Food Microbionet: a tool for the visualisation and exploration of food microbial communities based on network analysis

Dear Author,

Please, note that changes made to the HTML content will be added to the article before publication, but are not reflected in this PDF.

Note also that this file should not be used for submitting corrections.

International Journal of Food Microbiology xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

International Journal of Food Microbiology

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

FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis

Eugenio Parente [a,*], Luca Cocolin [b], Francesca De Filippis [c], Teresa Zotta [d], Ilario Ferrocino [b], Orla O'Sullivan [e,f], Erasmo Neviani [g], Maria De Angelis [h], Paul D. Cotter [e,f], Danilo Ercolini [c]

[a] Dipartimento di Scienze, Università degli Studi della Basilicata, Potenza, Italy
[b] Department of Agricultural, Forest and Food Science, University of Torino, Grugliasco, Italy
[c] Department of Agricultural Sciences, Division of Microbiology, University of Naples Federico II, Portici, Italy
[d] Istituto di Scienze dell'Alimentazione, CNR, Avellino, Italy
[e] Teagasc Food Research Centre, Moorepark, Fermoy, Ireland, UK
[f] APC Microbiome Institute, Cork, Ireland, UK
[g] Department of Food Science, Parma University, Parco Area delle Scienze 48, /A, Parma, Italy
[h] Department of Soil, Plant and Food Science, University of Bari Aldo Moro, Bari, Italy

ARTICLE INFO

Article history:
Received 17 September 2015
Received in revised form 27 November 2015
Accepted 4 December 2015
Available online xxxx

Keywords:
Food bacterial communities
Network analysis
16S rRNA amplicon-based high-throughput sequencing

ABSTRACT

Amplicon targeted high-throughput sequencing has become a popular tool for the culture-independent analysis of microbial communities. Although the data obtained with this approach are portable and the number of sequences available in public databases is increasing, no tool has been developed yet for the analysis and presentation of data obtained in different studies. This work describes an approach for the development of a database for the rapid exploration and analysis of data on food microbial communities. Data from seventeen studies investigating the structure of bacterial communities in dairy, meat, sourdough and fermented vegetable products, obtained by 16S rRNA gene targeted high-throughput sequencing, were collated and analysed using Gephi, a network analysis software. The resulting database, which we named FoodMicrobionet, was used to analyse nodes and network properties and to build an interactive web-based visualisation. The latter allows the visual exploration of the relationships between Operational Taxonomic Units (OTUs) and samples and the identification of core- and sample-specific bacterial communities. It also provides additional search tools and hyperlinks for the rapid selection of food groups and OTUs and for rapid access to external resources (NCBI taxonomy, digital versions of the original articles). Microbial interaction network analysis was carried out using CoNet on datasets extracted from FoodMicrobionet: the complexity of interaction networks was much lower than that found for other bacterial communities (human microbiome, soil and other environments). This may reflect both a bias in the dataset (which was dominated by fermented foods and starter cultures) and the lower complexity of food bacterial communities.
Although some technical challenges exist, and are discussed here, the net result is a valuable tool for the exploration of food bacterial communities by the scientific community and food industry.

1. Introduction

The degree of complexity of the microbiota that potentially impacts on food quality and safety is extremely variable. In both fermented and non-fermented foods, the type of contaminating microbiota can initially be rather diverse, reflecting mainly on the original microbiota of the raw material, processing, handling and storage conditions and the level of good manufacture practice (Bokulich et al., 2012; Bokulich and Mills, 2013; Chaillou et al., 2015, Cocolin and Ercolini, 2015; De Filippis et al., 2013). However, depending on the storage conditions and other extrinsic factors, only a few species and strains will be able to develop sufficiently in the food matrix to significantly affect the food quality (by spoilage or fermentation) or safety.

The methods employed to study microbes and microbial diversity in foods have evolved, and in turn, have also revolutionized our overall understanding of the microbial ecology of foods (Cocolin and Ercolini, 2015). The culture-independent evaluation of food microbial diversity by high-throughput rRNA gene sequencing has become an increasingly popular approach to food microbiology. After microbial nucleic acid extraction from the food, the DNA (or cDNA in cases where RNA was targeted) is used as a template to amplify variable regions within or across the rRNA genes of bacteria (16S) or fungi (internal transcribed spacer [ITS] or other target) and an amplicon library is then sequenced using high-throughput sequencing (HTS; Ercolini, 2013) platforms.

* Corresponding author at: Dipartimento di Scienze, Università degli Studi della Basilicata, Viale dell'Ateneo Lucano, 10, 85100 Potenza, Italy.
E-mail address: eugenio.parente@unibas.it (E. Parente).

http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001
0168-1605/

The result is a food (sample)-specific profile of the microbiota where all the microbial entities are identified at variable taxonomic depth. Based on the number of sequence reads assigned to a given taxon, the relative abundance of each identified operational taxonomic unit (OTU) can be determined. Therefore, for each food sample analysed, a clear understanding of the composition and relative abundances of the microorganisms populating the food at that time can be provided. The advantages and disadvantages of this methodology have been discussed elsewhere (Bokulich and Mills, 2013; Ercolini, 2013).

Many different sequencing technologies are available for the generation of sequence data (Glenn, 2011. http://www.molecularecologist.com/next-gen-fieldguide-2014). Regardless of how the data is generated, accurate data analysis tools are pivotal in any study of microbial ecology; from quality filtering to graphical representations, the software and the algorithms selected can greatly impact on the results and interpretation. Essential steps in any analysis pipeline include post-sequencing quality checking (based on both length and quality scores), clustering into OTUs, chimera removal, alignment, taxonomical assignment and diversity analysis. The choice of the pipeline has been proven to significantly affect the results in terms of estimated diversity and microbial community structure (May et al., 2014). Diversity can be calculated from both a within sample (alpha diversity) and between sample perspective (beta diversity). Numerous packages have been developed for rRNA gene amplicon data analysis, primarily designed for UNIX based operating systems. The most widely used packages are QIIME (Caporaso et al., 2010) and MOTHUR (Schloss et al., 2009). They have become popular because they provide a pre-compiled and user-friendly analysis pipeline, but also due to their constant maintenance and updates.

Beyond the power linked to the sensitivity and the throughput of sequencing-based microbiota analysis, a fundamental advantage is the possibility of using the raw sequence data in meta-studies. In fact, in contrast to previous culture-independent approaches, the sequencing-based tools offer the unprecedented advantage of making the results readily available for the scientific community through the deposit of the sequences in public databases (e.g. the Sequence Read Archive (SRA) of the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/Traces/sra) or the European Nucleotide Archive of the European Bioinformatics Institute (http://www.ebi.ac.uk/ena). This allows researchers to easily access datasets corresponding to diverse food samples generated by different laboratories and with different scopes.

Network analysis (Newman et al., 2006) tools have recently been used to provide effective and information dense displays of microbial communities for several environments (de Menezes et al., 2014; Deng et al., 2012; Muegge et al., 2011; Zhou et al., 2011a), including foods (Chaillou et al., 2015; De Filippis et al., 2013, 2014; Dolci et al., 2014; Ercolini et al., 2013; Oakley et al., 2013). In a network representation, objects (OTUs and/or samples) represent the nodes (or vertices), and are connected by links (edges). The edges can be directed (i.e. when the direction of the connection is of importance) or undirected and are usually associated with a weight. The latter can store information on the abundance of an OTU in a sample or the probability of a significant co-occurrence/co-exclusion relationship.

