., BURKEPILE, D. E.; VEGA THURBER, R. L.; KNIGHT, R.; BEIKO, R. G. & HUTTENHOWER, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology*, 31(9), 814-821.
- LAUBER, C. L.; HAMADY, M.; KNIGHT, R. & FIERER, N. (2009). Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Applied and environmental microbiology*, 75(15), 5111-5120.
- LEE, J. J.; KANG, M. S.; JOO, E. S.; JUNG, H. Y. & KIM, M. K. (2016). *Hymenobacter sedentarius* sp. nov., isolated from a soil. *The Journal of Microbiology*, 54(4), 283.
- LESTER, E. D., SATOMI, M. & PONCE, A. (2007). Microflora of extreme arid Atacama Desert soils. *Soil Biology and Biochemistry*, 39(2), 704-708.
- LÓPEZ-LÓPEZ, O.; CERDÁN, M. E. & GONZÁLEZ-SISO, M. I. (2013). Hot spring metagenomics. *Life*, 3(2), 308-320.
- MARIZCURRENA, J. J.; MOREL, M. A.; BRAÑA, V.; MORALES, D.; MARTINEZ-LÓPEZ, W. & CASTRO-SOWINSKI, S. (2017). Searching for novel photolyases in UVC-resistant Antarctic bacteria. *Extremophiles*, 21(2), 409-418.
- MAZAR, Y.; CYTRYN, E.; EREL, Y. & RUDICH, Y. (2016). Effect of dust storms on the atmospheric microbiome in the Eastern Mediterranean. *Environmental science & technology*, 50(8), 4194-4202.
- MEADOW, J. F.; ALTRICHTER, A. E.; KEMBEL, S. W.; KLINE, J.; MHUIREACH, G.; MORIYAMA, M.; NORTHCUTT, D.; O'CONNO, T. K.; WOMACK, A. M.; BROWN, G. Z. & GREEN, J. L. (2014). Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source. *Indoor Air*, 24(1), 41-48.
- MEOLA, M., LAZZARO, A., & ZEYER, J. (2015). Bacterial composition and survival on Sahara dust particles transported to the European Alps. *Frontiers in microbiology*, 6.
- NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA) (2015). *NASA Satellite Reveals How Much Saharan Dust Feeds Amazon's Plants*, Greenbelt, Maryland. Editor: Rob Garner. Visto el 15 de Mayo de 2017. <https://www.nasa.gov/content/goddard/nasa-satellite-reveals-how-much-saharan-dust-feeds-amazon-s-plants>
- NEILSON, J. W.; QUADE, J.; ORTIZ, M.; NELSON, W. M.; LEGATZKI, A.; TIAN, F.; LACOMB, M.; BETANCOURT, J. L.; WING, R. A.; SODERLUND, R. A. & MAIER, R. M. (2012). Life at the hyperarid margin: novel bacterial diversity in arid soils of the Atacama Desert, Chile. *Extremophiles*, 16(3), 553-566.

