International Journal of Food Microbiology 294 (2019) 1–9

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Compositional assessment of bacterial communities in probiotic supplements by means of metagenomic techniques

Gabriele Andrea Lugli[a,1], Marta Mangifesta[a,b,1], Leonardo Mancabelli[a], Christian Milani[a], Francesca Turroni[a,c], Alice Viappiani[b], Douwe van Sinderen[d], Marco Ventura[a,c,*]

[a] Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, Parma, Italy
[b] GenProbio srl, Parma, Italy
[c] Microbiome Research Hub, University of Parma, Parma, Italy
[d] APC Microbiome Institute and School of Microbiology, Bioscience Institute, National University of Ireland, Cork, Ireland

ARTICLE INFO

Keywords:
Genomics
Metagenomics
Probiotics
ITS
Next generation sequencing

ABSTRACT

Health promoting or probiotic bacteria are commonly incorporated into a variety of functional foods and drug formulations, due to their purported ability to confer benefit to host health. Despite the extensive commercial exploitation of probiotic formulations there are still major knowledge gaps regarding the precise molecular mechanism of action and corresponding genetic/genomic properties of probiotic bacteria. In the current study, we describe a metagenomic approach which allows determination of the composition of probiotic supplements through next-generation sequencing analyses based on rRNA-associated sequences to assess bacterial composition of the product combined with a shotgun metagenomics approach directed to decode the genome sequences of the probiotic strains for each product assayed. The here developed approach has been tested for 10 probiotic supplements, revealing inconsistencies between the identified probiotic strains and the declared strains as indicated by the producers. Furthermore, the decoded bacterial genome sequence of *Bifidobacterium animalis* subsp. *lactis* BB-12 from a 1995 frozen dried stock revealed genetic evidence for genome evolution and stability of this microorganism when compared with the re-constructed genome of the identical strain from a probiotic supplement of 2017.

1. Introduction

Probiotic products are defined as "live microorganisms which when administered in adequate amounts confer a health benefit on the host" by the Food and Agriculture Organization of the United Nations and the WHO (FAO/WHO) (Hill et al., 2014). In order for a probiotic microorganism to be included in the Qualified Presumption of Safety (QPS) list, by means of a generic risk assessment applied by the European Food Safety Authority (EFSA), it should not carry any transmissible antibiotic resistance genes, not produce harmful toxins, not evoke any undesirable side effects on the host, nor induce unwanted (e.g. pro-inflammatory) immune responses (Doron and Snydman, 2015; Sanders et al., 2014). Probiotics are supplemented in various formulations, both conventional pharmaceutical products, which include tablets, capsules and powders, as well as non-conventional foods, e.g., cheeses, yogurts, creams and milk (Govender et al., 2014; Sanders et al., 2014). Notably, the safety of a "pharmaceutical" probiotic strain should be demonstrated by means of phase 1 clinical trial, in order to validate its safety in both pathological and healthy conditions (Besselink et al., 2008). The currently used bacterial taxa incorporated in such formulations encompass members of the *Bacillus*, *Bifidobacterium*, *Enterococcus*, *Escherichia* and *Lactobacillus* genera (Govender et al., 2014). Such microorganisms had to fulfill strict genetic requirements, such as an unambiguous taxonomical identification and a precise evaluation of their genetic content (Donelli et al., 2013; Fontana et al., 2013; Johnson and Klaenhammer, 2014).

Recently, a genetic survey of probiotic products commercialized in the USA highlighted that the determined microbial composition of just four of the 13 products tested were in accordance with the strains declared by the producers (Drago et al., 2010). A similar investigation, aimed to explore the number of viable bacteria in probiotic products from the European market, unveiled that 10 out of the 24 examined

* Corresponding author at: Laboratory of Probiogenomics, Department of Chemistry, Life Sciences, and Environmental Sustainability, University of Parma, Parco Area delle Scienze 11a, 43124 Parma, Italy.
E-mail address: marco.ventura@unipr.it (M. Ventura).
[1] These authors contributed equally.

https://doi.org/10.1016/j.ijfoodmicro.2019.01.011
Received 18 July 2018; Received in revised form 31 October 2018; Accepted 20 January 2019
Available online 21 January 2019
0168-1605/

products did not contain the indicated number of viable bacteria, while in four of such products any viable bacteria of at least one of the declared species could be retrieved (Toscano et al., 2013). However, the large part of these analyses are based on culture-dependent approaches aimed to isolate and enumerate the microorganisms contained in the probiotic products by the use of selective media (Drago et al., 2010; Goldstein et al., 2014; Toscano et al., 2013). Conversely, metagenomic approaches have been performed to evaluate the bacterial relative abundance within the probiotic products by sequencing of the 16S rRNA sequence using Next-Generation Sequencing (NGS), and thus unveiling bacterial DNA contaminations (Morovic et al., 2016; Patro et al., 2016). Another crucial aspect of probiotic products that has pretty much been ignored is the genetic stability of the incorporated microorganisms. In fact, it is very well known that bacterial genomes evolve through the accumulation of changes (e.g. point mutations, horizontal gene transfer events, reshuffling of their genetic content, gene decay events) (Ventura et al., 2007). Such phenomena may happen at very low frequency under natural conditions but in industrial bacterial strains these events may occur at much higher rates as a consequence of repetitive cultivation under very specific and possible highly selective conditions. In this context, it has been shown that the continuous cultivation of *Bifidobacterium longum* NCC2705 for 1,000 generations in laboratory, lead to a massive reshaping of its chromosome characterized by two large deletions (Lee et al., 2008). In addition, comparative genome analyses of the industrial *Corynebacterium glutamicum* ATCC 13032 strain obtained from two different groups revealed intriguing genetic differences represented by a putative prophage island and three insertion sequences (Ventura et al., 2007). In a similar fashion, the genome instability of *Lactobacillus rhamnosus* GG has been reported, identifying a missing DNA segment from the genome constituting a deletion of 34 and 84 genes from two isolates (Sybesma et al., 2013).

In the current study, we describe a metagenomic approach for the identification of probiotic strains that have been incorporated into probiotic supplements. This methodology, referred to here as Genetic Identity Card (GIC), is able to reveal the microbiota composition of probiotic supplements starting from different probiotic formulations. A combination of NGS analyses that target specific rRNA sequences for compositional analysis and shotgun metagenomics was employed to assess the relative abundance and the overall genetic repertoire of the probiotic strains. Notably, the GIC approach, when applied to 10 probiotic supplements, showed several inconsistencies between identified probiotic strains and strains declared by the producers.

Furthermore, we evaluated the genetic stability of one of the most intensively used probiotic strains, *Bifidobacterium animalis* subsp. *lactis* BB-12, by comparative genomic analyses involving the genome sequences of this strain reconstructed from the shotgun metagenomics data and the chromosome sequences that were decoded from an ancestral stock of the identical strain.

2. Results and discussion

2.1. Taxonomical composition of probiotic supplements

The taxonomical composition of 10 probiotic products (Table S1) was investigated through 16S rRNA-based profiling (Fig. 1). Illumina-mediated 16S rRNA microbial profiling produced a total of 750,124 sequencing reads with an average of filtered reads of 60,119 per sample (Table S2). At genus level, the analysis showed that probiotic products B, C, D, E, F, G, I and J reflected the bacterial composition declared by the producer, while probiotic products A and H revealed the presence of additional bacteria. In detail, probiotic products A and H appeared to contain bacteria belonging to the *Lactobacillus* and *Streptococcus* genera, which suggest possible bacterial contamination (Fig. 2).