Two types of displays have been used in microbial ecology. In OTU-sample nets, the network is bipartite i.e. two types of nodes exist, sample and OTU nodes, and connections occur only between samples and associated OTUs. Conversely, co-occurrence/co-exclusion networks, which show significant positive or negative interactions among members of microbial communities have rarely (Chaillou et al., 2015; Mounier et al., 2008; Oakley et al., 2013) been used in food microbial ecology, but have been successfully applied to the study of the microbial communities of a variety of environments (miscellaneous environments: Deng et al., 2012; water: Liu et al., 2014; soil: Zhou et al., 2011a) and for the human microbiome (Faust et al., 2012). The methods and models used to derive interaction networks have been reviewed by Faust and Raes (2012) and, although they are riddled by pitfalls related to the structure of the data and to the sensitivity to the methods and parameters selected in the analysis (Faust and Raes, 2012; Kuczynski et al., 2010), they offer significant advantages with respect to detecting biologically and ecologically relevant relationships among members of microbial communities.

The number of HTS studies of food microbial communities has been increasing steadily in recent years (Mayo et al., 2014), but the information is dispersed in a large number of papers, each analysing a single or a limited range of foods. It is therefore tempting and timely to collect and integrate data from several studies in such a way that the results can be readily searched and visualized, even by relatively inexperienced users. With the aim of providing flexible means for meta-studies in food microbial ecology, here we present FoodMicrobionet, a database and visualisation tool based on network analysis, and some examples of the potential of the tool in terms of data display and analysis.

2. Material and methods

2.1. Data sources

FoodMicrobionet 1.0 includes data from 17 studies, on dairy products, dairy starter cultures, raw and fermented meat, doughs and sourdoughs, or fermented vegetables. The list of studies, with information on the sequencing platforms, software employed for bioinformatics analysis, and the databases used for OTU assignment is shown in Table 1.

2.2. Data tables

Abundance tables, including taxonomic lineages for each OTU, were obtained from each contributor and transformed in tab-delimited nodes and edges tables, which were collated and curated to remove duplicates. The node and edge tables and their specifications are provided in section 1 of the Supplementary Material.

2.3. Network analysis

2.3.1. OTU - food network

The edges tables were imported in Gephi 0.8.2-beta (http://gephi.github.io/; Bastian and Jacomy, 2009) using the “Import spreadsheet” feature. Nodes tables were then imported to retrieve the metadata for each node. Statistics (degree and weighted degree, centrality statistics, network diameter, graph density, average path length) were then calculated for each node and for the network using the statistical module of Gephi. A glossary of terms for node and network statistics is provided in Table 2. Styles were then applied to the nodes to enhance the display: the colour of the node was attributed on the basis of a custom field containing families for OTUs and Food subgroup for samples; the size of the nodes was made proportional to the weighted degree of the node; edge thickness was made proportional to the weight of the connection. A Yfan Hu force based layout algorithm was finally applied (Hu, 2006). Simplified versions of the networks were obtained by filtering. The whole network was then exported for web visualisation using the Sigmajs exporter plugin of Gephi.

2.3.2. Microbial interaction networks

Microbial interaction networks were generated for selected groups of samples extracted from FoodMicrobionet using the CoNet app (Faust et al., 2012) of Cytoscape 3.2.1. OTU abundance (as number of sequences per sample) tables were then imported using CoNet, and five methods (Pearson, Spearman, Mutual information, Bray Curtis, Kullback–Leibler) were used to mine for significant co-occurrence/co-exclusion relationships. Null distributions were generated using the edge-scores routine and random distributions using the bootstrap routine. Brown's method was used to merge method specific p-values and

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001

Table 1
List of studies used in FoodMicrobionet 1.0. The FoodID code of the FoodEx 2 classification (EFSA, 2011) is shown in the short description.

Target/region	Platform[1]2	Read length (bp) [2]	Sequencing center[3]	Data analysis software	OTU picking method	Assign taxonomy method	Database	Sequence accession number	Samples	Short description	Reference
16S DNA V1–V3	GSJ	504	DA-UNINA	QIIME 1.8.0	De novo, uclust pipeline	RDP	Greengenes	SRP052240	29	Commercial high-moisture Mozzarella cheese produced with different acidification methods (A02QJ)	Guidone et al. (2016)
16S DNA V1–V3	GSJ	498	DA-UNINA	QIIME 1.8.0	De novo, uclust pipeline	RDP	Greengenes	SRP057506	24	Undefined strain starters (milk cultures) for high-moisture Mozzarella cheese (A048Y)	Guidone et al., unpublished data; Parente et al. (2016)
16S DNA V1–V3	GSJ	465	DA-UNINA	QIIME 1.6.0	De novo, uclust pipeline	RDP	Greengenes	SRP033419	50	Undefined strain starters (whey cultures, A048Y) and cheese curds for water-buffalo Mozzarella (A02QJ), Grana Padano (A02ZQ) and Parmigiano Reggiano cheese (A02ZS)	De Filippis et al. (2014)
16S DNA V1–V3	GSJ	492	DA-UNINA	QIIME 1.5.0	De novo, uclust pipeline	RDP	Greengenes	SRP014821	10	Raw milk (A02MD), whey culture (A048Y), curd before and after fermentation and water-buffalo Mozzarella cheese (A02QJ) from two different manufactures	Ercolini et al. (2012)
16S DNA V4–V5	4GSFLX	346	TFRC	QIIME 1.6.0	De novo, uclust pipeline	BLASTN	SILVA v100	ERP002650	46	Kefir grain and kefir milk from different sources (A02NV)	Marsh et al. (2013)
16S RNA (cDNA) V1–V3	GSJ	485	DA-UNINA	QIIME 1.6.0	De novo, uclust pipeline	RDP	Greengenes	SRP038100	11	Milk (A02MC), curd and ewe's milk Canestrato cheese during ripening (A02YH).	De Pasquale et al. (2014a)
16S RNA (cDNA) V1–V3	GSFLX	515	RTL	QIIME 1.6.0	De novo, usearch pipeline	BLASTN + distributed. NET algorithm	Greengenes	NA	11	Milk (A02LY), curd and Caciocavallo cheese during ripening (A02ZL).	De Pasquale et al. (2014b)
16S RNA (cDNA) V1–V3	GSJ	490	DA-UNINA	QIIME 1.7.0	De novo, uclust pipeline	RDP	Greengenes	SRP040575	27	Milk (from different lactation stages, A02LY), curd and Fontina cheese (A02TY) from three different dairies	Dolci et al. (2014)
16S RNA (cDNA) V1–V3	GSJ	469	DA-UNINA	QIIME 1.8.0	De novo, uclust pipeline	RDP	Greengenes	SRP044294	39	Piedmont hard cheese made from raw milk: milk (A02LV), curd and cheese throughout ripening (A02ZQ)	Cocolin et al., unpublished data
16S DNA and 16S RNA (cDNA) V1–V3	GSJ	601	DA-UNINA	QIIME 1.8.0	De novo, uclust pipeline	RDP	Greengenes	NA	68	Caciocavallo Silano cheese manufacture: starter culture (A048Y), milk (A02LY), curd and cheere throughout ripening (A02ZL).	De Filippis et al., 2013 unpublished data
16S DNA V1–V3	GSJ	457	DA-UNINA	QIIME 1.6.0	De novo, uclust pipeline	RDP	Greengenes	SRP021108	86	Swabs from bovine carcass/butchery environment and samples of fresh (t0) and spoiled (t6) beefsteaks (A01QX).	De Filippis et al. (2013)
16S RNA (cDNA) V1–V3	GSJ	469	DA-UNINA	QIIME 1.8.0	De novo, uclust pipeline	RDP	Greengenes	PRJNA264135	30	Piedmontese fermented meat during ripening (A024V)	Greppi et al. (2015)
16S RNA (cDNA) V1–V3	GSJ	472	DA-UNINA	QIIME 1.8.0	De novo, uclust pipeline	RDP	Greengenes	PRJNA272131	52	Hamburgers (controls or with added preservatives, nisin + EDTA) during storage in vacuum packaging (A049S)	Cocolin et al., unpublished data; Ferrocino et al. (2016)
16S DNA and 16S RNA (cDNA) V1–V3	GSFLX	432	RTL	QIIME 1.6.0	De novo, usearch pipeline	BLASTn	Greengenes	SRP019475	20	Olive surfaces (A01BP) and brine during fermentation. Samples were treated or untreated with NaOH 1% v/w	Cocolin et al. (2013)
16S DNA V1–V3	GSJ	496	DA-UNINA	QIIME 1.9.0	De novo, uclust pipeline	RDP	Greengenes	NA	16	Sourdoughs for croissant production (A008K)	Ercolini et al., unpublished data
16S RNA (cDNA) V1–V3	GSJ	486	DA-UNINA	QIIME 1.6.0	De novo, uclust pipeline	RDP	Greengenes	SRP044788	8	Durum wheat flour (cultivar Senatore Cappelli) grown under organic and conventional farming and related sourdoughs (A008K)	Rizzello et al. (2015)