- OH, T. J.; HAN, S. R.; AHN, D. H.; PARK, H. & KIM, A. Y. (2016). Complete genome sequence of *Hymenobacter* sp. strain PAMC26554, an ionizing radiation-resistant bacterium isolated from an Antarctic lichen. *Journal of biotechnology*, 227, 19-20.
- ÓRDENES-AENISHANSLINS, N.; ANZIANI-OSTUNI, G.; VARGAS-REYES, M.; ALARCÓN, J.; TELLO, A. & PÉREZ-DONOSO, J. M. (2016). Pigments from UV-resistant Antarctic bacteria as photosensitizers in Dye Sensitized Solar Cells. *Journal of Photochemistry and Photobiology B: Biology*, 162, 707-714.
- OULAS, A.; PAVLOUDI, C.; POLYMENAKOU, P.; PAVLOPOULOS, G. A.; PAPANIKOLAOU, N.; KOTOULAS, G.; ARVANITIDIS, C. & ILIOPOULOS, I. (2015). Metagenomics: tools and insights for analyzing next-generation sequencing data derived from biodiversity studies. *Bioinformatics and biology insights*, 9, 75.
- PARKS, D. H.; TYSON, G. W.; HUGENHOLTZ, P. & BEIKO, R. G. (2014). STAMP: statistical analysis of taxonomic and functional profiles. *Bioinformatics*, 30(21), 3123-3124.
- PAULINO-LIMA, I. G.; FUJISHIMA, K.; NAVARRETE, J. U.; GALANTE, D.; RODRIGUES, F.; AZUA-BUSTOS, A. & ROTHSCCHILD, L. J. (2016). Extremely high UV-C radiation resistant microorganisms from desert environments with different manganese concentrations. *Journal of Photochemistry and Photobiology B: Biology*, 163, 327-336.
- QUAST, C.; PRUESSE, E.; YILMAZ, P.; GERKEN, J.; SCHWEER, T.; YARZA, P.; PEPLIES, J. & GLÖCKNER, F. O. (2012). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic acids research*, gks1219.
- RANAWAT, P. & RAWAT, S. (2017). Radiation resistance in thermophiles: mechanisms and applications. *World Journal of Microbiology and Biotechnology*, 33(6), 112.
- RIDEOUT, J. R.; HE, Y.; NAVAS-MOLINA, J. A.; WALTERS, W. A.; URSELL, L. K.; GIBBONS, S. M.; CHASE, J.; MCDONALD, D.; GONZALEZ, A.; ROBBINS-PIANKA, A.; CLEMENTE, J. C.; GILBERT, J. A.; HUSE, S. M.; ZHOU, H.; KNIGHT, R. & CAPORASO, J. G. (2014). Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences. *PeerJ*, 2, e545.
- ROSSELLI, R.; FIAMMA, M.; DELIGIOS, M.; PINTUS, G.; PELLIZZARO, G.; CANU, A.; DUCE, P.; SQUARTINI, A.; MURESU, R. & CAPPUCCINELLI, P. (2015). Microbial immigration across the Mediterranean via airborne dust. *Scientific Reports*, 5(1).
- SALVUCCI, E. (2016). Microbiome, holobiont and the net of life. *Critical reviews in microbiology*, 42(3), 485-494.
- SAVAGE, A. M.; HILLS, J.; DRISCOLL, K.; FERGUS, D. J.; GRUNDEN, A. M. & DUNN, R. R. (2016). Microbial diversity of extreme habitats in human homes. *PeerJ*, 4, e2376.
- SCHOLZ, M. B.; LO, C. C. & CHAIN, P. S. (2012). Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. *Current opinion in biotechnology*, 23(1), 9-15.
- SOGIN, M. L.; MORRISON, H. G.; HUBER, J. A.; WELCH, D. M.; HUSE, S. M.; NEAL, P. R.; ARRIETA, J. M. & HERNDL, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". *Proceedings of the National Academy of Sciences*, 103(32), 12115-12120.
- STACKEBRANDT, E. & GOEBEL, B. M. (1994). Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology. *International Journal of Systematic and Evolutionary Microbiology*, 44(4), 846-849.

- STEVEN, B.; GALLEGOS-GRAVES, L. V.; BELNAP, J. & KUSKE, C. R. (2013). Dryland soil microbial communities display spatial biogeographic patterns associated with soil depth and soil parent material. *FEMS microbiology ecology*, 86(1), 101-113.
- TEMPERTON, B., & GIOVANNONI, S. J. (2012). Metagenomics: microbial diversity through a scratched lens. *Current opinion in microbiology*, 15(5), 605-612.
- THOMAS, T.; GILBERT, J & MEYER, F. (2012). Metagenomics-a guide from sampling to data analysis. *Microbial informatics and experimentation*, 2(1), 3.
- TIAN, B. & HUA, Y. (2010). Carotenoid biosynthesis in extremophilic *Deinococcus-Thermus* bacteria. *Trends in microbiology*, 18(11), 512-520.
- TRIADÓ-MARGARIT, X.; VEILLETTE, M.; DUCHAINE, C.; TALBOT, M.; AMATO, F.; MINGUILLÓN, M. C.; MARTINS, V.; DE MIGUEL, E.; CASAMAYOR, E. O. & MORENO, T. (2016). Bioaerosols in the Barcelona subway system. *Indoor Air*.
- TURNBAUGH, P. J.; LEY, R. E.; HAMADY, M.; FRASER-LIGGETT, C.; KNIGHT, R. & GORDON, J. I. (2007). The human microbiome project: exploring the microbial part of ourselves in a changing world. *Nature*, 449(7164), 804.
- VARGA, G. (2012). Spatio-temporal distribution of dust storms—a global coverage using NASA TOMS aerosol measurements. *Hungar. Geograph. Bull*, 61, 275-298.
- VILANOVA, C. (2016). From genes to microbial consortia: multi-level bioprospecting of natural and artificial environments. Tesis Doctoral en Biotecnología. Universitat de València. 13 pp.
- VILANOVA, C.; IGLESIAS, A. & PORCAR, M. (2015). The coffee-machine bacteriome: biodiversity and colonisation of the wasted coffee tray leach. *Scientific reports*, 5.
- WILLIAMSON, A. J.; MORRIS, K.; LAW, G. T.; RIZOULIS, A.; CHARNOCK, J. M. & LLOYD, J. R. (2014). Microbial reduction of U (VI) under alkaline conditions: implications for radioactive waste geodisposal. *Environmental science & technology*, 48(22), 13549-13556.
- WOOLEY, J. C., & YE, Y. (2010). Metagenomics: facts and artifacts, and computational challenges. *Journal of computer science and technology*, 25(1), 71-81.
- YOO, K.; LEE, T. K.; CHOI, E. J.; YANG, J.; SHUKLA, S. K.; HWANG, S. I. & PARK, J. (2017). Molecular approaches for the detection and monitoring of microbial communities in bioaerosols: A review. *Journal of Environmental Sciences*, 51, 234-247.
- YU, L. Z. H.; LUO, X. S.; LIU, M. & HUANG, Q. (2015). Diversity of ionizing radiation-resistant bacteria obtained from the Taklimakan Desert. *Journal of basic microbiology*, 55(1), 135-140.

