



CREACIÓN DE LIBRERÍAS MEDIANTE AMPLIFICACIÓN DEL ARN RIBOSOMAL 16S PARA LA CARACTERIZACIÓN DE LA MICROBIOTA

LIBRARY GENERATION THROUGH RIBOSOMAL RNA 16S GENE AMPLIFICATION FOR DETERMINING MICROBIOTA COMPOSITION

Autores

Emilio J. Laserna Mendieta
Marcus J. Claesson

Centro

School of Microbiology and
APC Microbiome Institute,
University College Cork, Cork,
Ireland

Fecha de publicación

29 diciembre 2016

Páginas

Páginas 9-13

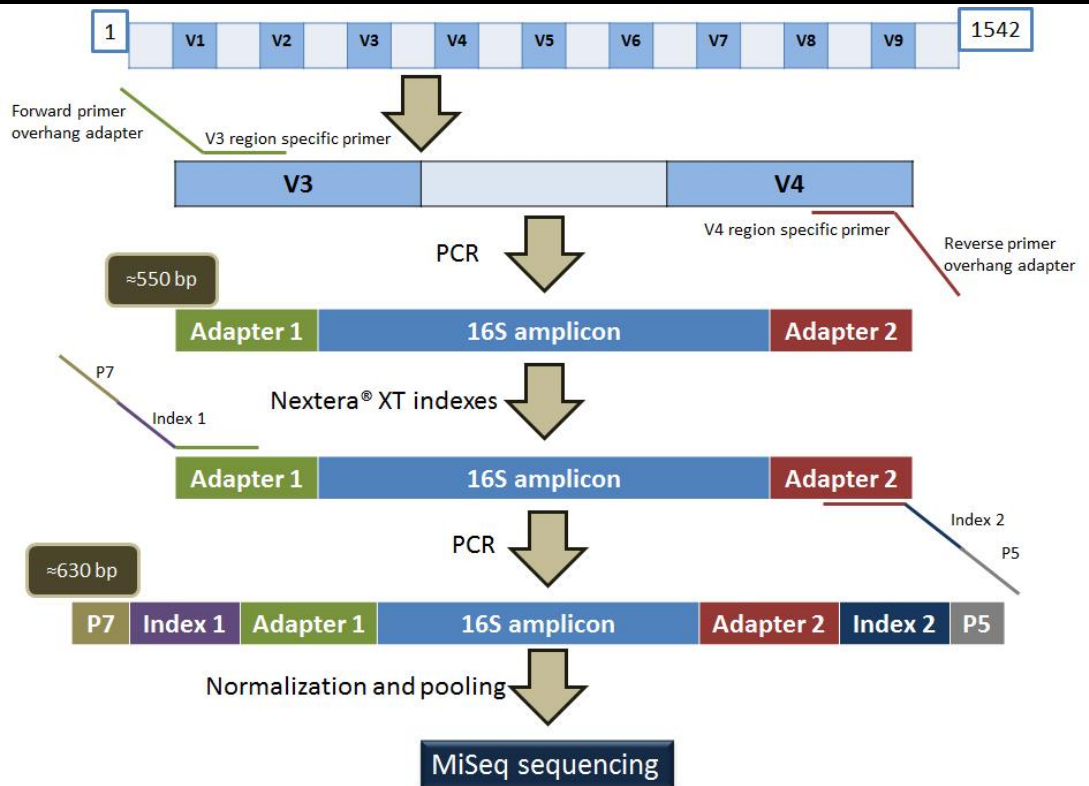


Figura 1. Esquema del proceso seguido para la creación de librerías mediante amplificación del gen del ARNr 16S. La región amplificada comprende a las regiones variables V3 y V4 y la región conservada que se encuentra entre ellas. Una primera PCR se realiza con los primers específicos V3-V4 que contienen los adaptadores para los índices Nextera XT. En una segunda PCR, se añaden los índices para introducir una combinación única para cada muestra y las secuencias P5 y P7 (que hibridan con los adaptadores de la célula de flujo del sistema MiSeq). Tras normalizar la cantidad de cada muestra que se añade al pool final, las librerías están listas para ser secuenciadas mediante MiSeq. **Figure 1.** Summary of the approach followed for the generation of 16S rRNA gene libraries. The amplified region comprised variable regions V3 and V4 and the conserved region between them. A first PCR was performed with specific primers for V3-V4 that contain adaptors for Nextera XT indexes. In a second PCR, a different combination of indexes for each sample and P5/P7 sequences (which hybridise with flow cell adaptor in the MiSeq system) were introduced. After normalization of the quantity of each sample that was included in the final pool, libraries were ready to be sequenced in the MiSeq system.