(continued on next page)

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001

Table 1 (*continued*)

Target/region	Platform[1]2	Read length (bp) [2]	Sequencing center[3]	Data analysis software	OTU picking method	Assign taxonomy method	Database	Sequence accession number	Samples	Short description	Reference
16S RNA (cDNA) V1–V3	GSJ	477	DA-UNINA	QIIME 1.6.0	De novo, uclust pipeline	RDP	Greengenes	SRP019996	15	Rye and wheat sourdough A008K)	Ercolini et al. (2013)

[1] GSJ: 454 GS Junior platform, 454 Roche Life Sciences, Branford, CT, USA; GSFLX 454 Genome Sequencer FLX Titanium System, 454 Roche Life Sciences, Branford, CT, USA

[2] Average sequence length after filtering.

[3] DA-UNINA: Dipartimento di Agraria, Università degli Studi di Napoli "Federico II", (Portici, Italy); TFRC: Teagasc Food Research Centre (Moorepark, Co. Cork, Ireland); RTL: Research and Testing Laboratories (Lubbock, TX, USA)

the Benjamini Hochberg method was used to adjust the p-values for multiple testing. Interactions were evaluated for both low level taxa and high level taxa, but a parent child exclusion filter was used to avoid interactions between a high level taxon and its members (e.g. between *Lactobacillaceae* and members of the genus *Lactobacillus*). To simplify network visualisations, interactions with high level taxa were included only if they did not duplicate interaction due to a lower level taxon (i.e. if *Lactobacillales* and *Lactobacillus delbrueckii* shared the same interactions, the former node was removed). Topological properties of the interaction networks were evaluated using the NetworkAnalyzer tool of Cytoscape.

Table 2
An informal glossary of terms for node and network statistics for FoodMicrobionet.

Property	Definition
Node statistics	
Degree	The number (k) of incoming or outgoing connections to a node. In FoodMicrobionet, a bipartite network with two types of nodes, it is the number of Operational Taxonomic Units (OTUs) associated with a food sample node or number of food samples in which an OTU is found for OTU nodes. In microbial interaction networks, which are unipartite, because they include only OTU nodes, it is the number of statistically significant interactions.
Weight	In FoodMicrobionet it is a continuous value (0–100%) assigned to an edge connecting an OTU and a food sample node, the abundance of an OTU in a sample node. In microbial interaction networks it is the q-value, showing the supporting evidence for the interaction.
Weighted degree	The sum of weights for a node. By definition in FoodMicrobionet this value is 100 for food sample nodes, while for OTU nodes it is the sum of abundances in all food samples in the whole network or in a subnetwork obtained by filtering.
Clustering coefficient	A measure of the probability that the nodes connected to a given node are also connected among them (varies from 0 to 1)
Network properties	
Components	The number of connected subnetworks in a network
Average degree	The average of the degrees of all nodes in a network
Node degree distribution	The frequency distribution of degrees for the nodes of a network or subnetwork. By convention the fraction of nodes with a given degree, P(k), is shown on the y axis, while the degree is shown on the x axis.
Average path length	The average of all the shortest paths connecting a node to any other node in the network. The longest shorter path is the network diameter. In small world networks, the average path length is shorter (typically <6) than that of a random network with the same average degree.
Average clustering coefficient	The average of clustering coefficient of all nodes in a network; for random graphs it is calculated as the ratio between the average degree and number of nodes.
Modularity	A measure of how well a network can be separated in modules (or clusters) of highly connected nodes
Scale-free network	In a scale-free network few nodes have a large degree, while many have a large degree. Typically the node degree distribution follows a power law $P(k) \sim k^{-\gamma}$. Scale-free networks usually also show the small world effect and have a highly modular, hierarchical structure, with nodes with high degree (hubs) connecting modules of nodes with low degree to the rest of the network.

For more formal definitions, formulas and algorithms, the interested readers are referred to Deng et al. (2012); Dunne et al. (2002); Newman et al. (2006).

3. Results and discussion

3.1. Building FoodMicrobionet: a bipartite OTU-sample network for food bacterial communities

The main purpose of FoodMicrobionet is to provide a user-friendly tool to explore multiple datasets generated by 16S rRNA gene amplicon HTS studies of food bacterial communities. The flowchart for the development of FoodMicrobionet and of its products (visualisations, tables, graphs) is shown in Fig. 1. Data from seventeen published and unpublished studies (Table 1) on dairy and meat products, starter cultures, sourdoughs or fermented vegetable products (olives) were assembled in a database and a network was generated using Gephi 0.8.2-beta. The network has 964 OTU nodes and 552 sample nodes, with 18,115 edges (sample-OTU relationships), and is by far the largest such collection of data of food bacterial microbiota.

Network analysis software packages such as Gephi or Cytoscape (edges and nodes files provided in supplementary material can be easily imported in Cytoscape) offer a wide range of filtering, statistical analysis and graphical representation options but require some informatics skills. Therefore an interactive network visualisation (http://www.foodmicrobionet.org/fmbn1_0_3web/) was created using a publicly available plugin. This visualisation allows even inexperienced users to explore FoodMicrobionet, to select individual sample or OTU nodes, or to carry out group selections for sample and OTU nodes. Relevant properties for both OTU and food sample nodes can be visualised by either clicking on nodes or by selecting them using a search field. A user manual for the web visualisation is provided in section 2 of the Supplementary material.

Because of the high number of sample and OTU nodes, the information cannot be easily presented into a readable graph. Therefore, simplified, filtered and node-labelled sub-networks for meats, sourdoughs and dairy foods are shown in Fig. 2. The common features found in OTU-sample networks previously published (De Filippis et al., 2013, 2014; Dolci et al., 2014; Ercolini et al., 2013) are evident: sample nodes with similar microbiota occupy defined areas of the graph and are close to the OTU nodes that dominate their microbiota. This allows identifying easily the dominant, core and minor OTUs, that can be clearly distinguished by their position and by their node size.

Taxon specific sub-networks can be easily extracted. Examples for members of the families *Pseudomonadaceae* and *Enterobacteriaceae* are presented in Supplementary Figs. S1 and S2. In this version of the display the size of sample nodes is related to the cumulative abundance of the taxon and the size of the OTU nodes is related to the cumulative abundance of the OTU in the subnetwork. Edge thickness gives an estimate of the abundance of a given OTU in each sample, while colours can be used to estimate the relative abundance of food groups in which the selected taxon is found.

