Shotgun metagenomics and volatilome profile of the microbiota of fermented sausages

Shotgun metagenomics and volatilome profile of the microbiota of fermented sausages

Ilario Ferrocino[1], Alberto Bellio[2], Manuela Giordano[1], Guerrino Macori[2], Angelo Romano[2], Kalliopi Rantsiou[1], Lucia Decastelli[2], Luca Cocolin[1*]

[1]DISAFA - Microbiology and food technology sector, University of Turin, Grugliasco (TO), [2]SC Controllo Alimenti e Igiene delle Produzioni, Istituto Zooprofilattico Sperimentale PVL, Torino, Italy

*Corresponding author: Luca Cocolin, lucasimone.cocolin@unito.it

Running title: Functional metagenomic and volatilome of fermented sausages

ABSTRACT

Changes in the microbial gene content and abundance can be analyzed to detect shifts in the microbiota composition due to the use of a starter culture in food fermentation process with the consequent shift of key metabolic pathways directly connected with product acceptance. Meat fermentation is a complex process involving microbes that metabolize the main components in meat. The breakdown of carbohydrates, proteins and lipids can lead to the formation of volatile organic compounds (VOCs) that can drastically affect the organoleptic characteristics of the final products. The present meta-analysis, performed with the shotgun DNA metagenomic approach, focuses on studying the microbiota and its gene content development in an Italian fermented sausage by using a commercial starter culture (a mix of *Lactobacillus sakei* and *Staphylococcus xylosus*), with the aim to discover the connection between microbiota and microbiome with the release of volatile metabolites during ripening.

The inoculated fermentation with the starter culture limited the development of *Enterobacteriaceae* and reduced the microbial diversity compared to spontaneous fermentation.

KEGG genes associated to the reduction of acetaldehyde to ethanol (EC:1.1.1.1), acetyl-P to acetate (EC:2.7.2.1), 2,3 butanediol to acetoin (EC:1.1.14) were found most abundant in inoculated samples (I) compared to spontaneous (S). The volatilome profiles were highly consistent with the genes abundance, elevated acetic acid (1173.85 μg/kg) ethyl acetate (251.58 μg/kg) and acetoin (1100.19 μg/kg) were observed in the presence of the starters at end of fermentation.

Significant differences were found in liking among samples based on flavor and odor suggesting a higher preference by consumers for the spontaneous samples. Inoculated samples exhibited the lowest scores for the liking data, clearly associated to the highest concentration of acetic acid.

IMPORTANCE

We present an advance in the understanding of meat fermentation by coupling the DNA-seq metagenomics and metabolomics approaches to describe the microbial function during this process.

Very few studies with this global approach have been dedicated to food but none in sausage fermentation underlying the originality of the study. The starter culture drastically affected the organoleptic properties of the products. This finding underlines the importance of starter culture selection that takes into consideration functional characteristics of the microorganism, in order to optimize production efficiency and product quality.

KEYWORDS Fermented sausages, Metabolic pathways, Shotgun metagenomics, volatile organic compounds (VOCs)

INTRODUCTION

During meat fermentation innumerable biochemical reactions take place that involve the breakdown of the main meat components (proteins, carbohydrates and lipids) and consequent conversion in metabolites that for the most part have an important organoleptic impact (1). In the last decades several culture-dependent and independent studies were carried out to describe the evolution of the microbiota in fermented meat and meat products (2–7). By using the recent 16S amplicon target sequencing it was found that meat can be contaminated by several microbial groups with spoilage potential coming from water, air, soil, from the workers and from the equipment involved in the processing or during the slaughtering procedures (8). Recently it was assessed that more than 30 different genera of Staphylococcaceae and Lactobacillaceae are present during the fermentation and ripening of fermented sausages (9) with the presence of several contaminant species, but only few of these taxa were recognized as metabolically active (10). More specifically the predominance of *L. sakei* during spontaneous fermentation was shown, usually associated with the presence of members of Leuconostaceae or with other bacteria (*L. curvatus*, *L. carnosum*, *S. xylosus*, *S. succinus*) (10). Development of coagulase-negative cocci (CNC) during meat fermentation can contribute to the proteolysis and lipolysis of meat components, while the lactic acid bacteria are responsible for the rapid decrease of pH, production of lactic acid, and production of small amounts of acetic acid, ethanol, acetoin, carbon dioxide and pyruvic acid (6, 9).

It has been recently shown that perturbation of the food system due to different ripening conditions (11), presence of starter culture, changes in the quality of the raw materials (initial microbiota composition) can change the genetic repertoire of the microbes. Changes in the composition of the microbiota have an impact on the sensorial characteristics of a product. Such changes may be due to differences in the abundance of genes encoding for enzymes that are involved in biochemical reactions leading to volatile compounds (11).