7. ANEXO

Tabla 7.1. Set de datos metagenómicos de las muestras seleccionadas en la búsqueda bibliográfica preliminar. WGS = *whole genome shotgun*; 16S = Secuenciación del gen 16S rRNA

Tipo de muestra	Localización	Referencia (URL)	Estrategia	Tecnología
Placas Solares	España	https://goo.gl/mltf89	16S/WGS	IonTorrent
Placas Solares	EE.UU	-	WGS	Illumina
Suelo La Gran Cuenca	EE.UU	https://goo.gl/NNFNqU	WGS	-
Suelo La Gran Cuenca	EE.UU	https://goo.gl/RyKbSF	16S	454
Suelo Sonora	EE.UU	https://goo.gl/Er8y3f	16S	454
Suelo Mojave	EE.UU	https://goo.gl/OK5Fkg	16S	Illumina
Suelo Mojave	EE.UU	https://goo.gl/1ZLss6	16S	454
Suelo Chihuahua	EE.UU	https://goo.gl/OK5Fkg	16S	Illumina
Suelo Sáhara	Chad	https://goo.gl/fk4yFT	16S	Illumina
Suelo Sáhara	Mauritania	https://goo.gl/MHpSLM	16S	454
Suelo Gobi	China	https://goo.gl/XV6Ujh	16S	454
Suelo Takla Makan	Mongolia	https://goo.gl/XV6Ujh	16S	454
Aire Control	EE.UU	https://goo.gl/3FM2Ar	WGS	Illumina
Aire Control	EE.UU	https://goo.gl/3XjU5D	16S	454
Aire Control	EE.UU	https://goo.gl/OHHu8S	16S	454
Polvo Desértico	-	https://goo.gl/x1a0e	WGS	Illumina
Polvo Desértico	España	https://goo.gl/MHpSLM	16S	454
Polvo Desértico	Israel	https://goo.gl/VxSTij	16S	Illumina
Polvo Desértico	Suiza	https://goo.gl/dnfNT5	16S	Illumina
Polvo Desértico	Israel	https://goo.gl/ncR4AO	16S	454
Polvo Desértico	Italia	https://goo.gl/LuONqb	16S	Illumina
Polvo Desértico	China/ Corea	https://goo.gl/0BJe7Y	16S	454
Polvo Desértico	Corea del Sur	https://goo.gl/Eg2pIX	16S	454
Polvo Agrícola	EE.UU	https://goo.gl/twqlgT	16S	454
Aire Urbano	Italia	https://goo.gl/092CvW	16S	454
Aire Urbano	Italia	https://goo.gl/CYU2GH	16S	Illumina
Aire Urbano	China	https://goo.gl/rCVqqn	WGS	Illumina
Aire Urbano	Italia	https://goo.gl/HKF6sM	16S	Illumina
Aire Urbano	EE.UU	https://goo.gl/iaDnzh	WGS	Illumina
Polvo Agrícola y Suburbano	EE.UU	https://goo.gl/a9iidf	16S	454
Filosfera	EE.UU	https://goo.gl/nQ4yKK	16S	454

Tabla 7.2. Muestra de los 15 taxones más abundantes en placas solares de Valencia y su abundancia relativa en cada muestra. Vlc = Valencia; Bk = Berkeley.