In order to further validate the possibility of bacterial contamination as identified by the first step of the GIC pipeline based on 16S rRNA sequencing, we performed PCR reactions on products A and H using primers specific for the *Lactobacillus* and *Streptococcus* genera (Moura et al., 2007; Picard et al., 2004). The PCR results of replicates based on three different products packaged individually for each probiotic formulation confirmed the occurrence of *Lactobacillus* and *Streptococcus* (DNA contamination) of the probiotic products A and H, respectively. These results confirmed DNA contamination in probiotic formulation during the production, excluding contaminations during the analysis.

2.2. Prediction of bifidobacterial and fungal composition

In order to characterize the bifidobacterial population present in the probiotic products A, C, D, E, F, G and H, we used the recently developed bifidobacterial ITS profiling protocol (Milani et al., 2014) (Fig. 1). Moreover, the indicated presence of *Saccharomyces cerevisiae* in the product I was verified by amplification and sequencing of fungal ITS as previously described (Bokulich and Mills, 2013) (Fig. 1). Quality filtering of the sequenced ITS produced an average of 7496 and 13,928 high-quality and full-length reads per sample, respectively (Table S2). The bifidobacterial analysis showed that all probiotic products had the microbial composition as declared by the producer except for probiotic product H that contained, in addition to the expected *Bifidobacterium longum* subsp. *longum* taxon, *Bifidobacterium animalis* subsp. *lactis* (Fig. 2). Fungal analysis of probiotic product I revealed the presence of *Saccharomyces cerevisiae*, thereby validating the fungal composition as outlined on the package leaflet.

2.3. Culturomics efforts directed to reconstruct bacterial contents of probiotic products

In order to confirm the obtained in silico metagenomic results, we performed a census of the bacterial populations of all assayed probiotics products by the use of a culturomics approach in which we attempted to isolate the bacterial strains from the assessed probiotic products on selective media (Table 1). These analyses were coupled with a taxonomical identification of the isolates through amplification of their 16S rRNA gene sequences followed by amplicon sequencing.

Such tests confirmed the bacterial composition identified with the metagenomics approaches described above with just a few deviations (see below) from what would have been expected based on the associated product declarations. In fact, the culture-dependent approach allowed the identification of isolates that belong to *Lactobacillus* genus in product A, and *Streptococcus* and *Lactobacillus* genera in the commercial probiotic formulation H. Furthermore, we isolated bacterial colonies belonging to *B. animalis* subsp. *lactis* and *Lb. paracasei* species in products H and J, respectively. These findings therefore confirmed the results obtained by metagenomic analyses, clearly indicating the occurrence of bacterial strains in product A, H and J that were not indicated in the product information.

2.4. Reconstruction of the probiotic genomes

In order to investigate the genomic repertoire of the strains identified in the metagenomic profiling and culturomics efforts, we subjected each probiotic product to shotgun metagenomics analysis (Fig. 1). The sequencing output ranged from 1.49 to 6.36 million of paired-end reads, based on the quantity of DNA loaded in the sequencer estimated by means of the number of probiotic strains declared by the producer of each sample, for a total of 33.89 million. Thus, the generated data from the last step of the GIC pipeline allowed us to perform a metagenomic assembly. Consequently, each reconstructed genome portion from each metagenomic assembly was taxonomically classified at species level. Accordingly, genome reconstruction from the metagenomics data allowed us to identify chromosomal sequences of 23 bacterial/yeast strains, including *B. animalis* subsp. *lactis*, *Bifidobacterium bifidum*, *B. longum* subsp. *longum*, *Enterococcus faecium*, *Lactobacillus acidophilus*,

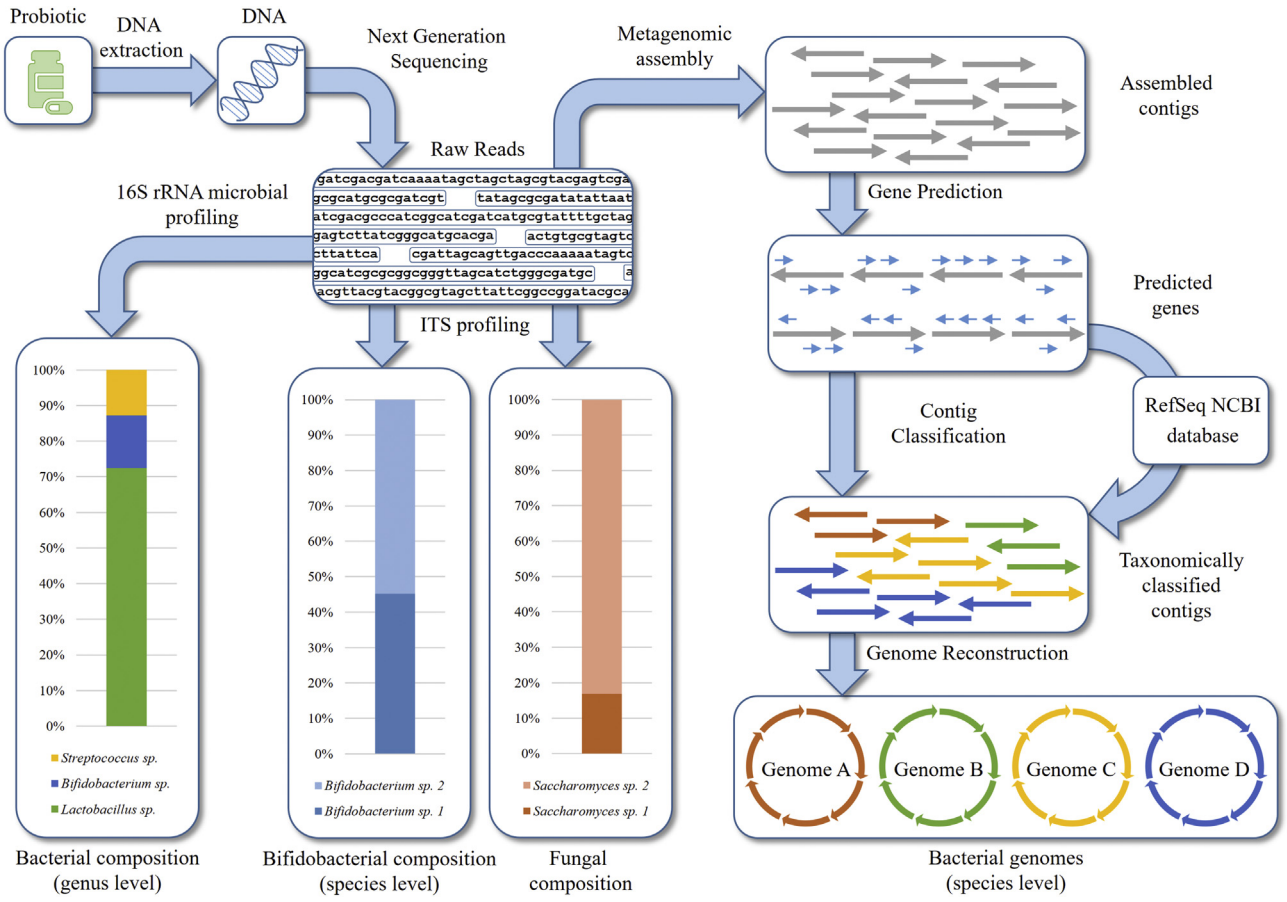

Fig. 1. Schematic representation of the metagenomic approach named Genetic Identity Card. The methodology involves three analyses based on NGS, i.e., from left to right, 16S rRNA microbial profiling, ITS profiling and assembly of shotgun metagenomic data.

Lactobacillus delbrueckii subsp. *bulgaricus*, *Lactobacillus helveticus*, *Lactobacillus paracasei*, *Lactobacillus plantarum*, *Lactobacillus rhamnosus*, *Streptococcus thermophilus* and a yeast strain belonging to the *S. cerevisiae* species (Fig. 3).