The evolution of the massive sequencing technologies like shotgun DNA-seq or RNA-seq can help to understand and characterize the microbial composition and function of the microbiota in

a food ecosystem. Unlike RNA-seq that only enables to profiling the transcriptome, shotgun DNA-seq offers the opportunity to concomitantly perform compositional analyses for the existing microbiota and gene pool (12). This technique is largely used in several environmental systems (human or agriculture) and only few studies have been applied on food to discover the presence of pathogens (13) and toxins (14), or to discover the gene content during food processing in vegetables (kefir grain (15), kimchi (16), soy (17)), broiler meat (18) and cheese (19).

The current study used the metagenomic DNA-seq sequence analyses to examine the microbiota composition and the gene content during ripening of the traditional Italian Felino sausages and to investigate how the use of starter culture can affect the bacterial gene abundance and the volatilome profile of the sausages.

RESULTS AND DISCUSSION

Microbial community profile. Microbial counts on De Man Rogosa Sharpe (MRS) agar and Mannitol Salt Agar (MSA) clearly showed the increase ($P < 0.05$) in the density of the lactic acid bacteria and Staphylococcaceae populations in the inoculated meat samples (I) as compared to the spontaneous fermentation (S). A significant reduction in the Enterobacteriaceae population ($P < 0.05$) was also observed early in the fermentation in the I samples (Table S1), likely due to the ability of the starter culture to lower the pH faster (Comparing S3 and I3 ($P < 0.05$), Table S1) and to compete with the indigenous microbiota.

Metagenomic shotgun sequencing data were used to estimate the relative abundance of microbial cells by mapping the non host clean reads against a set of clade-specific marker sequences by using MetaPhlAn2 that enables the estimation of relative abundance for individual species (20). Taxonomical assignment in I samples showed the dominance of *L. sakei* and *S. xylosus* followed in reduced proportion by *L. curvatus*, *S. equorum* and *Acinetobacter* sp. (Fig. 1).

The Spontaneous samples showed among the most abundant taxa the presence of *L. sakei* (varied from 37 to 56% of relative abundance) and *L. curvatus* (10 to 20%) followed by a number

of minor taxa identified as *S. xylosus*, *Leuconostoc* sp., *Lactococcus garvieae*, *Lactococcus lactis* as well as *Acinetobacter*, *Pseudomonas* and *Propionibacterium*.

The isolation from agar plates, identification and fingerprinting of LAB and staphylococci showed the presence of the *L. sakei* and *S. xylosus* in 90% of the colonies purified from the plate in both I and S samples with few strains belonging to *Weissella hellenica*, *Enterococcus faecalis* and *Kocuria* sp. The REP fingerprints analysis showed the ability of the starters to dominate the LAB and staphylococci microbiota since the beginning, while on the opposite several REP biotypes of the *L. sakei* and *S. xylosus* were found in the S samples (data not shown). The higher presence of sequences belonging to *L. sakei* and the finding of several REP-biotypes in the spontaneous fermentation clearly indicated that fermented meat is an ecological niche for *L. sakei*. The presence of several strains of *L. sakei* in the spontaneous samples most probably originated from the animal (21) and/or introduced during food processing (10).

Sequencing and assembly of the sausages metagenome. A total of 29,47 Gbp of raw reads were generated from 16 samples, which yielded 1,84 Gbp/sample. The quality reads after trimming was 22,04 Gbp of sequences. After host sequence removal, 8,59 Gbp of clean reads were analyzed and for each samples about ~2 Mbp to ~6 Mbp of reads were obtained (Table S2). Rarefaction curves obtained through MEGAN were used to determine genes abundance richness. These revealed that bacterial diversity was well represented as they are nearly parallel with the x-axis (Fig. S1), although Inoculated samples showed a lower genes abundance compared to Spontaneous fermentation in particular at time 0.

Sausage samples displayed higher proportions of reads of host origin (ranging from 43 to 87%), especially at the beginning of the fermentation. This is due to the higher abundance of mammalian cell and the low microbial biomass especially at the beginning of the ripening. Similar results have already been displayed in other food matrixes (18, 22).

A de novo performed assembly generated a total of 9,755 contigs of more than 1,000 bp in length, with an average N50 of 1,587 bp for Spontaneous fermentation and 83,842 bp for the Inoculated ones (Table S2). Consistent with the reduced microbiota diversity in I samples (Fig. 1), there were significantly fewer total genes predicted, and higher N50 values (in the assembled metagenomic data of I samples compared to S).

Exploration of metabolic potential of the sausages microbiota. To explore the metabolic potential we classified the predicted genes by aligning them to the integrated NCBI-NR database of non-redundant protein sequences. 11,402 predicted genes were identified, of which 10% were assigned to KEGG pathways by Megan, that compares genes using blastx and then assigns them to their latest common ancestors (LCA) of targeted organisms. The KEGG analysis assigned 1,774 genes to 21 pathways (Fig. 2) and the results gave a highly integrated picture of the global sausages microbiota metabolism. Consider the KEGG annotations at level 2, the KEGG categories Carbohydrate metabolism, Amino acid metabolism, Translation, Nucleotide metabolism (Fig. 2) were found as most abundant throughout the ripening period, consistently with those observed on a cheese matrix using a similar approach (19, 23). However only in Spontaneous sausages genes for carbohydrate and lipid metabolism increased significantly ($P < 0.05$) across the ripening time.