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001

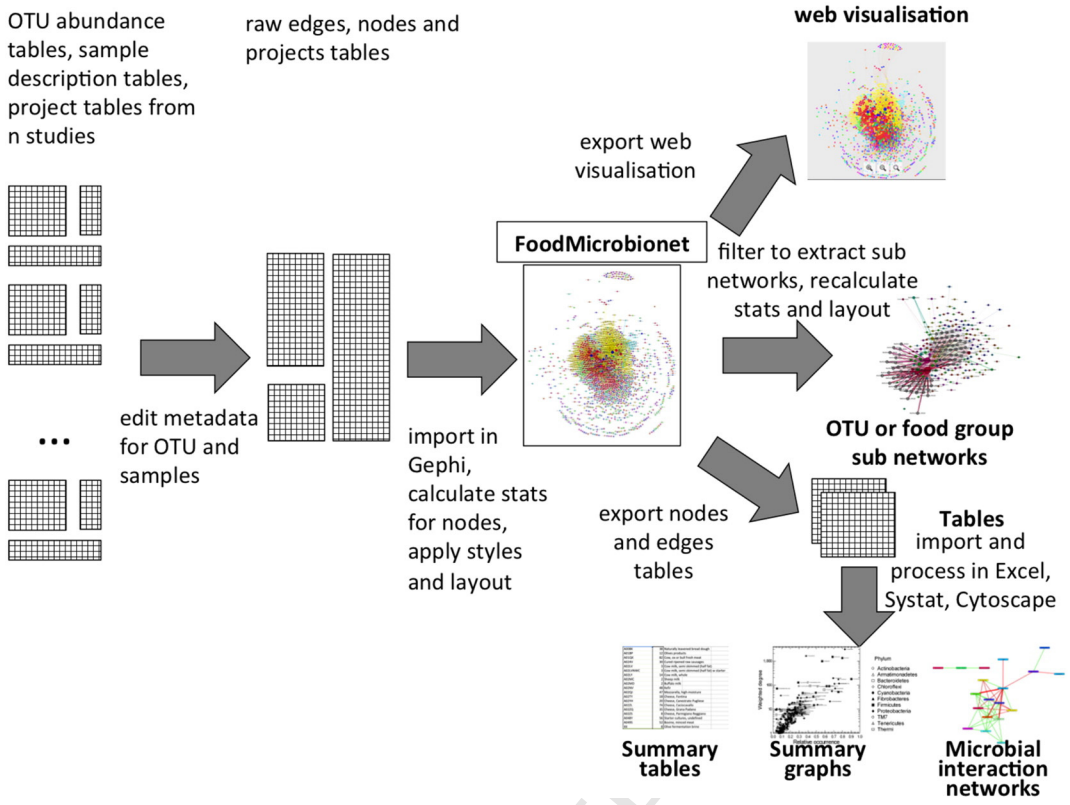

Fig. 1. Flowchart of the development of FoodMicrobionet v 1.0. FoodMicrobionet is a curated database of HTS studies on food bacterial communities which is implemented in Gephi 0.8.2, a network analysis software. The network file can then be used to generate a variety of products for visual and statistical analysis.

The node degree distribution for OTU nodes is shown in Supplementary Fig. S3. The distribution fits, albeit with a relatively low R^2 (0.832), a power law distribution with an exponent (γ) of 1.12 ± 0.04. Node degree power law distributions are indicative of a scale-free network (Dunne et al., 2002; Newman et al., 2006). Such networks are widely distributed in all fields (social networks, internet networks, power grids, bibliographic networks) and share several properties. They are usually large and complex, highly connected (large average degree), with a high number of nodes with low degree (in the case of FoodMicrobionet OTUs which are found only in one or few food samples) but with a small number of OTUs connected to a large number of samples (i.e. the 'signature' OTUs which make the core microbiota of a given group of food samples). Because FoodMicrobionet 1.0 includes different food groups, several signature OTUs with high degree are found, and this may affect the fit of the power law distribution.

FoodMicrobionet can also be used to obtain further information on distribution of taxa in different food groups by filtering and recalculation from nodes and edges tables. Information on dominating OTUs can be gathered by plots showing the weighted degree distribution (i.e. how abundant an OTU is in the whole dataset or in a subset) as a function of relative occurrence (i.e. the fractions of samples in which an OTU is found). An example for raw meat is shown in Fig. 3. Further examples for raw milk and mozzarella are shown in Supplementary Figs. S4 and S5. More traditional plots for OTU distribution can also be obtained. An example of the distribution of OTU belonging to different phyla in different food groups is shown in Supplementary Fig. S6.

3.2. Microbial interaction networks

Microbial interaction networks may help in formulating inferences on the phenomena underlying the structure of food microbial communities, from co-occurrence or co-exclusion patterns due to the occupation of different niches or to selective conditions allowing the growth of a subset of taxa, to relationships such as amensalism, commensalism, and symbiosis. The inference of microbial interactions is still affected by pitfalls: the results may be strongly affected by the level of coverage of the microbial community, by the bioinformatics pipelines used with specific options for clustering of the sequences and taxonomic assignment, by the procedures used in normalization and by the methods used to estimate the relationships, etc. However, robust methods have been developed to perform this analysis (Faust et al., 2012). Since OTU abundance tables were available for all datasets included in FoodMicrobionet, we explored co-occurrence/mutual exclusion patterns for all datasets for which a high enough number of samples was available. Statistics for all interaction networks are shown in Table 3 but a detailed discussion of microbial interactions is beyond the scope of this paper and only two examples are discussed below.

A microbial interaction network for the kefir dataset (Marsh et al., 2013) is shown in Fig. 4. The dataset included milk kefir and grains from different sources. Due to the very simple structure of bacterial communities in kefir and kefir grains only a few interactions were significant. The network has a very low complexity (7 nodes, average degree 3.41 and average path length 2.19), with a clustering coefficient of 0.714, and no fit of the power law for the node degree distribution. The occurrence of *Acetobacter* was negatively related with the occurrence of *Lactobacillales* and that of *Lactobacillus* with *Leuconostoc*, *Lactococcus* and *Streptococcus*. In fact, while *Lactobacillus* dominated the kefir grain microbiota, the latter genera showed a better ability to grow in milk kefir. Members of the family *Lachnospiraceae*, a minor group in the kefir microbiota, also systematically occurred in kefir grains, while they were almost always absent in milk kefir. On the other hand, the co-exclusion relationship with *Acetobacter* was observed in both grains and milk and may reflect conditions for storage and production of kefir.

A very complex interaction network was obtained for the beef dataset (De Filippis et al., 2013). The dataset included swabs from

different points of bovine carcasses cuts and beefsteaks obtained thereof, sampled at 0 day and after 7 days of aerobic storage at 4 °C, for two different samplings. The full network is shown in Supplementary Fig. S7, while a simplified version, including only the most abundant taxa, is shown in Fig. 5A, together with the interaction network inferred for spoiled beef steak samples (Fig. 5B). The complexity of co-occurrences and mutual exclusions in fresh raw meat mainly reflects the high diversity of bacterial communities (Fig. S7, Fig. 5A). Significant interactions among the most abundant taxa (*Moraxellaceae*, *Pseudomonadaceae*, *Aerococcaceae*, *Staphylococcaceae*, *Flavobacteriaceae*, *Rhodobacteriaceae* and *Corynebacteriaceae* on carcass swabs and freshly cut beefsteaks; *Pseudomonaceae*, *Listeriaceae*, *Moraxellaceae* and *Enterobacteriaceae* on spoiled steaks) confirm the co-occurrence and mutual exclusion patterns due to different samplings, different cuts, and