The performed genome reconstruction of probiotic products A, B, C, E, F, G and I disclosed the declared bacterial composition in the product information (Fig. 3 and Table S1). However, the contamination identified in probiotic supplement A through the 16S rRNA profiling and culturomics analyses was not detected in the metagenome analysis, probably due to the low abundance of these strains within the samples. In fact, the number of reads needed for the reconstruction of a bacterial genome must be sufficiently high in order to obtain a minimum genome coverage for the read assembly. Based on the assembled genomes of the analyzed products, we retrieved genome coverages ranging from 3.2 of *B. animalis* subsp. *lactis* (product A) to 84.2 of *Lb. helveticus* (product F) (Table S1). Thus, the success in genome reconstruction is correlated with genome length and abundance of the strain in the product. The same issue also affected metagenomic data of probiotic products D and J, where we were unsuccessful in identifying chromosomal sequences of *Lb. rhamnosus* and *Lactobacillus fermentum*, respectively. Notably, both of these two strains were acknowledged to be present by the producers (Fig. 3 and Table S1). Therefore, these results suggest that the latter strains are present at very low abundance with respect to other bacterial strains that were present in sample D and J, or were essentially absent. To validate this notion, qPCR analysis was performed on serial dilutions starting from 1 g of probiotic powder from products D and J in order to identify these apparently missing microbes, as well as on products A, D, G and H so as to compare the abundance of a number of genomes that we were able to re-construct, i.e., *B. animalis* subsp. *lactis*, *B. bifidum* and *S. thermophilus*. The normalized results show the presence of *Lb. rhamnosus* in product D at an abundance that was $< 10^5$ colonies forming unit (CFU), while *Lb. fermentum* was not detected confirming its presumed absence in product J. Furthermore, qPCR analysis of the *Streptococcus* strain present in product H, for which we could not reconstruct the genome, displayed a similar CFU value to that of *Lb. rhamnosus* in sample D (7.87×10^4 CFU). In contrast, *B. animalis* subsp. *lactis*, *B. bifidum* and *S. thermophilus* present in products A, D and G exhibited higher CFUs within the products, i.e., 1.5×10^6, 5.89×10^8 and 3.19×10^5 CFU. Thus, as demonstrated by qPCR analyses, where the relative abundance of a strain does not exceed 10^5 CFU, genome reconstruction could not be accomplished.

The GIC pipeline also allowed us to identify various bacterial contaminations or undeclared microbes in probiotic products. In this context, metagenomics analysis of probiotic supplements H and J allowed the reconstruction of *B. animalis* subsp. *lactis* and *Lb. paracasei* genomes, respectively (Fig. 3). Thus, a high level of undeclared or contaminant bacteria was present in these two probiotic supplements, suggesting that proper controls in the supply chain of production should be performed to insure the probiotic strain content of these products.

Furthermore, genome reconstruction of strains present in probiotic products as obtained through the GIC pipeline will permit comparative genome analyses with sequences from other strains belonging to the same species. This application may disclose unique genes of the probiotic strains with respect to publicly available genomes of the same species, perhaps unveiling probiotic characteristics of such microorganisms.

2.5. Genome stability and evolution

During evolution, bacterial genomes are subjected to genetic events

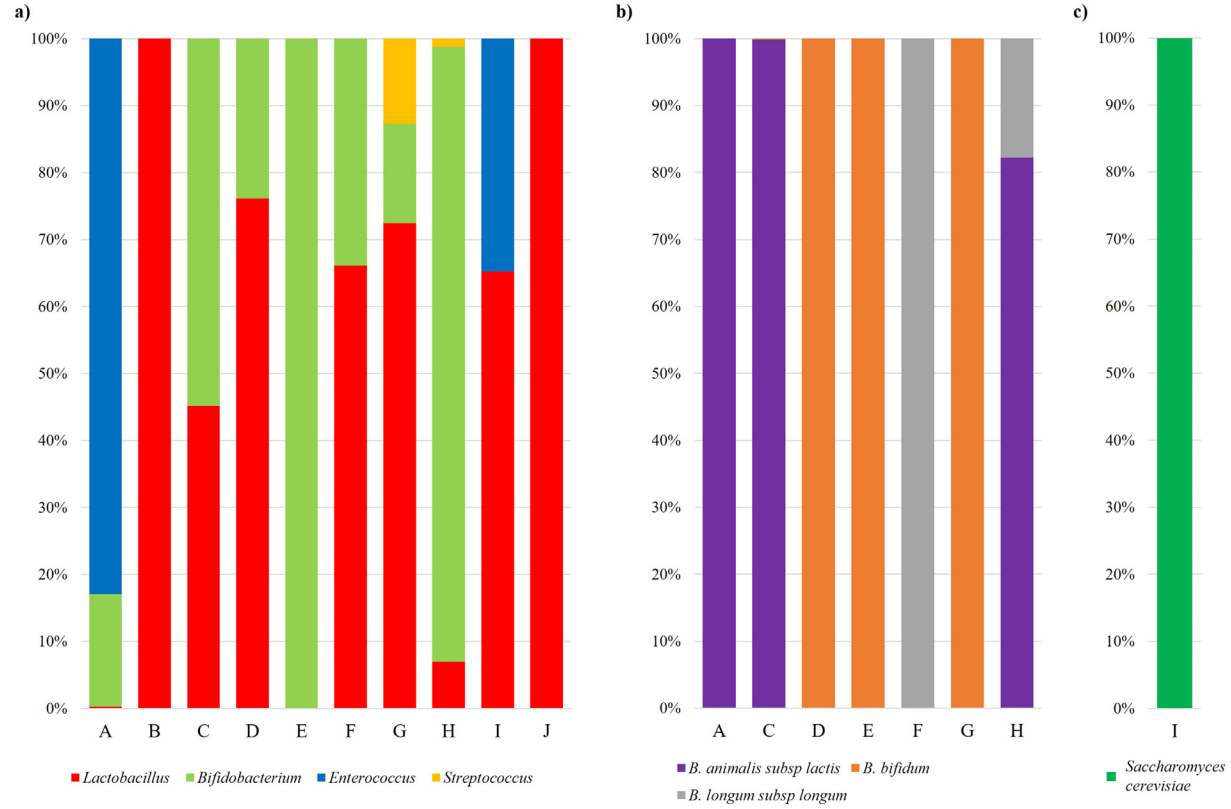

Fig. 2. The 16S rRNA gene and ITS profiling of probiotic products. Panel a shows a bar plot of each probiotic sample representing the OTUs obtained through 16S rRNA gene profiling. Panel b displays a bar plot for each probiotic sample that include bifidobacterial strains in the formulations representing the results achieved through ITS bifidobacterial profiling. Panel c exhibits the fungal ITS profiling of sample I.

shaping genome sequences, i.e., gene duplication, gene loss, chromosomal rearrangements and horizontal gene transfer (Magadum et al., 2013; Soucy et al., 2015). These events, which may occur at different rates, are responsible for bacterial speciation and are driven by the competitive environment in which these bacterial cells reside (Koonin, 2015). Among bacteria, industrial strains used as started cultures as well as probiotic bacteria appear to be subjected to rapid genome evolution processes, due to the continuous cultivation of these

Table 1
Culturomics analyses.