Within the Translation and Nucleotide metabolism categories, genes related to ribosomal protein and DNA polymerase were respectively the most abundant in all the datasets. The higher abundance genes related to Nucleotide metabolism is a consequence of the sugar consumption in meat during ripening and nucleoside metabolism that might improve LAB survival on meat during ripening after sugar consumption (24).

Taking into account the KEGG pathways related to carbohydrate, aminoacid and lipid categories, the pathway network (Fig. 3) shows the evolution of those pathways over time and condition. Sample node sizes are made proportional to a pathway's abundance in a given sample. In particular from the size of the edges, it was possible to visualize the relative abundance of numerous

functions related to energy metabolism, including all enzymes in the pathways related glycolysis/gluconeogenesis, pentose phosphate pathway, fructose and mannose metabolism, amino sugar and nucleotide sugar metabolism. These were most abundant after 3 days of fermentation in the inoculated samples. On the other hand the same pathways in the S samples were found to be most abundant after 7 days of fermentation; most likely due to the delayed evolution of the microbiota (lactic acid bacteria and Staphylococaceae) in the spontaneous as compared to the inoculated samples (Fig. 3).

Further taking into the account the KEGG gene differences between S and I samples were also observed by using the principal Component Analysis (Fig. S2). The PCA showed a separation of the metagenome content between spontaneous and inoculated samples. Genes involved in glycolysis were the most abundant and in particular acetaldehyde dehydrogenase had the highest number of sequence assignments in the entire dataset, followed by alcohol dehydrogenase, enolase, acetate kinase, phospho ketolase and glucono kinase.

Metagenomic content boosted the production of VOCs during the ripening time. Going more deeply into the metagenome content, the DESeq2 analysis identified 340 KEGG genes differentially abundant between Spontaneous fermented sausages and Inoculated ones (false-discovery rate [FDR], < 0.05) (Table S3). According to the DESeq2 analysis, the most prominent differences during fermentation among the two sausages types involved key KEGG genes in carbohydrate metabolism (pyruvate metabolism and glycolysis).

KEGG genes responsible for the reduction of acetaldehyde to ethanol (EC:1.1.1.1), acetyl-P to acetate (EC:2.7.2.1), 2,3 butanediol to acetoin (EC:1.1.14) were found most abundant in I samples compared to S (Fig. S3a). Along with it, elevated acetyl-P, acetaldehyde and 2,3 butanediol production increased acetic acid (1,173.85 μg/kg), ethyl acetate (251.58 μg/kg) and acetoin (1,100.19 μg/kg) during fermentation in the presence of starter ($P < 0.05$) as observed from GC analysis (Fig. 4 and Table S4).

Acetoin is the keton mostly found in fresh meat stored in different conditions and is referred as a product of the carbohydrate catabolism of LAB and staphylococci associated with cheesy odors of meat (25, 26). Moreover, even though we found a high presence of this molecule at the end of ripening in I samples (1,100 μg/kg, Table S4). Dainty et al. (27) have reported that the higher presence of acetoin in meat is not unpleasant. Carbohydrate metabolism was the most prevalent pathway in I samples most likely due to the predominant presence of the heterofermentative *L. sakei* (28) and carbohydrate pathways was found to be one of the main precursors of many VOCs such as: acetate, acetoin, diacetyl, acetic acid, iso-butyric acid (1).

The ability of the heterofermentative carbohydrate metabolism (alternative degradation of pyruvate) of the starter culture is well explained (29) and it was found that inoculated fermentation highest abundance of Acetate kinase (EC:2.7.2.1) that can lead to the formation of acetic acid, this compound is a typical aroma compound of dry fermented sausages (30). *L. sakei* utilizes glucose, fructose as well as several hexoses and amino acids as primary as energy sources during the initial growth stage (31). In particular sugars are fermented through different metabolic pathways: sugar hexose fermentation is homolactic and proceeds via the glycolytic pathway leading to lactate, whereas pentoses are fermented through the heterolactic phosphoketolase pathway ending with lactate and other end products such as acetate (32).

Glucose is the preferred carbon source for *L. sakei* (24) in meat under chilled storage and after its utilization several substrates are metabolized, such as lactate, gluconate, glucose-6-phospate, pyruvate, propionate, formate, ethanol, acetate, amino acids, nucleotides, etc. (1).