Una gran cantidad de estudios metagenómicos, cuyo objetivo es la determinación de la composición bacteriana de un determinado tipo de muestra, se han realizado mediante el análisis por secuenciación masiva de librerías resultantes de la amplificación por PCR del gen procariota del ARN ribosómico (ARNr) 16S. Este gen posee una longitud de 1542 pares de bases y está formado por nueve regiones variables entre las que se intercalan regiones conservadas. Son precisamente los fragmentos resultantes de la amplificación de una o dos regiones variables del ARNr16S las que mayoritariamente se han empleado en la caracterización filogenética de las diversas especies bacterianas que constituyen la microbiota humana¹.

Una de las fuentes principales de variabilidad de este tipo de estudios, aparte del método de extracción del ADN y de la plataforma de secuenciación, lo constituye la elección de los *primers* empleados en la amplificación del ARNr 16S². La evaluación *in silico* que Klindworth y colaboradores³ llevaron a cabo sobre este aspecto, demostró que una pareja de *primers* degenerados (es decir, que tienen en varias posiciones dos o más bases nucleotídicas) para las regiones V3-V4 mostraba los resultados más prometedores, lo que ha extendido su uso en posteriores trabajos (casi 300 citas a finales de septiembre de 2016).

En nuestro laboratorio hemos venido realizando diversos proyectos encaminados a la caracterización de la microbiota presente en muestras humanas de heces y biopsias intestinales, de modo que el ADN es extraído y posteriormente amplificado en dos PCR consecutivas (Figura 2). El objetivo es la construcción de librerías que contienen los amplicones del gen del ARNr16S de diversas muestras, protocolo que se esquematiza en la Figura 1. Dicho protocolo toma como referencia el método

Many metagenomic studies are aimed at determining the bacterial composition in a specific type of sample. This is usually performed by next-generation sequencing analysis of the data obtained from libraries created by PCR amplification of the prokaryotic 16S ribosomal RNA (rRNA) gene. This gene has a length of 1542 base pairs and it is composed by nine variable regions and interspaced by more conserved regions. The amplification of one or two of these variable regions of the 16S rRNA gene is a common approach employed in the phylogenetic identification of hundreds of species that are present in the human microbiota¹.

One of the main sources of variation in these sorts of studies, apart from the DNA extraction method and the sequencing platform, is the choice of primers used in 16S rRNA gene amplification². Klindworth and co-authors³ carried out an *in silico* evaluation about this issue and described a couple of degenerate primers (that is, with two or more nucleotide bases in several positions) for regions V3-V4 as those showing the most promising results. As a result, these primers have been widely employed in later studies (almost 300 cites at the end of September 2016).

In our laboratory, we have been analysing microbiota composition across multiple studies from hundreds of stool and biopsy human samples. Here, DNA was extracted from those samples and employed for two consecutive PCR amplifications (Figure 2). In this way and following the protocol summarized in Figure 1, we have created libraries that contain amplicons of the 16S rRNA gene from several samples. This approach was based in the method recommended by Illumina⁴, although some modifications were introduced.

recomendado por Illumina⁴ aunque introduciendo algunos cambios.