Probiotics	Medium	Anaerobic condition	Aerobic condition	Identified 16S ribosomal RNA sequences
A	MRS agar	37 °C for 48 h	–	*B. animalis* subsp. *lactis*
	Brain heart infusion	–	37 °C for 48 h	*E. faecium*
	MRS agar	–	37 °C for 48 h	*Lactobacillus* spp.
B	MRS agar	–	37 °C for 48 h	*Lb. paracasei*
C	MRS agar	37 °C for 48 h	–	*B. animalis* subsp. *lactis*
	MRS agar	37 °C for 48 h	–	*Lb. acidophilus*
D	MRS agar	37 °C for 48 h	–	*B. bifidum*
	MRS agar	37 °C for 48 h	–	*Lb. acidophilus*
E	MRS agar	37 °C for 48 h	–	*B. bifidum*
F	MRS agar	37 °C for 48 h	–	*B. longum*
	MRS agar	–	37 °C for 48 h	*Lb. helveticus*
G	MRS agar	37 °C for 48 h	–	*B. bifidum*
	MRS agar	37 °C for 48 h	–	*Lb. acidophilus*
	MRS agar	37 °C for 48 h	–	*Lb. rhamnosus*
	Brain heart infusion	–	37 °C for 48 h	*S. thermophilus*
	MRS agar	37 °C for 48 h	–	*Lb. delbrueckii*
H	MRS agar	–	37 °C for 48 h	*Lactobacillus* spp.
	MRS agar	37 °C for 48 h	–	*B. animalis* subsp. *lactis*
	MRS agar	37 °C for 48 h	–	*B. longum* subsp. *longum*
	Brain heart infusion	–	37 °C for 48 h	*Streptococcus* spp.
I	Brain heart infusion	–	37 °C for 48 h	*E. faecium*
	YPD agar	–	30 °C for 48 h	*S. cerevisiae*
	MRS agar	37 °C for 48 h	–	*Lb. acidophilus*
J	MRS agar	37 °C for 48 h	–	*Lb. acidophilus*
	MRS agar	–	37 °C for 48 h	*Lb. paracasei*
	MRS agar	37 °C for 48 h	–	*Lb. plantarum*
	MRS agar	37 °C for 48 h	–	*Lb. rhamnosus*

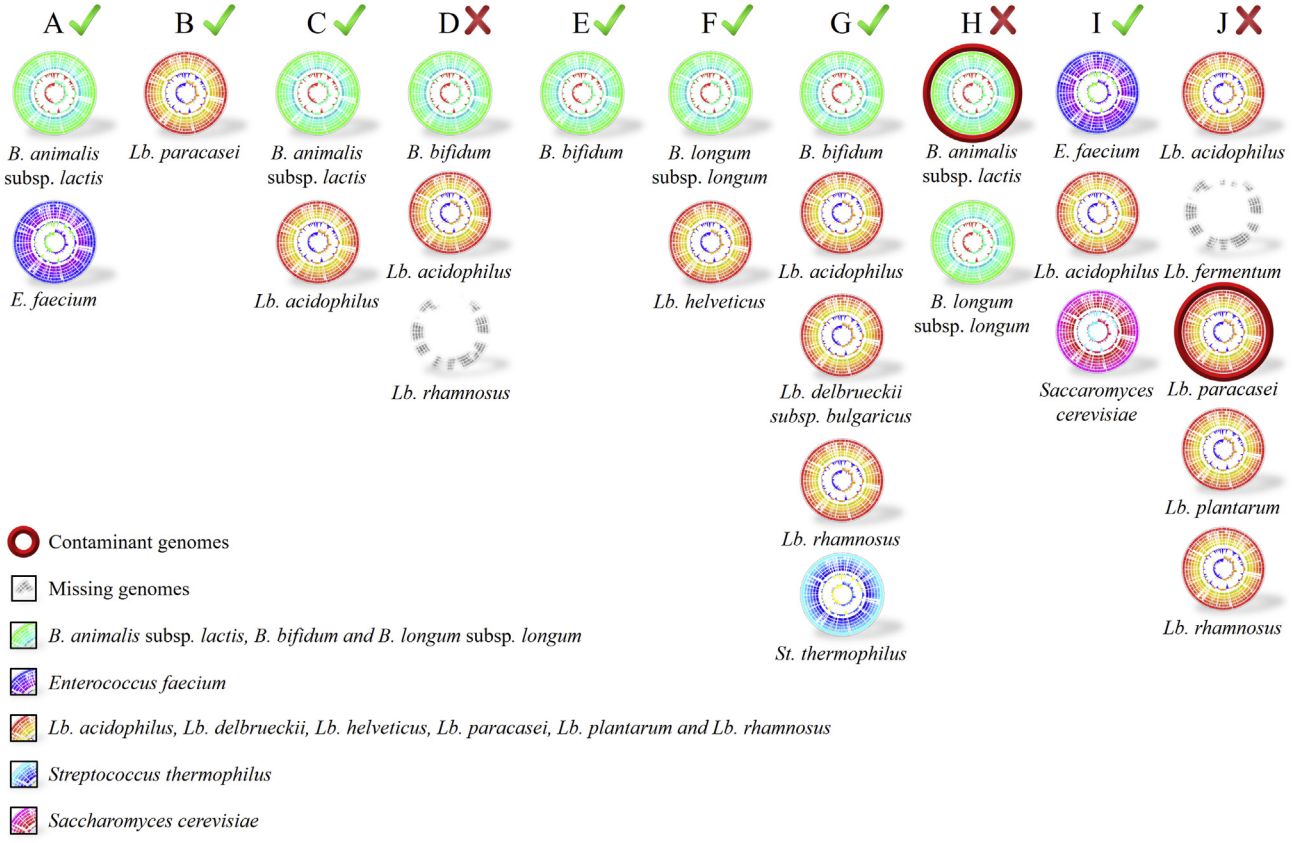

Fig. 3. Re-constructed genomes from probiotic products by means of the GIC metagenomic pipeline. Each re-constructed genome is represented by a genetic map arranged in vertical position by means of the 10 probiotic supplements. Apparently missing genomes were illustrated as fading genomes, while undeclared bacteria genomes are highlighted with red circles. Green check mark indicates probiotic products that showed the re-construction of the strains declared by the producer, while red letter x reveal products with undeclared or contaminant strains. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 2
SNPs between re-constructed and reference genomes.

Probiotics	Re-constructed genomes	Reference strain	SNPs	Intergenic SNPs	Coding SNPs
C	*B. animalis* subsp. *lactis*	BB-12	418	227	191
H	*B. longum* subsp. *longum*	CECT7347	184	10	174
I	*Lb. acidophilus*	La-14	9	4	5
J	*Lb. acidophilus*	La-14	5	4	1

microorganisms (Lee et al., 2008; Sybesma et al., 2013). Usually, companies maintain their products as “master cultures” in order to reduce the genome variations. However, specific growth conditions may significantly influence the genetic stability of the strains used (Machielsen et al., 2010).

Public database contain complete genomes of *B. animalis* subsp. *lactis* BB-12 (present in product C), *B. longum* subsp. *longum* CECT7347 (present in product H) and *Lb. acidophilus* La-14 (occurring in products I and J). The reconstructed genomes as obtained by the GIC pipeline were aligned with previously sequenced genomes of these commercially exploited probiotic strains (Fig. S1). The genome comparisons resulted in the identification of several single nucleotide polymorphisms (SNPs) characterizing the genome evolution of these strains, ranging from five to 418 between sample J and C, respectively (Table 2). Notably, SNPs placed at the very extreme border of the contigs were excluded from our analyses in order to report consistent polymorphisms between sequences. Thus, the evaluation of the SNPs that affect coding sequences highlighted that the published *Lb. acidophilus* La-14 genome sequence when compared to the re-constructed genomes of samples I and J showed just five and one SNPs, respectively (Table 2). Similarly, *B. animalis* subsp. *lactis* BB-12 and *B. longum* subsp. *longum* CECT7347 genomes, when compared with the re-constructed genomes of samples C and H, identified 191 and 174 SNPs, respectively (Table 2). These data disclosed a high frequency of SNPs occurrence in the *Bifidobacterium* species in respect to *Lb. acidophilus* genomes. Nonetheless, the high frequency in SNPs occurrence may be correlated with the quality of the genome sequences deposited. In fact, while *Lb. acidophilus* was sequenced at a 149-fold coverage (Stahl and Barrangou, 2013), *B. animalis* subsp. *lactis* BB-12 and *B. longum* subsp. *longum* CECT7347 genomes were sequenced at 11-fold coverage with Sanger methodology and 32-fold coverage using a 454 Roche sequencer, respectively (Chenoll et al., 2013; Garrigues et al., 2010).