Regarding the evolution of VOCs through ripening (Fig. 5) we observed that I3 and I7 cluster together with S7 and display a significantly higher abundance ($P < 0.05$) of short chain esters such as ethyl acetate, ethyl 2-methylbutanoate, ethyl isovalerate and ethyl butanoate and some SCFA (acetic acid and butanoic acid) (Fig. 5 and Table S4). Samples at the end of ripening (I40 and S40) cluster together however I40 display a significantly higher abundance ($P < 0.05$) of ethyl alpha-hydroxybutyrate, Ethyl ester, 3-methyl-2-buten-1-ol and acetoin while S40 display the

higher presence of ethyl decanoate and 2-Heptanol ($P < 0.05$) (Fig. 5 and Table S4). This is in agreement with the finding that the Inoculated samples boosted the development of micro-organisms that can induce the high formation of ethanol and acetic acid at the beginning of ripening ($P < 0.05$). In addition the most abundant acetate kinase (EC:2.7.2.1) in I samples (false-discovery rate [FDR], < 0.05) may be involved in the interconversion of 2-oxobutanoate to propanoate from aminoacid metabolism (serine and aspartate) that can lead to the production of short chain volatile esters that we found most abundant in I samples compared to S (Fig. S3b). In detail it was shown that *L. sakei* is auxotrophic for all amino acids except for aspartic and glutamic acids and needs to absorb them after the aminoacid metabolism (33). Meat provides an environment rich in amino acids and *L. sakei* was displayed to be able to use it as an energy source. In particular inosine metabolism can lead to the formation of acetic acid as well as ethanol (34).

The indigenous microbiota of the S samples (*L. lactis*, *L. citreum*, *L. gelidum*, *S. xylosus* and *L. sakei*) displayed highter counts of KEGG genes involved in ex-novo fatty acids biosynthesis from pyruvate and aminoacid metabolism (Fig. S4). Consistently with this, a high production of long chain esters at the end of ripening such as ethyl octanoate and decanoate was found more abundant in S samples compared to I (Table S4) ($P < 0.05$), associated with the fruity and sweet odor description.

The differential abundance analysis showed that S samples displayed a high abundance KEGG genes compared to I samples (Table S3). In particular it was possible to find several proteases and amino acid catabolism encoding genes, but no clear association could be found between the genes and the volatilome profile. We also observed unusual functions of *L. sakei* in detail: key genes in folate synthesis dihydrofolate synthase/folylpolyglutamate synthase (EC:6.3.2.12 and EC:6.3.2.17) and formate tetrahydrofolate ligase (EC:6.3.4.3) associated with the indigenous presence of *L. sakei* were found in S samples. The ability of several strains of *L. sakei* to display genes related to folate biosynthesis was elsewhere assessed (16). In addition nucleotide metabolism serves as one of the important sources of energy for *L. sakei.* As recently pointed out

(24) this alternative energy route can be explained by the presence of unusual functions for a LAB because of the presence of a methylglyoxal synthase encoding gene [EC:4.2.3.3] that in this study we observed more abundant in I compared to S samples (Table S4) (FDR < 0.05).

Correlation between metagenome, volatilome and sensorial characteristics

Spearman's correlations (false-discovery rate [FDR], < 0.05) between KEGG pathways, taxa and volatilome (Fig. 6) clearly suggests that *L. sakei* was positively correlated to several pathways involved in carbohydrate and aminoacid metabolism: aminosugar, fructose, glycolysis, pentose phosphate pathway and valine leucine and isoleucine degradation as well as with abundance of acetoin, ethyl 2-methylbutanoate, 3-methyl-2-buten-1-ol. Branched-chain esters derived from amino acids like isoleucine and leucine that are precursors of the important aroma compounds (branched-chain alcohols, aldehydes, acids and their corresponding esters) and the majority of branched chain flavour compounds in sausages usually generated at the end of ripening and after the growth of the staphylococci has ceased (35, 36) due to the proteolitic activity of the LAB that after sugar consumption absorbs nutrients by hydrolyzing proteins present in the environment (37). In agreement with this we also found several association between LAB and esters compound. *L. lactis* was found positively correlated with galactose and butanoate metabolism as well as with 2-3-octanedione and ethyl-alpha-hydroxybutyrate (Fig. 6), *L. brevis* with ethyl esters, *L. citreum* with ethyl isovaletate and *Leuconostoc* sp. displayed the highest negative correlation with the volatile esters (ethyl butanoate, octanoate and pentanoate) as well as with the fatty acid metabolism. However esters are also formed through the esterification of alcohols (ethanol in particular) and carboxylic acids found in meat (38) following microbial esterase activity (30, 39).

The higher loads of the LAB and the finding of a higher presence of *L. sakei* in Inoculated fermentation, along with the concentration of volatile metabolites (in particular of the acids class), allowed a definite separation of the samples according to the presence of the starter culture. Significant differences were found in liking among samples based on flavor and odor (Fig. 7, $P <$

0.05) suggesting a higher preference by consumers for the S samples. The radar plot depicted in Fig 7 clearly showed that I samples exhibited the lowest scores for the liking data, clearly associated to the highest presence of acetic acid and its related odor descriptors: pungent, acidic, cheesy, vinegar (1).