Taxón	Abundancia relativa (%)					
	vlc.1	vlc.2	vlc.3	bk.1	bk.2	bk.3
f__Cytophagaceae;g__Hymenobacter	40.0	11.8	11.0	18.6	12.5	22.9
f__Sphingomonadaceae;g__Kaistobacter	11.2	11.1	7.9	0.0	0.0	0.0
f__Deinococcaceae;g__Deinococcus	8.9	11.1	37.8	0.7	2.4	0.5
f__Geodermatophilaceae;g__Modestobacter	4.0	0.2	0.2	0.4	0.5	0.7
f__Acetobacteraceae;g__	3.3	4.0	4.8	0.1	0.1	0.0
f__Erythrobacteraceae;g__	3.0	8.0	1.8	0.1	0.0	0.0
f__Sphingomonadaceae;g__Sphingomonas	3.0	2.5	1.9	1.4	0.8	0.7
f__Oxalobacteraceae;g__	2.8	1.3	1.8	0.1	0.1	0.1
f__Rhodobacteraceae;g__Rubellimicrobium	2.5	7.6	4.8	0.0	0.0	0.0
f__Sphingomonadaceae;g__	2.4	4.3	4.4	0.9	0.5	0.4
f__Chitinophagaceae;g__Flavisolibacter	1.6	8.5	5.5	0.0	0.0	0.0
f__Sphingomonadaceae;g__Novosphingobium	0.7	2.4	0.5	0.1	0.1	0.1
f__Sporichthyaceae;g__	0.6	0.6	0.6	0.0	0.0	0.0
f__Kineosporiaceae;g__	0.6	0.1	0.2	0.4	0.5	0.3
f__Oxalobacteraceae;g__Janthinobacterium	0.5	1.4	0.7	0.0	0.0	0.0

Tabla 7.3. Muestra de los 15 taxones más abundantes en placas solares de Berkeley y su abundancia relativa en cada muestra.

Taxón	Abundancia relativa (%)					
	vlc.1	vlc.2	vlc.3	bk.1	bk.2	bk.3
f__Cytophagaceae;g__Hymenobacter	40.0	11.8	11.0	18.6	12.5	22.9
f__Methylobacteriaceae;g__Methylobacterium	0.0	0.2	0.1	3.4	4.6	2.7
f__Sphingomonadaceae;g__Sphingomonas	3.0	2.5	1.9	1.4	0.8	0.7
f__Methylobacteriaceae;g__	0.0	0.0	0.0	1.3	1.7	1.0
f__Kineosporiaceae;g__Kineococcus	0.1	0.0	0.0	1.2	1.5	0.9
f__Streptomycetaceae;g__Streptomyces	0.0	0.0	0.0	1.0	1.4	0.9
f__Deinococcaceae;g__Deinococcus	8.9	11.1	37.8	0.7	2.4	0.5
f__Nocardioideae;g__Friedmanniella	0.1	0.1	0.1	0.7	0.6	0.9
f__Pseudomonadaceae;g__Pseudomonas	0.0	0.0	0.0	0.6	0.8	0.6
f__Nocardioideae;g__	0.3	0.3	0.3	0.5	0.5	0.5
f__Streptomycetaceae;g__	0.0	0.0	0.0	0.4	0.6	0.4
f__Geodermatophilaceae;g__Modestobacter	4.0	0.2	0.2	0.4	0.5	0.7
f__Kineosporiaceae;g__	0.6	0.1	0.2	0.4	0.5	0.3
f__Burkholderiaceae;g__Burkholderia	0.0	0.0	0.0	0.4	0.4	0.3
f__Mycobacteriaceae;g__Mycobacterium	0.0	0.0	0.0	0.3	0.4	0.3