Further insights related to evolution of probiotic strains was provided by the decoding of genome sequences of a freeze dried stock of *B. animalis* subsp. *lactis* BB-12 (strain named as BB-12-1995), which was produced in 1995 and had since then been maintained in a freezer. The genome sequence of this “ancestral” BB-12 strain was compared with the previous sequenced *B. animalis* subsp. *lactis* BB-12 strain (strain named as BB-12-2010), whose genome was decoded in 2010. In the same fashion, the chromosome sequences belonging to *B. animalis*

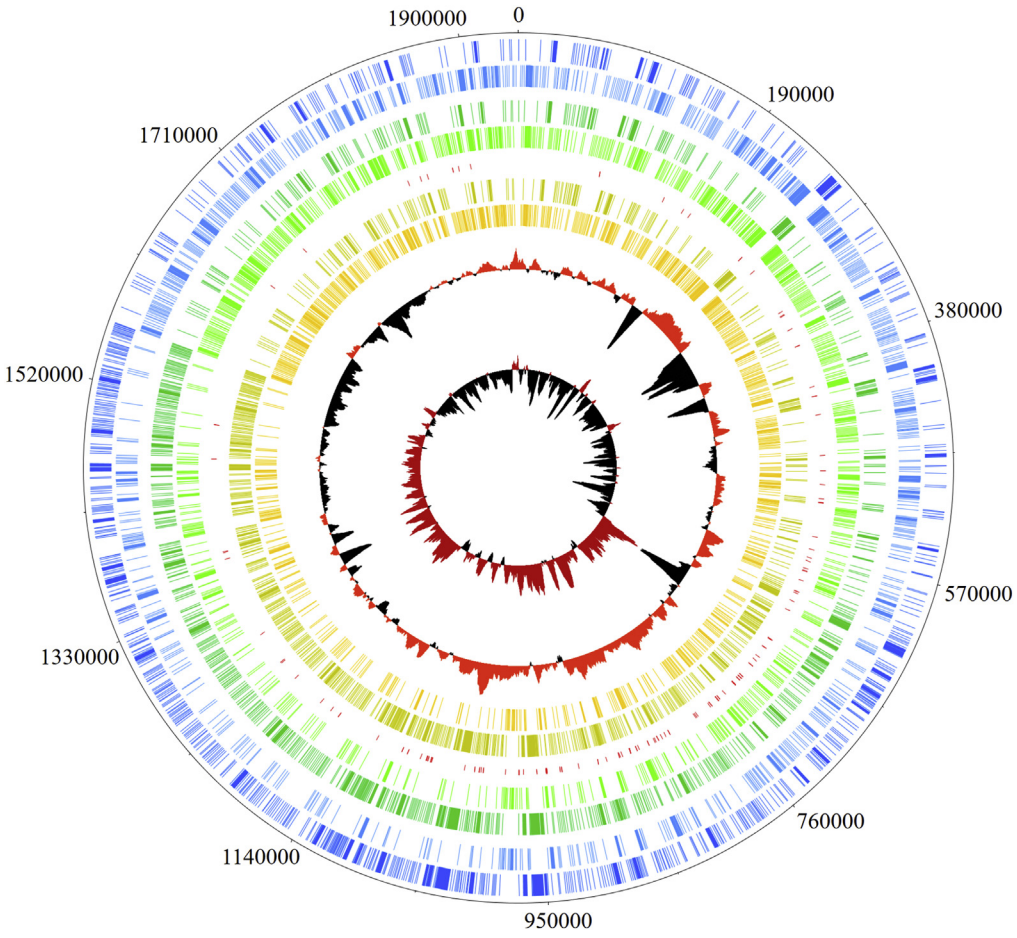

Fig. 4. Genome evolution of strain *B. animalis* subsp. *lactis* BB-12. Circular genome atlas of *B. animalis* subsp. *lactis* BB-12-1995, BB-12-2010 and BB-12-2017 with SNP positions between genomes represented by red lines. Starting from the external circle, BB-12-2010 gene locations (blue circles), BB-12-1995 gene positions (green circles), SNPs spot between BB-12-1995 and BB-12-2017 (red lines), BB-12-2017 gene locations (yellow circles), BB-12 G + C% deviation (black/orange) and BB-12 GC skew (G − C/G + C) (black/red) are reported. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

subsp. *lactis*, which were reconstructed from the metagenomic data of probiotic C (strain named as BB-12-2017), thus constituting a 'modern' version of *B. animalis* subsp. *lactis* BB-12 strain was compared to BB-12-1995. Such analysis revealed that strain BB-12-1995 has been moderately modified during this time span by accumulating several point mutations within the genome sequence without showing any deletion (Fig. 4). The genome of BB-12-1995 was assembled at 249-fold coverage, resulting in 161 SNPs when compared with BB-12-2017 genome, of which 148 SNPs reside in open reading frames (ORFs) (Fig. 4 and Table S3). SNPs identified at the very edge of contigs were discarded, obtaining an average of SNP frequency of 99.4% between the identified nucleotide variations (Table S3). Notably, the 148 SNPs that affect coding regions, involved 89 ORFs of which 80 ends in nonsynonymous substitutions that change the codon that encodes for a different amino acid (Table S3). Therefore, this analysis clearly demonstrates that the genome of *B. animalis* subsp. *lactis* BB-12 has been subjected to modification in the course of the evolution of the last 22 years.

3. Conclusions

Health promoting or probiotic bacteria are commonly incorporated into a variety of probiotic supplements and dairy products, due to their purported ability to confer benefit to host health (Govender et al., 2014; Sanders et al., 2014). Metagenomic based analyses performed on commercialized probiotic products reveal inconsistency in the bacterial composition, highlighting the need for a consistent strategy to perform quality checks during the production of the probiotic supplements (Morovic et al., 2016; Patro et al., 2016). Thus, we have established a metagenomic approach named Genetic Identity Card, which characterizes the microbial content of probiotic products by means of a metagenomics approach based on rRNA-associated sequences and shotgun metagenomics. Validation of this analysis was performed using 10 different supplements, revealing inconsistency of the determined bacterial content of four probiotic formulations on the one hand and the microbial content as declared by the producers on the other. Genome comparison with previously sequenced strains used in the supplements allowed us to investigate genome stability of these strains. Remarkably, comparative genomic analyses based on the genome of the probiotic *B. animalis* subsp. *lactis* BB-12 strain collected in 1995 and the current version on the market revealed 162 SNPs that may have been introduced and selected during cultivation of this commercial strain.

Thus, the GIC pipeline is a collection of metagenomic methodologies aimed to identify and characterize the bacterial composition as well as the genetic features, of probiotic products including both single- and multiple strain formulations, without the need of a direct isolation of strains from these probiotic supplements. In the close future, the GIC approach may be implemented by the enumeration of the bacterial cells using a flow cytometric approach that will allow a normalization of the reads for each bacterial strain determined by metagenomics analyses (Chiron et al., 2018; Vandeputte et al., 2017).