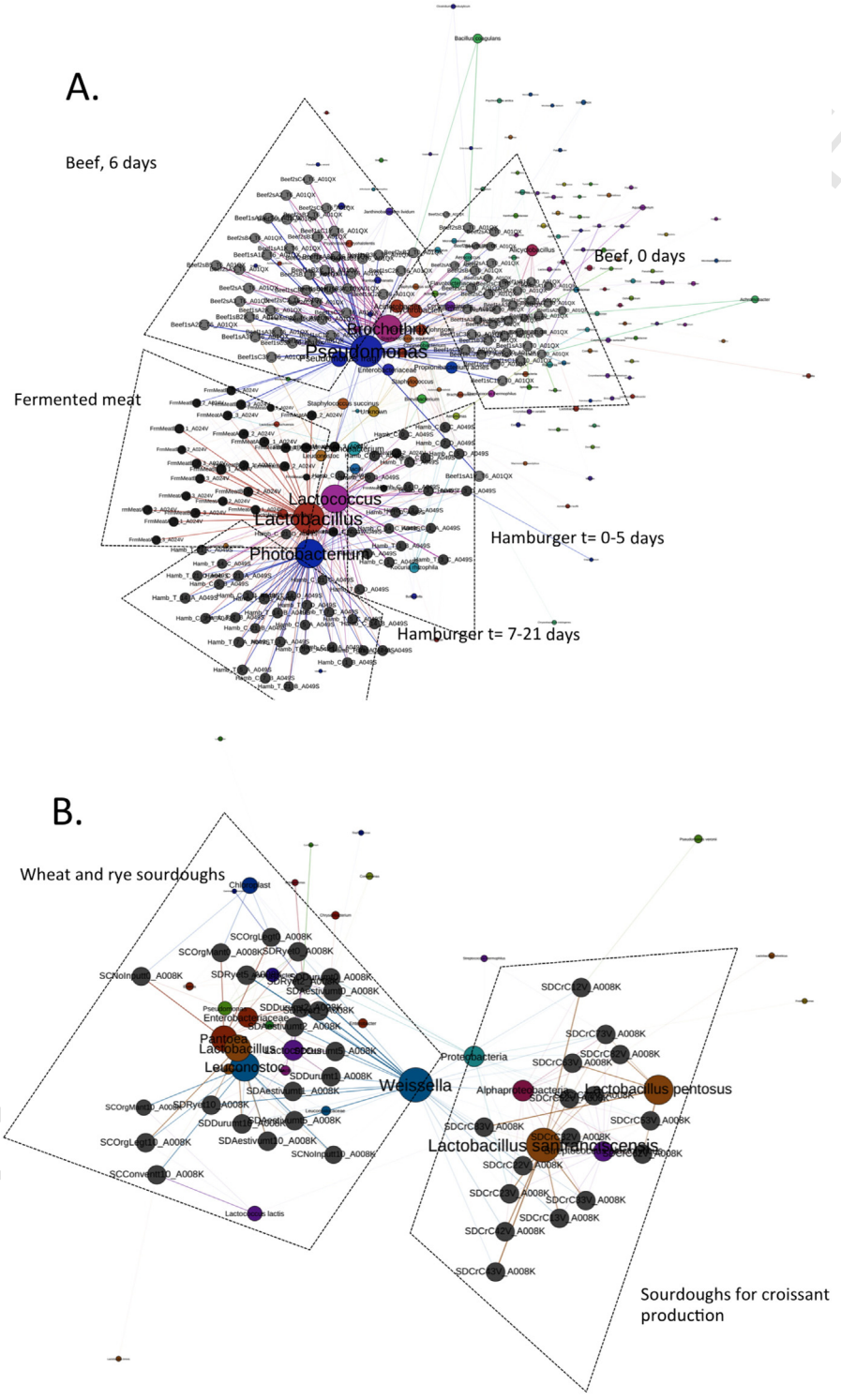

Fig. 2. Filtered (only OTU nodes with a cumulative abundance >5% are shown) for raw meat (A) sourdough samples (B) and dairy products and starters (C) extracted from FoodMicrobionet 1.0. Node colour (grey scale for food subgroups for sample nodes or bacterial family of OTU nodes) is used to highlight different sample and OTU nodes. Style features are used to enhance the graph: node size is related to the weighted degree (i.e. cumulative abundances for OTUs) while edge thickness is proportional to the abundance of an OTU in a given sample. Areas of the graph in which samples belonging to a given group are more abundant are enclosed by dashed lines.

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001

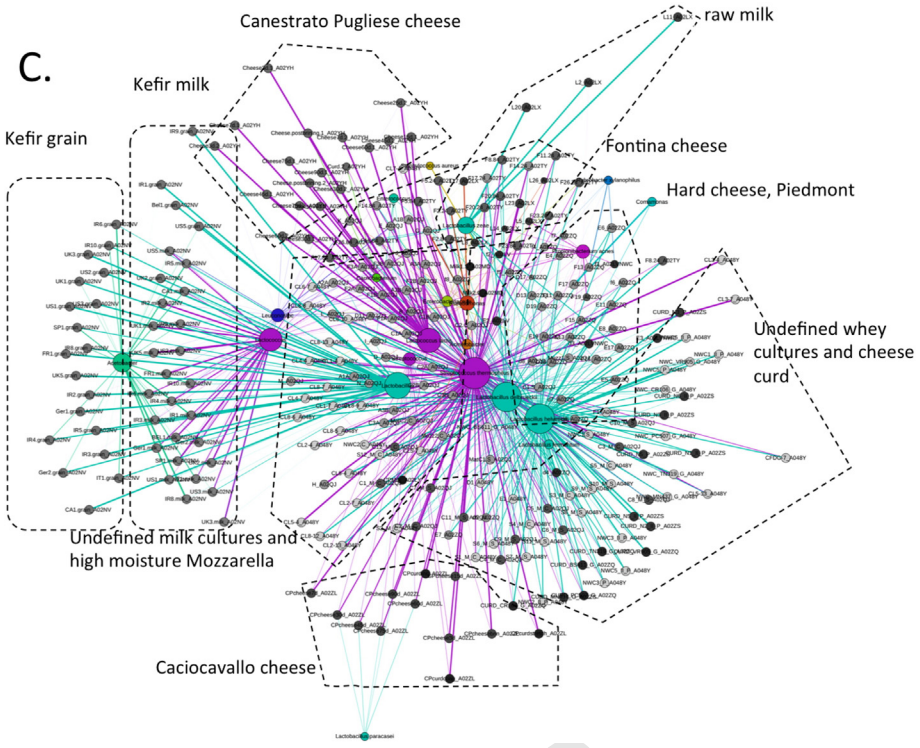

Fig. 2 (*continued*).

spoilage described by De Filippis et al. (2013). Spoilage dramatically reduced diversity (De Filippis et al., 2013) and simplified the microbial interaction network (Fig. 5B). The co-occurrence relationship between *Acinetobacter guillouiae* (a species occurring at low abundance) and *Enterobacteriaceae* is independent of the sampling and of the cut. On the other hand the mutual exclusion relationship between *Staphylococcus equorum* and *Serratia* is clearly related to the contamination patterns of the beef cuts, with the former species occurring systematically in thick flank cuts and members of the genus *Serratia* occurring in brisket and chuck cuts. The last set of interactions reflects different spoilage environments. In fact, the dominating spoilage organism was *Pseudomonas* in sampling 1 and *Brochothrix* in sampling 2 (De Filippis et al., 2013). *Carnobacterium*, *Acinetobacter johnsonii*, *Chryseobacterium* and members of the class *Actinobacteria* also occurred more frequently in beef steaks from sampling 1.