CONCLUSION

In this study we present an integrated analysis of a typical Italian fermented sausage with a strict link between the volatilome profile, microbiota, gene content and consumers acceptability. A robust standard bioinformatics pipeline to process, annotate and realize the sausages gene catalog was assessed according to several pipelines in different environments. We found that the presence of the starter culture, in particular the presence of *L. sakei*, assure a fast and predominant growth, high acidification rate and fast consumption of fermentable substrates. A decline in numbers of Enterobacteriaceae was observed as well as the decline in microbial diversity. In addition the pH endpoint of sausages made with the addition of starter cultures was lower than the pH of the spontaneous sausages. On the opposite, starter cultures used in this study had a negative impact on the sensory properties of the product, as confirmed by the consumers test, due to the faster metabolic activity implied from metagenomic and confirmed by the meta-metabolomics data. Spontaneous fermented sausages made without the addition of starter cultures were found generally more acceptable and displayed a higher variety of genes with valuable potential (like the ones involved in the folate biosynthesis).

The multi-omics approach followed, in this case DNA-seq metagenomic coupled with metabolomics data, was effective in providing unprecedented insight into fermentation mechanisms that can affect the final characteristics of products.

MATERIALS AND METHODS

Sausages manufacturing and sample collection. Felino type sausages were manufactured in a local meat factory in the area of Torino according to the standard recipe. The formulation used

295 in the manufacture included pork meat (77%), lard (23%), salt (2.9%), spices (including pepper,
296 coriander, nutmeg and cinnamon 0.2%), sucrose (0.4% w/w), nitrate salt (E252, 0.01%) and wine
297 (0.3%). A commercial starter culture composed by *Lactobacillus sakei* and *Staphylococcus xylosus*
298 (SA8-400M, Veneto Agricoltura, Thiene, Vicenza, Italy) was added to the meat batter to reach the
299 final concentration of 5 log CFU/g. Meat batter was stuffed into synthetic casings resulting in 12
300 sausages of about 5 cm diameter and 500 g weight. Fermentation and ripening was carried out in a
301 climatic chamber; time and relative humidity/temperature conditions are reported in Table S5.
302 Another series of 12 sausages were prepared as described above, without using the starter culture
303 and used as a control. Three samples of the meat mixture prior to filling (0) and three sausages
304 samples obtained after 3, 7, and 40 days of fermentation/ripening were analyzed.
305 Two independent batches were analyzed for a total of 24 Inoculated (I) and 24 Spontaneous
306 sausages (S). Both batches were prepared with meat from the same meat factory in two different
307 periods of time, the second batch one week after the first. At each sampling point 3 sausages for
308 both conditions (I and S) were removed from the casing and individually placed in a stomacher bag
309 (Sto-circul-bag, PBI, Milan, Italy) and gently mixed. Aliquots were then used for microbial count,
310 pH and a_w determination, DNA extraction and Volatile Organic Compounds (VOCs) analysis.
311
312 **Microbiological analysis pH and a_w determination.** About 25 g from each of the three
313 sausages at every sampling time were then homogenized with 225 mL of Ringer's solution (Oxoid,
314 Milano, Italy) for 2 min in a stomacher (LAB Blender 400, PBI, Italy). Decimal dilutions in
315 quarter-strength Ringer's solution were prepared, and aliquots of 0.1 ml of the appropriate dilutions
316 were spread in triplicate on the following media: (i) gelatin peptone agar (GPA, Oxoid) for total
317 aerobic bacteria incubated for 48 to 72 h at 30 °C; (ii) De Man Rogosa and Sharpe agar (MRS,
318 Oxoid) for Lactobacillaceae and Lactococcaceae (LAB), incubated at 30 °C for 48 h; (iii) mannitol
319 salt agar (MSA, Oxoid) for Staphylococcaceae incubated at 30 °C for 48 h; (iv) violet red bile agar
320 (VRBA, Oxoid) for Enterobacteriaceae, incubated at 30 °C for 24–48 h. Results were calculated as