4. Materials and methods

4.1. Probiotics included in this study

We randomly selected 10 powder-based probiotic supplements in order to analyze products containing either a single-strain or multi-strains, which were retrieved from supermarkets and processed through the GIC pipeline before the expiration date indicated in the products' information. We recoded the commercial name of each probiotic product to keep the anonymity of the product. The complete amount of lyophilizate contained in a daily dose for each probiotic product, was

collected and mixed in Phosphate Buffered Saline (Sigma-Aldrich, St Louis MO) solution in order to obtain a homogeneous representation of each sample. Therefore, a portion of the mixed lyophilizate, consisting of 200 mg of product, was obtained from each probiotic. DNA was extracted from each probiotic using DNeasy PowerFood Microbial Kit (Qiagen Ltd, Strasse, Germany) following the manufacturer's instructions (Qiagen Ltd, Strasse, Germany).

4.2. Isolation of probiotic from probiotic supplements

One milliliter of each sample was mixed with 9 mL sterile Phosphate-buffered saline (PBS) to make an initial dilution. Serial dilutions were made for each sample and 1 mL of the appropriate dilution was mixed with different media agar plates. For the isolation of probiotics, different media were assessed, including de Man, Rogosa and Sharpe (MRS) agar (Difco Laboratories, USA) supplemented with 0.05% (wt/vol) L-cysteine hydrochloride and 50 μg/mL mupirocin (Delchimica, Italy), M17 agar (Oxoid Ltd., UK) with modification (2% lactose after autoclaving) and YPD (Sigma-Aldrich, St Louis MO). The M17 agar plates were incubated at 37 °C for 48 h in aerobic condition, MRS agar plates were incubated in anaerobic atmosphere (2.99% H_2, 17.01% CO_2, and 80% N_2) in a chamber (Concept 400; Ruskin) at 37 °C for 48 h. YPD agar plates were incubated at 30 °C in aerobic condition. Colonies with distinct morphological differences (based on color, shape, size, rough or smooth surface) were selected and then purified using another agar plate of the same culture medium. DNA was extracted from each isolate through rapid mechanic cell lysis as described previously (Turroni et al., 2009). Isolates from each sample were stored at −80 °C in the presence of glycerol (30%, v/v).

4.3. Genus-specific PCR and taxonomic identification of probiotic isolates

For all bacterial species tested we used a specific couple of primers. All PCR amplifications were performed from 2 μL of a genomic DNA preparation at 10 ng/μL which was transferred directly to a 12.5 μL PCR mixture containing Platinum PCR SuperMix 1× (Invitrogen, USA) and 100 pM of each oligo. PCR reactions were performed on a Verity Thermocycler (Applied Biosystems, USA). Electrophoretic profiles were visualized by SYBR Safe DNA gel stain (Invitrogen). For *Streptococcus*-specific PCR we used primers Str1 (5′-GTACAGTTGCTTCAGGACGTATC-3′) and Str2 (5′-ACGTTCGATTTCATCACGTTG-3′) (Picard et al., 2004). For *Lactobacillus*-specific and *Bifidobacterium*-genus PCR we used the couples of primers Lab 0677F (5′-CTCCATGTGTAGCGGTG-3′), Lact71R (5′-TCAAAACTAAACAAAGTTTC-3′) and BIF-specific (5′-GGTGTGAAAGTCCATCGCT-3′), 23S_bif (5′-GTCTGCCAAGGCATCCACCA-3′) (Moura et al., 2007; Turroni et al., 2009) respectively. For *S. cerevisiae* detection, we used SC-FW (5′-GGACTCTGGACATGCAAGAT-3′) and SC-RV (5′-ATACCCTTCTTAACACCTGGC-3′) (Salinas et al., 2009).

The identification of each selected isolate was then performed by PCR amplification of a portion of the 16S rRNA gene using primers P0 (5′-GAAGAGTTTGATCCTGGCTCAG-3′) and P6 (5′-CTACGGCTACCTTGTTACGA-3′). Each 25 μL PCR reaction contained approximately 30 ng of genomic DNA, Platinum PCR SuperMix 1× (Invitrogen, USA) and 100 pM of each oligo. PCR reactions were performed on a Verity Thermocycler (Applied Biosystems, USA). Electrophoretic profiles were visualized by SYBR Safe DNA gel stain (Invitrogen). PCR product purification was performed using the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel) following manufacturer instructions. Each 16S rRNA gene generated from individual colonies originating from each probiotic sample was sequenced and it was then subjected to a BLAST search against the GenBank database.

4.4. Evaluation of bacterial cell density by qPCR

Real-Time PCR reactions were performed on MicroAmp optical plates sealed with MicroAmp optical caps (Applied Biosystems, Foster City, 2 CA) and amplifications were carried out on a CFX96 system (BioRad, CA, USA). Primers used in this study are Ban_Lac_fw (5′-CGAAGAGGACCATGACTTCC-3′) and Ban_Lac_rv (5′-CGGTGAGCTCCACATTGA-3′) to determine *B. animalis* subsp. *lactis*, Bbif_0282Fw (5′-GCGAACAATGATGGCACCTA-3′) and Bbif_0282Rv (5′-GTCGAACACCACGACGATGT-3′) to enumerate *B. bifidum* (Turroni et al., 2016), Rha I (5′-CAGACTGAAAGTCTGACGG-3′) and Rha II (5′-GCGATGCGAATTTCTATTATT-3′) to determinate *Lb. rhamnosus* (Byun and Yoon, 2003) and Str1 (5′-GTACAGTTGCTTCAGGACGTATC-3′) and Str2 (5′-ACGTTCGATTTCATCACGTTG-3′) for *Streptococcus* genus-specific identification (Picard et al., 2004). Each PCR reaction mix contained 7.5 μL 2× SYBR SuperMix Green (BioRad, CA, USA), 5 μL of DNA dilution, forward and reverse primers at 10pM, and nuclease-free water was added to obtain a final volume of 15 μL. PCR amplicons were detected with SYBR Green fluorescent dye and amplified according to the following protocol: one cycle of 95 °C for 2 min, 40 cycles of 95 °C for 5 s and 60 °C for 30 s. Amplicon specificity was judged based on the dissociation curve of PCR end products by increasing the temperature at a rate of 1 °C every 30 s from 60 to 95 °C. Negative controls composed by water were included for each primer set in each run.

4.5. 16S rRNA gene amplification

Partial 16S rRNA gene sequences were amplified from extracted DNA using primer pair Probio_Uni and/Probio_Rev, which targets the V3 region of the 16S rRNA gene sequence (Milani et al., 2013). Illumina adapter overhang nucleotide sequences were then added to the partial 16S rRNA gene-specific amplicons, which in turn were further processed by employing the 16S Metagenomic Sequencing Library Preparation Protocol (Part no. 15044223 Rev. B—Illumina; see also below). The PCR conditions used were 5 min at 95 °C, 35 cycles of 30 s at 94 °C, 30 s at 55 °C and 90 s at 72 °C, followed by 10 min at 72 °C. Amplifications were carried out using a Verity Thermocycler (Applied Biosystems). The integrity of the PCR amplicons was analyzed by electrophoresis on a 2200 TapeStation Instrument (Agilent Technologies, USA).

4.6. MiSeq sequencing of 16S rRNA gene-based amplicons

PCR products obtained following amplification of part of the 16S rRNA gene sequences were purified by a magnetic purification step involving Agencourt AMPure XP DNA purification beads (Beckman Coulter Genomics GmbH, Bernried, Germany) in order to remove primer dimers. The DNA concentration of the amplified sequence library was estimated through fluorimetric Qubit quantification system (Life Technologies). Amplicons were diluted to 4 nM and 5 μL of each diluted DNA amplicon sample was mixed to prepare the pooled final library. Sequencing was performed using an Illumina MiSeq sequencer with MiSeq Reagent Kit v3 chemicals.