The interaction networks inferred in our study (Table 3) are less complex (sometimes dramatically) than those inferred for environmental bacterial communities (Deng et al., 2012) or for the human microbiome (Faust et al., 2012). In addition, they do not show any fit of the power law for either the node degree distribution or the clustering coefficient/degree relationship, showing that they are neither scale-free nor show a hierarchical structure. However, the average clustering coefficient is often higher than that of random networks with the same size and average degree. In general, interaction networks for fermented

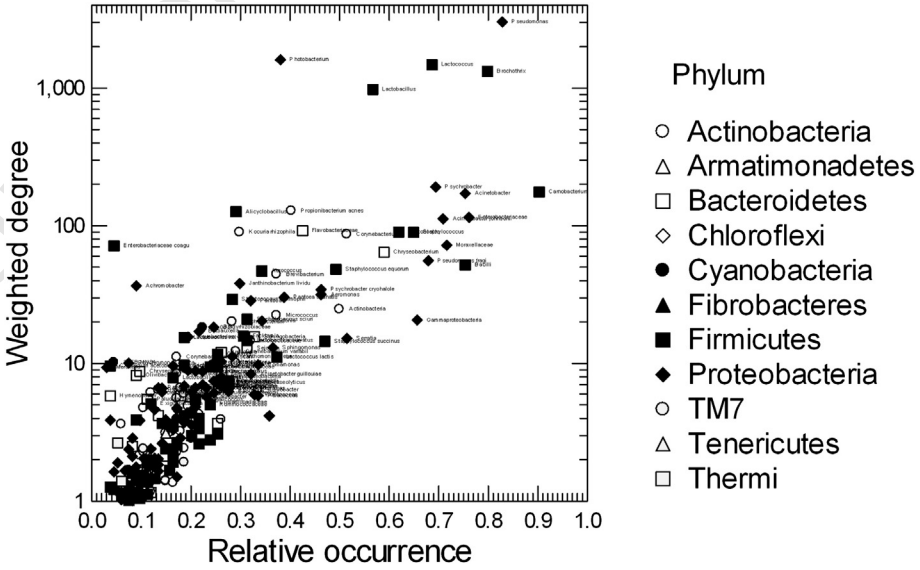

Fig. 3. Relative occurrence/weighted degree scatterplot for OTU nodes in raw meat samples (De Filippis et al., 2013). Only nodes with weighted degree > 1 are shown. Different symbols are used for members of different phyla and the identity of nodes with a weighted degree > 5 is shown.

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001

Table 3
Statistics for microbial interaction networks inferred for selected datasets in FoodMicrobionet. See Table 2 for definitions of network statistics.

Dataset	Reference	Samples	OTU	Nodes	Components	Diameter	Average degree	Average path length	Clustering coefficient	Clustering coefficient random network
Kefir	Marsh et al. (2013)	48	61	7	2	1	3.14	1	0.712	0.449
Mozzarella	Guidone et al. (2016)	29	77	22	2	5	4.18	2.19	0.453	0.19
Fontina, all samples milk and cheese	Dolci et al. (2014)	27	296	27	2	4	5.19	1.92	0.235	0.192
Fontinae, cheese only	Dolci et al. (2014)	18	158	11	2	4	2	2.95	0.152	0.181
NWC Natural whey culture (NWC) and curd	De Filippis et al., 2014	50	49	2	1	1	1	1	0	0
NWC and curd, Mozzarella	De Filippis et al. (2014)	24	34	2	1	1	1	1	0	0
NWC and curd, Grana type	De Filippis et al. (2014)	26	33	4	2	1	1	1	0	0
Piedmont hard cheese, all samples	Cocolin et al., unpublished	39	344	9	2	2	3.11	1.36	0.463	0.282
Beef, all samples	De Filippis et al. (2013)	108	827	118	1	5	35.15	1.83	0.544	0.297
Beef, spoiled samples	De Filippis et al. (2013)	45	108	11	3	5	1.46	2.35	0	0
Hamburger	Cocolin et al., unpublished	52	129	10	1	3	3	1.60	0.709	0.300
Fermented meat	Greppi et al. (2015)	30	74	12	1	4	3.67	1.86	0.462	0.305

or spoiled foods show the lowest complexity, and a high correlation ($r = 0.92$) was found between the number of OTUs detected in the dataset and the number of nodes in the interaction network. More complex networks, with >20 nodes were obtained when raw foods (milk, meat) were included in the dataset or when the dataset reflected different environments (milk and cheese, Dolci et al., 2014; Mozzarella produced with different acidification methods, Guidone et al., 2016; raw and spoiled meat, different cuts and samplings, De Filippis et al., 2013). In contrast, Deng et al. (2012) published figures on a wide range of complex bacterial interaction networks from environmental or human sources: the network size ranged from 107 to 254 nodes, the node degree distribution showed a good fit of the power law for all networks and the modularity (which measures the occurrence of modules which are strongly interconnected) was significantly higher than that of random networks, while the occurrence of a hierarchical structure was variable. Moreover, the bacterial interaction network of the human microbiome (Faust et al., 2012) included 197 phylotypes with 3005 significant interactions. The network showed a good fit of the power law model for the node degree distribution but did not show a strong hierarchical structure, although the occurrence of body site modules was found. Several factors may contribute to the lower complexity of microbial interaction networks for food. Food microbial communities, and fermented food communities in particular, are dramatically less complex than those found in environmental samples or in the human or animal microbiome, and therefore a lower number of sequences is generally sufficient to obtain a high coverage as it can be predicted by alpha rarefaction analyses (Ercolini, 2013). This may prevent the detection of significant interactions for minor OTUs. Finally, the interactions detected in this study mainly reflect co-occurrence and mutual exclusion patterns in different food environments, and although they in some cases may suggest true positive (commensalism, mutualism) or negative (competition, amensalism) interactions (Gram et al., 2002; Ivey et al., 2013), these should be confirmed in independent experiments.

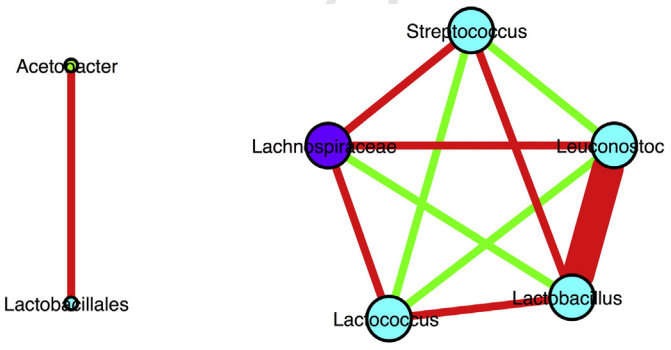

Fig. 4. Microbial interaction network for the kefir dataset (Marsh et al., 2013). Each node represents an OTU. Interactions were evaluated at different taxonomic levels. Only significant interactions are shown ($p < 0.0004$; $q < 4 \times 10^{-4}$). Edges showing negative interactions (co-exclusion) are coloured red, those for positive interactions in green. The colour of nodes corresponds to the class. The thickness of the edges reflects the level of significance of the supporting evidence for the association (as q-values, $0–4 \times 10^{-4}$), while the size of the nodes is proportional to their degree.

3.3. Future perspectives

Both the web visualisation and the full or filtered networks obtained from Gephi, although visually pleasing and informative, are somewhat naïve and should be interpreted with caution. The meta-analyses based on sequencing data published by different laboratories carry some inevitable bias due to differences in data generation and processing. These include possible differences from sample handling through nucleic acid extraction, variable 16S region chosen as target, library purification and preparation, sequencing technology and parameters, sequencing depth/sample coverage (Ercolini, 2013). Furthermore, it is important to underline that the exact bioinformatics path chosen for the analysis can have a strong impact too (May et al., 2014), and will have to be taken into account for a possible standardization of the data handling and usage. In addition, detection of rare OTUs might be affected by biases and reproducibility and repeatability issues (Benson et al., 2014; Guidone et al., 2016; Pinto and Raskin, 2012; Zhou et al., 2011b). To take this into account it may be advisable to compare different studies at a lower taxonomic resolution and exclude rare OTU from the comparisons. This can be easily done by processing data tables from FoodMicrobionet. An example of a filtered network is shown in Supplementary Fig. S8. In this case, OTUs belonging to the same genus were merged, and the interactive web visualisation is available at http://www.foodmicrobionet.org/fmbn1_0_3gweb/.