321 the means of log colony-forming units (CFU) for three independent determinations. The pH was
322 measured by immersing the pH probe of a digital pH meter (micropH2001, Crison, Barcelona,
323 Spain) in a diluted and homogenized sample containing 10 g of sausages and 90 ml of distilled
324 water. Water activity (a_w) was measured with a calibrated electric hygrometer (HygroLab, Rotronic,
325 Bassersdorf, Switzerland) according to the manufacturer's instructions. Fifteen colonies from MRS
326 and MSA agars at each sampling point were randomly isolated and purified. The purified isolates
327 were preliminarily characterized by Gram staining and microscopic observations as well as catalase
328 and oxidase reactions. Working cultures were maintained in brain heart infusion (BHI, Oxoid) or
329 MRS broth (Oxoid) with 25% glycerol and stored at −20 °C.
330
331 **Molecular typing by rep-PCR and RSA and cluster analysis and identification.** LAB
332 and *Staphylococcus* sp. isolates were subjected to DNA extraction as previously reported (10). The
333 molecular identification of the LAB isolates was performed by PCR 16S–23S rRNA gene spacer
334 analysis (RSA) and 16S rRNA gene sequencing. The RSA was carried out with primers G1
335 (GAAGTCGTAACAAGG) and L1 (CAAGGCATCCACCGT) under conditions reported by
336 Bautista-Gallego et al. (39). LAB isolates displaying the same RSA profile were then subjected to
337 identification. LAB and Staphylococcaceae molecular fingerprints were obtained by using repetitive
338 extragenic palindromic PCR (rep-PCR) with the primer (GTG)5 according to Iacumin et al. (41).
339 The rep-PCR profiles were normalised and cluster analysis was performed using Bionumerics
340 software (version 6.1, Applied Maths, Sint-Martens-Latem, Belgium). The dendrograms were
341 calculated on the basis of DICE coefficient of similarity with the unweighted pair group method
342 using arithmetic averages (UPGMA) clustering algorithm (42). For *Staphylococcus*, after cluster
343 analysis 2 isolates from each cluster at 70 % of similarity were selected and subjected to
344 identification. The identification of LAB and Staphylococcaceae was performed by amplifying the
345 16S rRNA gene. The oligonucleotide primers described by Weisburg et al. (43) FD1 (50-AGA GTT
346 TGA TCC TGG CTC AG-30) and RD1 (50-AAG GAG GTG ATC CAG CC-30) (*Escherichia coli*

positions 8–17 and 1540–1524, respectively) were used. PCR conditions were chosen according to Ercolini et al. (44). 16S rRNA amplicons were sent for sequencing to GATC-Biotech (Cologne, Germany). To determine the closest known relatives of the 16S rRNA gene sequences obtained, searches were performed in public data libraries (GenBank) with the Blast search program (http://www. ncbi.nlm.nih.gov/blast/).

Analysis of Volatile Organic Compounds. The volatile organic compounds (VOCs) in the sausages samples were extracted using headspace (HS) solid phase micro-extraction (SPME) and analyzed by gas chromatography/mass spectrometry (GC/MS). All samples were analyzed in triplicate. The analysis was conducted using a 20 mL vial filled with 3 g of sample to which was added 10 μL of 3-octanol in ultrapure water (323 ppb) and methyl caproate (3383 ppb) as an internal standard for ester chemical class. After an equilibration time of 5 min at 40 °C, the extraction was performed using the same temperature for 30 min with a 50/30 μm DVB/CAR/PDMS fiber (Supelco, Milan, Italy) with stirring (250 rpm) using an SPME autosampler (PAL System, CombiPAL, Switzerland). The fiber was desorbed at 260 for 1 min in splitless mode. GC/MS analysis was performed with a Shimadzu GC-2010 gas chromatograph equipped with a Shimadzu QP-2010 Plus quadruple mass spectrometer (Shimadzu Corporation, Kyoto, Japan) and a DB-WAXETR capillary column (30m × 0.25 mm, 0.25 μm film thickness, J&W Scientific Inc., Folsom, CA). The carrier gas (He) flow rate was 1 mL/min. The temperature program began at 40 °C for 5 min, and then the temperature was increased at a rate of 10 °C/min to 80 °C and 5 °C/min to 240 °C for 5 min. The injection port temperature was 260 °C, the ion source temperature was 240 °C and the interface temperature was 240 °C. The detection was carried out by electron impact mass spectrometry in total ion current mode, using ionization energy of 70 eV. The acquisition range was m/z 33–330. The identification of volatile compounds was confirmed by injection of pure standards and the comparison of their retention indices (a mixture of a homologous series of C5–C28 was used), MS data reported in the literature and in the database (http://

webbook.nist.gov/chemistry/). Compounds for which pure standards were not available were identified on the basis of mass spectra and retention indices available in the literature. Semi-quantitative data (μg/kg) were obtained by measuring the relative peak area of each identified compound in relation to that of the added internal standard.

DNA extraction, library preparation and sequencing. At each sampling point, 2 ml of the first 10-fold serial dilution was collected and directly centrifuged at maximum speed for 30 s. Total DNA was extracted from the pellet by using the MasterPure Complete DNA & RNA Purification kit (Illumina Inc, San Diego, CA) following the manufacturer's instructions. Three biological replicates (3 different sausages at each sampling point and for each condition) were subjected to DNA extraction and total DNA was pooled before further processing. The DNA was further purified using the Agencourt AMPure XP (Beckman Coulter, USA) according to the manufacturer's protocol. DNA quality was checked on the NanoDrop 2000c instrument (Thermo Scientific, USA) and quantified on Qubit 2.0 Fluorometer (Invitrogen, USA). Sequence libraries were fragmented and tagged with sequencing adapters by using the Nextera XT library preparation kit (Illumina) following the manufacturer's instruction. The libraries were quantified using Qubit 2.0 Fluorometer. The quality and the size distribution of the libraries were determined by using High Sensitivity DNA chips and DNA Reagents on BioAnalyzer 2100 (Agilent, USA). Sequencing was performed in the MiSeq (Illumina) system for a 151 cycle paired-end run by sequencing 4 sample/run. Base calling and Illumina barcode demultiplexing processes were performed by the MiSeq Control Software V2.3.0.3, the RTA v1.18.42.0 and the CASAVA v1.8.2.