4.7. Analysis of 16S rRNA microbial profiling datasets

The fastq files were processed using QIIME (Caporaso et al., 2010) as previously described (Milani et al., 2013). Paired-end reads were merged, and quality control implementation allowed the retention of sequences with a length between 140 and 400 bp, mean sequence quality score > 25 and with truncation of a sequence at the first base if a low quality within a rolling 10-bp window was found. Sequences with mismatched forward and/or reverse primers were omitted. 16S rRNA operational taxonomic units (OTUs) were defined at ≥99% sequence homology using uclust (Edgar, 2010). All reads were classified to the lowest possible taxonomic rank using QIIME (Caporaso et al., 2010) and a reference dataset from the SILVA database v. 128 (Quast et al., 2013). The classification at species level was confirmed through manual alignment of the OTUs using online blastn tool (https://blast.ncbi.nlm.

nih.gov/Blast.cgi).

4.8. ITS microbial and fungal profiling analysis

Profiling of known bifidobacterial species was performed using the primer pair Probio_bif_uni/Probio_bif_rev, an improved bifidobacterial ITS database encompassing all publicly available bifidobacterial genomes and a custom bioinformatics script, as described previously (Milani et al., 2014). Moreover, prediction of the fungal composition at species level was completed using the primer pair BITS/B58S3 and based on the UNITE database (Koljalg et al., 2005).

4.9. Shotgun metagenomics

DNA library preparation was performed using the Nextera XT DNA sample preparation kit (Illumina, San Diego, CA) according to the manufacturer's instructions. In brief, 1 ng input DNA from each sample was used for library preparation. The isolated DNA underwent fragmentation, adapter ligation and amplification. The ready-to-go libraries were pooled equimolarly, denaturated and diluted to a sequencing concentration of 1.8 pM. Sequencing was performed on NextSeq 550 instrument (Illumina, San Diego, CA), according to the manufacturer's instructions, using the 2 × 150 bp High Output sequencing kit, and spike-in of 1% PhiX control library.

4.10. Genome reconstruction of the probiotic strains

The generated paired fastq files were used as input for SPAdes assembler v3.9 (Bankevich et al., 2012), for the novo metagenomic assemblies using default parameters coupled with k-mer sizes of 21, 33, 55 and 77, and enabling the metagenomic flag option (-meta). ORFs of each assembled contig were predicted with Prodigal (Hyatt et al., 2010) and annotated by means of the software MEGAnnotator (Lugli et al., 2016). The annotated contigs were taxonomically classified based on the gene hit obtained through the NCBI RefSeq databases using the script Contig Classifier (CoCla) (Lugli et al., 2017). Species-specific contigs were subdivided in different files, obtaining the reconstructed genomes of each bacterial strain included in each probiotic sample.

4.11. Genome analyses and sequence comparisons

Genome sequence alignments were performed using MAUVE software (Darling et al., 2010), while the SNPs count was obtained through read mapping by means of Burrows-Wheeler Aligner v0.7.15 (Li and Durbin, 2009).

4.12. Availability of data and materials

The 16S rRNA- and ITS-profiling data sequenced in this study were deposited in the Sequence Read Archive (SRA) database under accession number PRJNA422025. Shotgun metagenomics data are accessible through SRA study accession number PRJNA422026.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijfoodmicro.2019.01.011.

Acknowledgements

This work was funded by the EU Joint Programming Initiative – A Healthy Diet for a Healthy Life (JPI HDHL, http://www.healthydietforhealthylife.eu/) to DvS (in conjunction with Science Foundation Ireland [SFI], Grant number 15/JP-HDHL/3280) and to MV (in conjunction with MIUR, Italy). We thank GenProbio srl for financial support of the Laboratory of Probiogenomics. DvS is a member of The APC Microbiome Institute funded by Science Foundation Ireland (SFI), through the Irish Government's National Development Plan (Grant number SFI/12/RC/2273). Part of this research is conducted using the High Performance Computing (HPC) facility of the University of Parma.

Competing interests

The authors declare that they have no competing interests.

References

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.

Besselink, M.G., van Santvoort, H.C., Buskens, E., Boermeester, M.A., van Goor, H., Timmerman, H.M., Nieuwenhuijs, V.B., Bollen, T.L., van Ramshorst, B., Witteman, B.J., Rosman, C., Ploeg, R.J., Brink, M.A., Schaapherder, A.F., Dejong, C.H., Wahab, P.J., van Laarhoven, C.J., van der Harst, E., van Eijck, C.H., Cuesta, M.A., Akkermans, L.M., Gooszen, H.G., Dutch Acute Pancreatitis Study, G., 2008. Probiotic prophylaxis in predicted severe acute pancreatitis: a randomised, double-blind, placebo-controlled trial. Lancet 371, 651–659.

Bokulich, N.A., Mills, D.A., 2013. Improved selection of internal transcribed spacer-specific primers enables quantitative, ultra-high-throughput profiling of fungal communities. Appl. Environ. Microbiol. 79, 2519–2526.

Byun, J.R., Yoon, Y.H., 2003. Identification and comparison of the nucleotide sequence of 16S-23S rRNA gene intergenic small SR(spacer region) of Lactobacillus rhamnosus ATCC 53103 with those of L. casei, L. acidophilus and L. helveticu. Asian-Aust. J. Anim. Sci. 16, 1816–1821.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, 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. Nat. Methods 7, 335–336.

Chenoll, E., Codoñer, F., Silva, A., Ibáñez, A., Martinez-Blanch, J., Bollati-Fogolín, M., Crispo, M., Ramírez, S., Sanz, Y., Ramón, D., Genovés, S., 2013. Genomic sequence and pre-clinical safety assessment of Bifidobacterium longum CECT 7347, a probiotic able to reduce the toxicity and inflammatory potential of gliadin-derived peptides. J. Probiotics Health 1 (2), 106.

Chiron, C., Tompkins, T.A., Burguiere, P., 2018. Flow cytometry: a versatile technology for specific quantification and viability assessment of micro-organisms in multistrain probiotic products. J. Appl. Microbiol. 124, 572–584.

Darling, A.E., Mau, B., Perna, N.T., 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One 5, e11147.

Donelli, G., Vuotto, C., Mastromarino, P., 2013. Phenotyping and genotyping are both essential to identify and classify a probiotic microorganism. Microb. Ecol. Health Dis. 24.

Doron, S., Snydman, D.R., 2015. Risk and safety of probiotics. Clin. Infect. Dis. 60 (Suppl. 2), S129–S134.

Drago, L., Rodighiero, V., Celeste, T., Rovetto, L., De Vecchi, E., 2010. Microbiological evaluation of commercial probiotic products available in the USA in 2009. J. Chemother. 22, 373–377.

Edgar, R.C., 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26, 2460–2461.

Fontana, L., Bermudez-Brito, M., Plaza-Diaz, J., Munoz-Quezada, S., Gil, A., 2013. Sources, isolation, characterisation and evaluation of probiotics. Br. J. Nutr. 109 (Suppl. 2), S35–S50.

Garrigues, C., Johansen, E., Pedersen, M.B., 2010. Complete genome sequence of Bifidobacterium animalis subsp. lactis BB-12, a widely consumed probiotic strain. J. Bacteriol. 192, 2467–2468.

Goldstein, E.J., Citron, D.M., Claros, M.C., Tyrrell, K.L., 2014. Bacterial counts from five over-the-counter probiotics: are you getting what you paid for? Anaerobe 25, 1–4.