Therefore, for a future larger scale meta-analysis it would be advisable to, as a minimum requirement, process the sequences with the same standardized flow in order to limit at least the post-sequencing bias of the analysis. Unfortunately an optimized bioinformatics pipeline is still not available for food microbial communities. Most studies in FoodMicrobionet were carried out using the same sequencing platform and similar or identical bioinformatics pipelines (Table 1) and direct comparisons among studies can be carried out with a good degree of confidence.

With these limitations in mind, the approach used here provides an appealing means for microbiologists and food scientists dealing with food microbial community metadata analysis by (a) providing access to a large set of curated data on the occurrence of different taxa in foods most of which were obtained from studies published in peer reviewed journals, thus facilitating the process of formulating and

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001

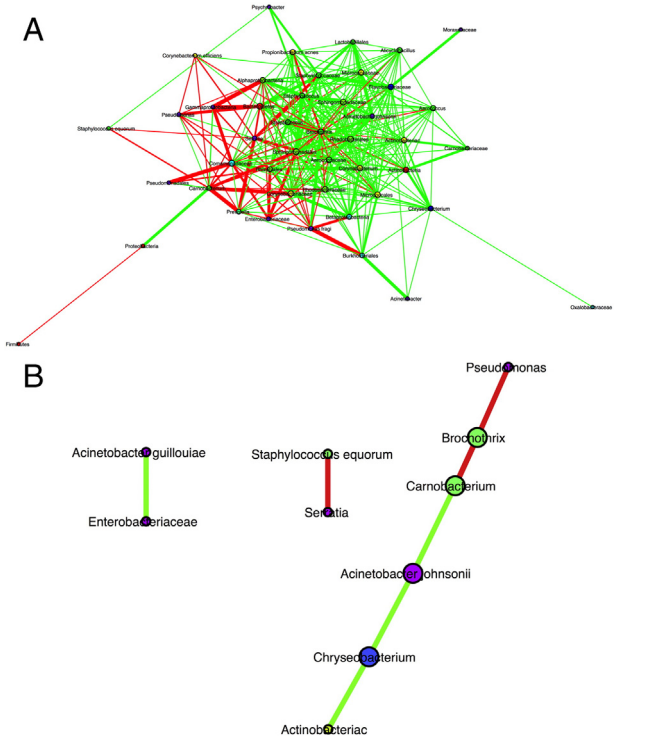

Fig. 5. Microbial interaction network for the beef dataset (De Filippis et al., 2013). A simplified network for all samples (non-spoiled and spoiled, A) and the full interaction network for spoiled beefsteaks (B) are shown. Each node represents an OTU (only low level taxa are shown). Interactions were evaluated at different taxonomic levels. Only significant interactions are shown ($p < 0.0004$; $q < 4 \times 10^{-4}$). Edges showing negative interactions (co-exclusion) are coloured red, those for positive interactions in green. The colour of nodes corresponds to the class. The thickness of the edges reflects the level of significance of the supporting evidence for the association (as q-values, $0–4 \times 10^{-4}$), while the size of the nodes is proportional to their degree. Actinobacteriac refers to the class *Actinobacteria*.

validating hypotheses on the structure and dynamics of food bacterial communities and writing original articles and reviews; (b) fostering open access to microbial ecology data by making curated nodes and edges tables publicly available; (c) improving our understanding of the ecology of spoilage-associated and beneficial microorganisms; and (d) providing information on the structure of bacterial communities in raw materials, fermented and spoiled foods which can be used for food process development.

While only part of this information is available through the online visualisation, the latter provides a simple interactive interface to explore the microbial ecology of the food environments included in FoodMicrobionet 1.0. Experienced users can import the nodes and edges files provided as supplementary material in a variety of spreadsheet, statistical or network analysis software packages to carry out graphical and statistical analyses or to generate their own networks.

Future plans include (a) expanding the network to other food matrices and food environments; (b) implementing an optimized data analysis pipeline to standardize the treatment of the raw data; and (c) the addition of metadata describing food properties in order to speculate on relevant ecological factors driving microbial interactions and to allow the selection of FoodMicrobionet subnetworks with defined range of specific ecological factors. Contributions from other research groups will be welcome. Details on the submission procedure are provided in section 3 of the Supplementary Materials.

Ultimately, FoodMicrobionet is a valuable resource which allows researchers in the food microbiology to benefit from the significant advances that HTS is providing in this key field of research.

Uncited reference

Mendes et al., 2014

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001.

References

Bastian, M.H.S., Jacomy, M., 2009. Gephi: An Open Source Software for Exploring and Manipulating Networks. International AAAI Conference on Weblogs and Social Media: San Jose, CA, USA.

Benson, A.K., David, J.R.D., Gilbreth, S.E., Smith, G., Nietfeldt, J., Legge, R., Kim, J., Sinha, R., Duncan, C.E., Ma, J., Singh, I., 2014. Microbial successions are associated with changes in chemical profiles of a model refrigerated fresh pork sausage during an 80-day shelf life study. Appl. Environ. Microbiol. 80, 5178–5194.

Bokulich, N.A., Mills, D.A., 2013. Facility-specific 'house' microbiome drives microbial landscapes of artisan cheesemaking plants. Appl. Environ. Microbiol. 79, 5214–5223.

Bokulich, N.A., Bamforth, C., Mills, D.A., 2012. Brewhouse-resident microbiota responsible for fermentation of American coolship ale. PLoS One 7, e35507.

Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Gonzalez Peña, A., 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.

Chaillou, S., Chaulot-Talmon, A., Caekebeke, H., Cardinal, M., Christieans, S., Denis, C., Desmonts, M.H., Dousset, X., Feurer, C., Hamon, E., Joffraud, J.-J., La Carbona, S., Leroi, F., Leroy, S., Lorre, S., Macé, S., Pilet, M.-F., Prévost, H., Rivollier, M., Roux, D., Talon, R., Zagorec, M., Champomier-Vergès, M.-C., 2015. Origin and ecological selection of core and food-specific bacterial communities associated with meat and seafood spoilage. ISME J. 9, 1105–1118.

Cocolin, L., Ercolini, D., 2015. Zooming into food-associated microbial consortia: a 'cultural' evolution. Curr. Opin. Food Sci. 2, 43–50.

Cocolin, L., Alessandria, V., Botta, C., Gorra, R., De Filippis, F., Ercolini, D., Rantsiou, K., 2013. NaOH-debittering induces changes in bacterial ecology during table olives fermentation. PLoS One 8, e69074.

De Filippis, F., La Storia, A., Stellato, G., Gatti, M., Ercolini, D., 2014. A selected core microbiome drives the early stages of three popular Italian cheese manufactures. PLoS One 9, e89680.

De Filippis, F., La Storia, A., Villani, F., Ercolini, D., 2013. Exploring the sources of bacterial spoilers in beefsteaks by culture-independent high-throughput sequencing. PLoS One 8, e70222.

De Pasquale, I., Calasso, M., Mancini, L., Ercolini, D., La Storia, A., De Angelis, M., Di Cagno, R., Gobbetti, M., 2014a. Causal relationship between microbial ecology dynamics and proteolysis during manufacture and ripening of protected designation of origin (PDO) cheese Canestrato Pugliese. Appl. Environ. Microbiol. 80, 4085–4094.

De Pasquale, I., Di Cagno, R., Buchin, S., De Angelis, M., Gobbetti, M., 2014b. Microbial ecology dynamics reveal a succession in the core microbiota involved in the ripening of Pasta Filata Caciocavallo Pugliese Cheese. Appl. Environ. Microbiol. 80, 6243–6255.

Deng, Y., Jiang, Y.-H., Yang, Y., He, Z., Luo, F., Zhou, J., 2012. Molecular ecological network analyses. BMC Bioinf. 13, 113.