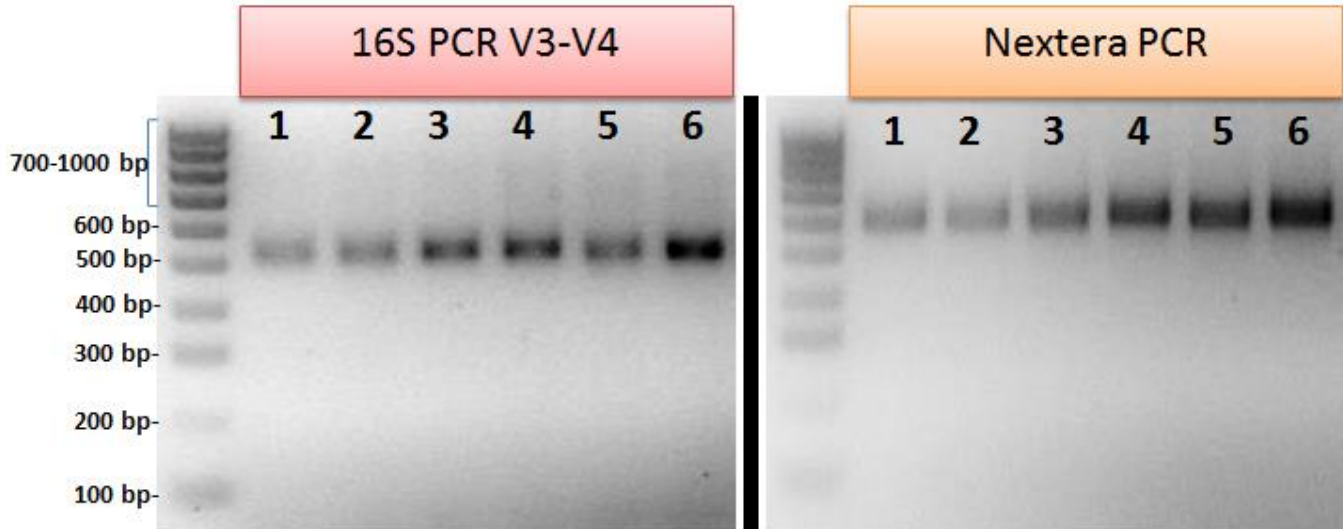


Figura 2. En la parte izquierda se puede observar el resultado de la PCR de amplificación del gen del ARNr16S en seis muestras de ADN extraído de heces humanas. Se distingue claramente una banda específica de tamaño entre 500 y 600 pares de bases. En la parte derecha, se puede visualizar el resultado de la PCR empleando los índices Nextera XT, de modo que ahora los amplicones poseen un tamaño superior a los 600 pares de bases. En ambos casos las muestras se corrieron en un gel de agarosa al 2% teñido con *Midori Green* (*NipponGenetics*). **Figure 2.** On the left, it is shown the PCR result for 16S rRNA gene amplification in six DNA samples from human faeces. Clear and specific bands are observed in the size range of 500-600 base pairs. On the right, it is shown the PCR result employing Nextera XT indexes, thus now amplicons have a size higher than 600 base pairs. In both cases, samples were run in a 2% agarose gel stained with Midori Green (*NipponGenetics*).

El primer paso consiste en una PCR con los primers V3-V4 a los que se les añade la secuencia adaptadora para los índices Nextera XT de Illumina. Así, la secuencia completa (nomenclatura IUPAC) de los *primers* empleados es:

Directo (*Forward*):

5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGCAG 3'

Reverso (*Reverse*):

5' GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATCTAATCC 3'

The first step consisted in a PCR with V3-V4 primers where an adaptor sequence for Nextera XT indexes (Illumina) was added to. Thus, the complete sequence for these primers was (IUPAC nomenclature):