Govender, M., Choonara, Y.E., Kumar, P., du Toit, L.C., van Vuuren, S., Pillay, V., 2014. A review of the advancements in probiotic delivery: conventional vs. non-conventional formulations for intestinal flora supplementation. AAPS PharmSciTech 15, 29–43.

Hill, C., Guarner, F., Reid, G., Gibson, G.R., Merenstein, D.J., Pot, B., Morelli, L., Canani, R.B., Flint, H.J., Salminen, S., Calder, P.C., Sanders, M.E., 2014. The International Scientific Association for Probiotics and Prebiotics consensus statement on the scope and appropriate use of the term probiotic. Nat. Rev. Gastroenterol. Hepatol. 11, 506.

Hyatt, D., Chen, G.L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinforma. 11, 119.

Johnson, B.R., Klaenhammer, T.R., 2014. Impact of genomics on the field of probiotic research: historical perspectives to modern paradigms. Antonie Van Leeuwenhoek 106, 141–156.

Koljalg, U., Larsson, K.H., Abarenkov, K., Nilsson, R.H., Alexander, I.J., Eberhardt, U., Erland, S., Hoiland, K., Kjoller, R., Larsson, E., Pennanen, T., Sen, R., Taylor, A.F., Tedersoo, L., Vralstad, T., Ursing, B.M., 2005. UNITE: a database providing web-based methods for the molecular identification of ectomycorrhizal fungi. New Phytol. 166, 1063–1068.

Koonin, E.V., 2015. The turbulent network dynamics of microbial evolution and the statistical tree of life. J. Mol. Evol. 80, 244–250.

Lee, J.H., Karamychev, V.N., Kozyavkin, S.A., Mills, D., Pavlov, A.R., Pavlova, N.V., Polouchine, N.N., Richardson, P.M., Shakhova, V.V., Slesarev, A.I., Weimer, B.,

O'Sullivan, D.J., 2008. Comparative genomic analysis of the gut bacterium Bifidobacterium longum reveals loci susceptible to deletion during pure culture growth. BMC Genomics 9, 247.
Li, H., Durbin, R., 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25, 1754–1760.
Lugli, G.A., Milani, C., Mancabelli, L., van Sinderen, D., Ventura, M., 2016. MEGAnnotator: a user-friendly pipeline for microbial genomes assembly and annotation. FEMS Microbiol. Lett. 363.
Lugli, G.A., Milani, C., Mancabelli, L., Turroni, F., Ferrario, C., Duranti, S., van Sinderen, D., Ventura, M., 2017. Ancient bacteria of the Otzi's microbiome: a genomic tale from the Copper Age. Microbiome 5, 5.
Machielsen, R., van Alen-Boerrigter, I.J., Koole, L.A., Bongers, R.S., Kleerebezem, M., Van Hylckama Vlieg, J.E., 2010. Indigenous and environmental modulation of frequencies of mutation in Lactobacillus plantarum. Appl. Environ. Microbiol. 76, 1587–1595.
Magadum, S., Banerjee, U., Murugan, P., Gangapur, D., Ravikesavan, R., 2013. Gene duplication as a major force in evolution. J. Genet. 92, 155–161.
Milani, C., Hevia, A., Foroni, E., Duranti, S., Turroni, F., Lugli, G.A., Sanchez, B., Martin, R., Gueimonde, M., van Sinderen, D., Margolles, A., Ventura, M., 2013. Assessing the fecal microbiota: an optimized ion torrent 16S rRNA gene-based analysis protocol. PLoS One 8, e68739.
Milani, C., Lugli, G.A., Turroni, F., Mancabelli, L., Duranti, S., Viappiani, A., Mangifesta, M., Segata, N., van Sinderen, D., Ventura, M., 2014. Evaluation of bifidobacterial community composition in the human gut by means of a targeted amplicon sequencing (ITS) protocol. FEMS Microbiol. Ecol. 90, 493–503.
Morovic, W., Hibberd, A.A., Zabel, B., Barrangou, R., Stahl, B., 2016. Genotyping by PCR and high-throughput sequencing of commercial probiotic products reveals composition biases. Front. Microbiol. 7, 1747.
Moura, P., Simoes, F., Girio, F., Loureiro-Dias, M.C., Esteves, M.P., 2007. PCR monitoring of Lactobacillus and Bifidobacterium dynamics in fermentations by piglet intestinal microbiota. J. Basic Microbiol. 47, 148–157.
Patro, J.N., Ramachandran, P., Barnaba, T., Mammel, M.K., Lewis, J.L., Elkins, C.A., 2016. Culture-independent metagenomic surveillance of commercially available probiotics with high-throughput next-generation sequencing. In: mSphere. 1.
Picard, F.J., Ke, D., Boudreau, D.K., Boissinot, M., Huletsky, A., Richard, D., Ouellette, M., Roy, P.H., Bergeron, M.G., 2004. Use of tuf sequences for genus-specific PCR detection and phylogenetic analysis of 28 streptococcal species. J. Clin. Microbiol. 42, 3686–3695.
Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glockner, F.O., 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596.
Salinas, F., Garrido, D., Ganga, A., Veliz, G., Martinez, C., 2009. Taqman real-time PCR for the detection and enumeration of Saccharomyces cerevisiae in wine. Food Microbiol. 26, 328–332.
Sanders, M.E., Klaenhammer, T.R., Ouwehand, A.C., Pot, B., Johansen, E., Heimbach, J.T., Marco, M.L., Tennila, J., Ross, R.P., Franz, C., Page, N., Pridmore, R.D., Leyer, G., Salminen, S., Charbonneau, D., Call, E., Lenoir-Wijnkoop, I., 2014. Effects of genetic, processing, or product formulation changes on efficacy and safety of probiotics. Ann. N. Y. Acad. Sci. 1309, 1–18.
Soucy, S.M., Huang, J., Gogarten, J.P., 2015. Horizontal gene transfer: building the web of life. Nat. Rev. Genet. 16, 472–482.
Stahl, B., Barrangou, R., 2013. Complete genome sequence of probiotic strain Lactobacillus acidophilus La-14. Genome Announc. 1.
Sybesma, W., Molenaar, D., van, I.W., Venema, K., Kort, R., 2013. Genome instability in Lactobacillus rhamnosus GG. Appl. Environ. Microbiol. 79, 2233–2239.
Toscano, M., de Vecchi, E., Rodighiero, V., Drago, L., 2013. Microbiological and genetic identification of some probiotics proposed for medical use in 2011. J. Chemother. 25, 156–161.
Turroni, F., Foroni, E., Pizzetti, P., Giubellini, V., Ribbera, A., Merusi, P., Cagnasso, P., Bizzarri, B., de'Angelis, G.L., Shanahan, F., van Sinderen, D., Ventura, M., 2009. Exploring the diversity of the bifidobacterial population in the human intestinal tract. Appl. Environ. Microbiol. 75, 1534–1545.
Turroni, F., Milani, C., Duranti, S., Mancabelli, L., Mangifesta, M., Viappiani, A., Lugli, G.A., Ferrario, C., Gioiosa, L., Ferrarini, A., Li, J., Palanza, P., Delledonne, M., van Sinderen, D., Ventura, M., 2016. Deciphering bifidobacterial-mediated metabolic interactions and their impact on gut microbiota by a multi-omics approach. ISME J. 10, 1656–1668.
Vandeputte, D., Kathagen, G., D'Hoe, K., Vieira-Silva, S., Valles-Colomer, M., Sabino, J., Wang, J., Tito, R.Y., De Commer, L., Darzi, Y., Vermeire, S., Falony, G., Raes, J., 2017. Quantitative microbiome profiling links gut community variation to microbial load. Nature 551, 507–511.
Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G.F., Chater, K.F., van Sinderen, D., 2007. Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. Microbiol. Mol. Biol. Rev. 71, 495–548.