Dolci, P., De Filippis, F., La Storia, A., Ercolini, D., Cocolin, L., 2014. rRNA-based monitoring of the microbiota involved in Fontina PDO cheese production in relation to different stages of cow lactation. Int. J. Food Microbiol. 185, 127–135.

Dunne, J.A., Williams, R.J., Martinez, N.D., 2002. Food-web structure and network theory: the role of connectance and size. Proc. Natl. Acad. Sci. U. S. A. 99, 12917–12922.

EFSA, 2011. The food classification and description system FoodEx 2 (draft-revision 1). Supporting publications 2011:215. http://www.efsa.europa.eu/en/supporting/pub/215e.htm.

Ercolini, D., 2013. High-throughput sequencing and metagenomics: moving forward in the culture-independent analysis of food microbial ecology. Appl. Environ. Microbiol. 79, 3148–3155.

Ercolini, D., De Filippis, F., La Storia, A., Iacono, M., 2012. "Remake" by high-throughput sequencing of the microbiota involved in the production of water buffalo mozzarella cheese. Appl. Environ. Microbiol. 78, 8142–8145.

Ercolini, D., Pontonio, E., De Filippis, F., Minervini, F., La Storia, A., Gobbetti, M., Di Cagno, R., 2013. Microbial ecology dynamics during rye and wheat sourdough preparation. Appl. Environ. Microbiol. 79, 7827–7836.

Faust, K., Raes, J., 2012. Microbial interactions: from networks to models. Nat. Rev. Microbiol. 10, 538–550.

Faust, K., Sathirapongsasuti, J.F., Izard, J., Segata, N., Gevers, D., Raes, J., Huttenhower, C., 2012. Microbial co-occurrence relationships in the human microbiome. PLoS Comput. Biol. 8, e1002606.

Ferrocino, I., Greppi, A., La Storia, A., Rantsiou, K., Ercolini, D., Cocolin, L., 2016. Impact of nisin-activated packaging on microbiota of beef burgers during storage. Appl. Environ. Microbiol. AEM.03093–15–36. http://dx.doi.org/10.1128/AEM.03093–15.

Glenn, T.C., 2011. Field guide to next generation DNA sequencers. Mol. Ecol. Resour. http://dx.doi.org/10.1111/j.1755-0998.2011.03024.x.

Gram, L., Ravn, L., Rasch, M., Bartholin Bruhn, J., Christensen, A.B., Givskov, M., 2002. Food spoilage—interactions between food spoilage bacteria. Int. J. Food Microbiol. 78, 79–97.

Greppi, A., Ferrocino, I., La Storia, A., Rantsiou, K., Ercolini, D., Cocolin, L., 2015. Monitoring of the microbiota of fermented sausages by culture independent rRNA-based approaches. Int. J. Food Microbiol. 212, 67–75.

Guidone, A., Zotta, T., Matera, A., Ricciardi, A., De Filippis, F., Ercolini, D., Parente, E., 2016. The microbiota of high-moisture mozzarella cheese produced with different acidification methods. Int. J. Food Microbiol. 216, 9–17.

Hu, Y., 2006. Efficient, high-quality force-directed graph drawing. Math. J. 10, 37–71.

Ivey, M., Massel, M., Phister, T.G., 2013. Microbial interactions in food fermentations. Ann. Rev. Food Sci. Technol. 4, 141–162.

Kuczynski, J., Liu, Z., Lozupone, C., McDonald, D., Fierer, N., Knight, R., 2010. Microbial community resemblance methods differ in their ability to detect biologically relevant patterns. Nat. Methods 7, 813–819.

Liu, J., Yang, H., Zhao, M., Zhang, X.-H., 2014. Spatial distribution patterns of benthic microbial communities along the Pearl estuary, China. Syst. Appl. Microbiol. 37, 578–589.

Marsh, A.J., O'Sullivan, O., Hill, C., Ross, R.P., Cotter, P.D., 2013. Sequencing-based analysis of the bacterial and fungal composition of Kefir grains and milks from multiple sources. PLoS One 8, e69371.

May, A., Abeln, S., Crielaard, W., Heringa, J., Brandt, B.W., 2014. Unraveling the outcome of 16S rDNA-based taxonomy analysis through mock data and simulations. Bioinformatics 30, 1530–1538.

Mayo, B., Rachid, C.T.C.C., Alegría, A., Leite, A.M.O., Peixoto, R.S., Delgado, S., 2014. Impact of next generation sequencing techniques in food microbiology. Curr. Genomics 15, 293–309.

Mendes, L.W., Kuramae, E.E., Navarrete, A.A., van Veen, J.A., Tsai, S.M., 2014. Taxonomical and functional microbial community selection in soybean rhizosphere. ISME J. 8, 1577–1587.

de Menezes, A.B., Prendergast-Miller, M.T., Richardson, A.E., Toscas, P., Farrell, M., Macdonald, L.M., Baker, G., Wark, T., Thrall, P.H., 2014. Network analysis reveals that bacteria and fungi form modules that correlate independently with soil parameters. Environ. Microbiol. 17, 2677–2689.

Mounier, J., Monnet, C., Vallaeys, T., Arditi, R., Sarthou, A.-S., Hélias, A., Irlinger, F., 2008. Microbial interactions within a cheese microbial community. Appl. Environ. Microbiol. 74, 172–181.

Muegge, B.D., Kuczynski, J., Knights, D., Clemente, J.C., Gonzalez, A., Fontana, L., Henrissat, B., Knight, R., Gordon, J.I., 2011. Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. Science 332, 970–974.

Newman, M., Barabási, A.-L., Watts, D.J., 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton.

Oakley, B.B., Morales, C.A., Line, J., Berrang, M.E., Meinersmann, R.J., Tillman, G.E., Wise, M.G., Siragusa, G.R., Hiett, K.L., Seal, B.S., 2013. The poultry-associated microbiome: network analysis and farm-to-fork characterizations. PLoS One 8, e57190.

Parente, E., Guidone, A., Matera, A., De Filippis, F., Mauriello, G., Ricciardi, A., 2016. Microbial community dynamics in thermophilic undefined milk starter cultures. Int. J. Food Microbiol. 217, 59–67.

Pinto, A.J., Raskin, L., 2012. PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets. PLoS One 7, e43093.

Rizzello, C.G., Cavoski, I., Turk, J., Ercolini, D., Nionelli, L., Pontonio, E., De Angelis, M., De Filippis, F., Gobbetti, M., Di Cagno, R., 2015. Organic cultivation of *Triticum turgidum* subsp. *durum* is reflected in the flour-sourdough fermentation-bread axis. Appl. Environ. Microbiol. 81, 3192–3204.

Schloss, P.D., Westcott, S.L., Ryabin, T., Hall, J.R., Hartmann, M., Hollister, E.B., Lesniewski, R.A., Oakley, B.B., Parks, D.H., Robinson, C.J., Sahl, J.W., Stres, B., Thallinger, G.G., Van Horn, D.J., Weber, C.F., 2009. Introducing MOTHUR: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75, 7537–7541.

Zhou, J., Deng, Y., Luo, F., He, Z., Yang, Y., 2011a. Phylogenetic molecular ecological network of soil microbial communities in response to elevated CO_2. mBio 2 (e00122–11).

Zhou, J., Wu, L., Deng, Y., Zhi, X., Jiang, Y.-H., Tu, Q., Xie, J., Van Nostrand, J.D., He, Z., Yang, Y., 2011b. Reproducibility and quantitation of amplicon sequencing-based detection. ISME J. 5, 1303–1313.

Please cite this article as: Parente, E., et al., FoodMicrobionet: A database for the visualisation and exploration of food bacterial communities based on network analysis, Int. J. Food Microbiol. (2015), http://dx.doi.org/10.1016/j.ijfoodmicro.2015.12.001