Bioinformatics and data analysis. Raw reads quality (Phred scores) was evaluated by using the FastQC toolkit (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The raw sequences were trimmed to Qscore of 30 with the SolexaQA++ software (45). Duplicate sequences and sequences less than 50 bp were discarded by using Prinseq (46). The obtained trimmed reads

were then mapped to the draft genome of *Sus Scrofa* (ftp://ftp.ncbi.nlm.nih.gov/genomes/Sus_scrofa/Assembled_chromosomes/) by using Bowtie2 (47) in end-to-end, sensitive mode. Clean, non host reads were then assembled with Velvet (48) with a minimum contig length set at 600 bp. Each contig was run through an automated gene annotation pipeline utilizing MetaGeneMark (49). Predicted genes for all the samples were concatenated and clustered using USEARCH (50) with the following criterion: identity ≥ 95% and alignment length > 90%. The sausages gene catalogs obtained were then aligned against the NCBI-NR database by using mblastx (51) in order to obtain the gene annotation. Clean reads were mapped back by using Bowtie2 to the annotated gene catalogue to enable semi-quantitative analysis and to check the quality of the assembly. The number of reads uniquely mapped to each gene (SAM file) were then used for functional analysis based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) (52) version April 2011 using MEGAN software version 5 (53). MEGAN specifically requires BLAST searches of nucleotide or protein sequences as input. The software then parses the BLAST results using the NCBI Refseq IDs to assign peptides to KEGG pathways. The KEGG classification in MEGAN is represented by a rooted tree (with approximately 13 000 nodes) whose leaves represent different pathways.

The functional gene count table was internally normalized in MEGAN by checking the "use normalized count" option. Rarefaction analysis was performed based on the leaves of the tree in MEGAN. The phylogenetic characterization of the shotgun sequences was achieved at species levels of taxonomy by using MetaPhlAn2 (20) with default parameters. Statistical analysis and plotting were carried out in R environment. Data normalization and determination of differentially abundant genes were then conducted using the Bioconductor DESeq2 package (54) in the statistical environment R. P-values were adjusted for multiple testing using the Benjamini-Hochberg procedure, which assesses the False Discovery Rate (FDR).

Pairwise Spearman's correlations between taxa, KEGG genes and volatile organic compounds were assessed by the R package psych and the significant correlation (false-discovery

rate [FDR], < 0.05) were plotted in a correlative network by using Cytoscape v. 2.8.143. Principal Component Analysis (PCA) and Hierarchical Clustering were carried out by using the made4 package in R. All the results are reported as mean values of the two batches.

Liking test. To assess the sensory acceptability of sausages samples at the end of the ripening, a consumer test was performed. A total of 15 regular consumers of sausages (7 males, 8 females, 28–56 years) voluntarily participated in the sensory evaluation. Sausages samples (10 g) were served under blind conditions in opaque white plastic cups and coded with a random three-digit number. Samples were served in completely randomized order, with the spontaneous fermented sausages served as the last sample for all subjects to limit the contrast effect. Consumers were asked to observe its appearance, smell and taste and rate the sausages for appearance, odor, taste, flavor, texture and overall acceptance. Liking was expressed on a 9-point hedonic scale ranging from “dislike extremely” (1) to “like extremely” (9). Purchase interest (Would you buy this sausages?) was also rated on a 7-point scale (1 = absolutely no, 7 = absolutely yes). Participants were required to rinse their mouth with still water for about 1 min between samples. Consumers took between 15 and 20 min to complete the evaluation. Liking data (appearance, odor, taste, flavor, texture and overall acceptance) and declared purchase interest from consumers were independently submitted to a Pairwayse Kruskal-Wallis Test on R enviromnment assuming samples as main factors.

Accession number. The raw reads data was deposited at the Sequence Read Archive of NCBI (accession number SRP092525)

Acknowledgements

We thank Angelica Laera for help in samples preparation and microbiological analysis.