Esta reacción de amplificación se realiza usando una polimerasa de alta fidelidad (en nuestro caso, *Phusion High-Fidelity DNA polymerase* suministrada por ThermoScientific). Las condiciones para la PCR son: 98°C durante 30 s, seguido de 25 ciclos de 98°C durante 10 s, 55°C durante 15 s y 72°C durante 20 s, con un ciclo final de 72°C durante 5 min. Los productos de la PCR son purificados mediante bolitas magnéticas (*AgencourtAMPure XP beads*, suministradas por Beckman-Coulter). A continuación, se usan 5 µL del ADN purificado para la segunda PCR con los primers de los índices Nextera XT. El programa para esta PCR es el mismo que para la PCR del 16S, con la salvedad de que sólo se realizan 8 ciclos de amplificación. Con esta PCR, se consigue que para cada muestra los amplicones 16S posean una combinación de índices única que la diferencia del resto. Tras realizar una nueva purificación del ADN resultante, se procede a la cuantificación de su concentración, para lo cual se emplean métodos de alta sensibilidad para el ADN de doble cadena como el *Quant-iTPicogreen* o *Qubit* (suministrados por ThermoScientific). Finalmente, se prepara un *pool*, es decir, se mezclan todas las muestras en un único tubo de modo que la cantidad de ADN sea la misma para todas ellas, y dicho *pool* se diluye a una concentración adecuada para secuenciación mediante el sistema MiSeq de Illumina.

Los resultados de la secuenciación han de ser procesados mediante las correspondientes herramientas bioinformáticas para obtener la composición bacteriana de cada muestra. Esta metodología ha sido empleada con éxito en el estudio sobre la comparación de microbiota en personas tratadas a largo plazo con inhibidores de la bomba de protones e individuos control, que mostró cambios en la microbiota intestinal (como el ratio

This PCR reaction was performed using a high-fidelity DNA polymerase (in our case, *Phusion High-Fidelity DNA polymerase*, purchased from Thermo Scientific). PCR conditions were: 98°C for 30 s, followed by 25 cycles of 98°C for 10 s, 55°C for 15 s and 72°C for 20 s, with a final extension step of 72°C for 5 min. PCR products were purified using magnetic beads (*Agencourt AMPure XP beads*, purchased from Beckman-Coulter). Then, 5 µL of the purified DNA were used in the second PCR that employed primers from Nextera XT indexes. The program for this second PCR was the same set up for 16S rRNA PCR, only changing the amplification cycles from 25 to 8. After this PCR, we achieved 16S amplicons that had a different and unique combination of indexes for every single sample. After a new purification of the second PCR products, DNA quantification is performed employing high sensitivity methods for double stranded DNA, like *Quant-iTPicogreen* or *Qubit* (both from Thermo Scientific). Finally, we pooled all samples, which should have the same final concentration, by mixing them in one single tube. This pool was afterwards diluted to achieve an adequate concentration to be sequenced in the MiSeq system (Illumina).

Sequencing results have to be processed using properly selected bioinformatic pipelines to get the bacterial composition of each sample. This methodology was successfully employed in a study carried out in our laboratory about stool microbiota comparison between people in long-term treatment with proton pump inhibitors and control individuals. The analysis of 16S rRNA gene amplicons recently revealed changes in the colonic microbiota, as the Firmicutes/Bacteroides ratio, associated to the long-term prescription of these types of drugs⁵.

entre los filos Firmicutes/Bacteroides) asociados al uso de estos fármacos⁵.

Bibliografía/References:

1. Cox MJ, Cookson WO, Moffatt MF. Sequencing the human microbiome in health and disease. *Hum Mol Genet.* 2013;22(R1):R88-94.
<http://hmg.oxfordjournals.org/content/22/R1/R88.long>
2. Fouhy F, Clooney AG, Stanton C, Claesson MJ, Cotter PD. 16S rRNA gene sequencing of mock microbial populations- impact of DNA extraction method, primer choice and sequencing platform. *BMC Microbiol.* 2016;16(1):123.
<https://bmcmicrobiol.biomedcentral.com/articles/10.1186/s12866-016-0738-z>
3. Klindworth A, Pruesse E, Schweer T, Peplies J, Quast C, Horn M, et al. Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 2013;41(1):e1.
<http://nar.oxfordjournals.org/content/41/1/e1.long>
4. http://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf
5. Clooney AG, Bernstein CN, Leslie WD, Vagianos K, Sargent M, Laserna-Mendieta EJ, et al. A comparison of the gut microbiome between long-term users and non-users of proton pump inhibitors. *Aliment Pharmacol Ther.* 2016;43(9):974-84.
<http://onlinelibrary.wiley.com/doi/10.1111/apt.13568/abstract>