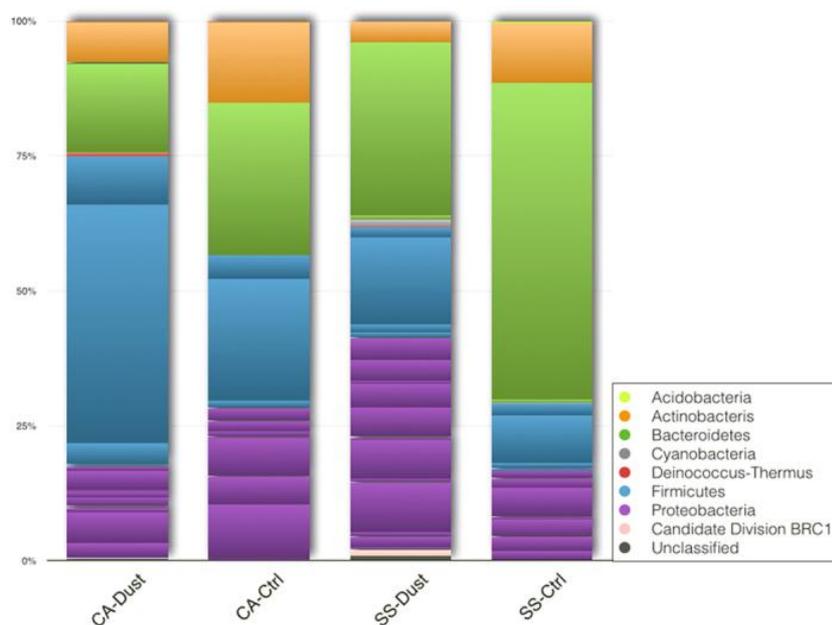


Figura 7.1. Perfil taxonómico (porcentaje de secuencias asignables a bacterias identificables en bases de datos) procedente de muestras de Cerdeña en días libres de polvo (2, 4) y días de tormenta de polvo (1,3). *Extraído de Roselli et al., 2015.*

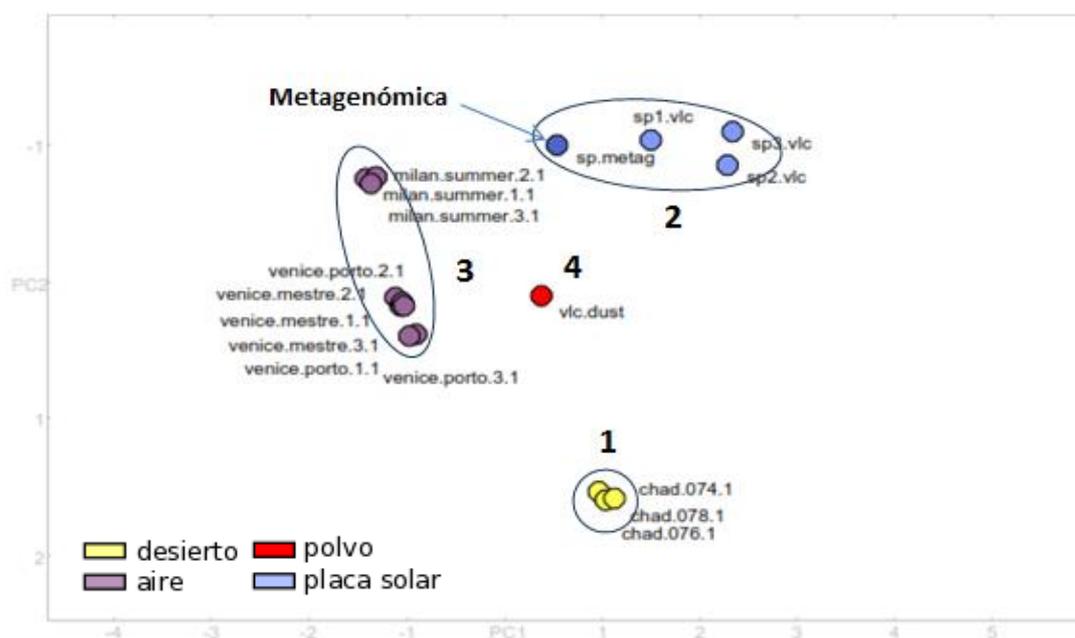


Figura 7.2. Análisis de Coordenadas Principales (PC1 (33,6%) vs PC2 (23,1%)) realizado a partir de perfiles taxonómicos (máximo nivel de asignación taxonómica: género). Se incorporó al análisis una muestra de placas solares de Valencia secuenciada mediante una estrategia metagenómica. Se observan cuatro grupos constituidos por: **1.** Suelo del Sáhara (Chad); **2.** Placas solares de Valencia; **3.** Aire urbano; **4.** Polvo del Sáhara recogido en Valencia.