Reference

1. **Casaburi A**, **Piombino P**, **Nychas G-J**, **Villani F**, **Ercolini D**. 2015. Bacterial populations and the volatilome associated to meat spoilage. Food Microbiol **45**:83–102.
2. **Cenci-Goga BT**, **Ranucci D**, **Miraglia D**, **Cioffi A**. 2008. Use of starter cultures of dairy origin in the production of Salame nostrano, an Italian dry-cured sausage. Meat Sci **78**:381–90.
3. **Mauriello G**, **Casaburi A**, **Blaiotta G**, **Villani F**. 2004. Isolation and technological properties of coagulase negative staphylococci from fermented sausages of Southern Italy. Meat Sci **67**:149–58.
4. **Cocolin L**, **Dolci P**, **Rantsiou K**, **Urso R**, **Cantoni C**, **Comi G**. 2009. Lactic acid bacteria ecology of three traditional fermented sausages produced in the North of Italy as determined by molecular methods. Meat Sci **82**:125–32.
5. **Blaiotta G**, **Pennacchia C**, **Villani F**, **Ricciardi A**, **Tofalo R**, **Parente E**. 2004. Diversity and dynamics of communities of coagulase-negative staphylococci in traditional fermented sausages. J Appl Microbiol **97**:271–84.
6. **Francesca N**, **Sannino C**, **Moschetti G**, **Settanni L**. 2012. Microbial characterisation of fermented meat products from the Sicilian swine breed "Suino Nero Dei Nebrodi." Ann Microbiol **63**:53–62.
7. **Kesmen Z**, **Yetiman A**, **Gulluce A**, **Kacmaz N**, **Sagdic O**, **Cetin B**, **Adiguzel A**, **Sahin F**, **Yetim H**. 2012. Combination of culture-dependent and culture-independent molecular methods for the determination of lactic microbiota in sucuk. Int J Food Microbiol **153**:428–35.
8. **Stellato G**, **La Storia A**, **De Filippis F**, **Borriello G**, **Villani F**, **Ercolini D**. 2016. Overlap of spoilage-associated microbiota between meat and the meat processing environment in small-scale and large-scale retail. Appl Environ Microbiol **82**:4045–4054.
9. **Połka J**, **Rebecchi A**, **Pisacane V**, **Morelli L**, **Puglisi E**. 2015. Bacterial diversity in typical Italian salami at different ripening stages as revealed by high-throughput sequencing of 16S rRNA amplicons. Food Microbiol **46**:342–356.
10. **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.
11. **De Filippis F, Genovese A, Ferranti P, Gilbert JA, Ercolini D**. 2016. Metatranscriptomics reveals temperature-driven functional changes in microbiome impacting cheese maturation rate. Sci Rep **6**:1–12.

12. **Sharpton TJ**. 2014. An introduction to the analysis of shotgun metagenomic data. Front Plant Sci **5**:1–14.
13. **Yang X**, **Noyes NR**, **Doster E**, **Martin JN**, **Linke LM**, **Magnuson RJ**, **Yang H**, **Georgnaras I**, **Woerner DR**, **Jones KL**. 2016. Use of metagenomic shotgun sequencing technology to detect foodborne pathogens within the microbiome of the beef production chain. Appl Environ Microbiol **82**:2433–2443.
14. **Leonard SR**, **Mammel MK**, **Lacher DW**, **Elkins CA**. 2015. Application of metagenomic sequencing to food safety: detection of shiga toxin-producing *Escherichia coli* on fresh bagged spinach. Appl Environ Microbiol **81**:8183–8191.
15. **Nalbantoglu U**, **Cakar A**, **Dogan H**, **Abaci N**, **Ustek D**, **Sayood K**, **Can H**. 2014. Metagenomic analysis of the microbial community in kefir grains. Food Microbiol **41**:42–51.
16. **Jung JY**, **Lee SH**, **Kim JM**, **Park MS**, **Bae J**, **Hahn Y**, **Madsen EL**, **Jeon CO**. 2011. Metagenomic analysis of kimchi, a traditional korean fermented food. Appl Environ Microbiol **77**:2264–2274.
17. **Sulaiman J**, **Gan HM**, **Yin WF**, **Chan KG**. 2014. Microbial succession and the functional potential during the fermentation of Chinese soy sauce brine. Front Microbiol **5**:1–9.
18. **Nieminen TT**, **Koskinen K**, **Laine P**, **Hultman J**, **Säde E**, **Paulin L**, **Paloranta A**, **Johansson P**, **Björkroth J**, **Auvinen P**. 2012. Comparison of microbial communities in marinated and unmarinated broiler meat by metagenomics. Int J Food Microbiol **157**:142–149.
19. **Dugat-Bony E**, **Straub C**, **Teissandier A**, **Onésime D**, **Loux V**, **Monnet C**, **Irlinger F**, **Landaud S**, **Leclercq-Perlat M-N**, **Bento P**, **Fraud S**, **Gibrat JF**, **Aubert J**, **Fer F**, **Guédon E**, **Pons N**, **Kennedy S**, **Beckerich JM**, **Swennen D**, **Bonnarme P**. 2015. Overview of a surface-ripened cheese community functioning by meta-omics analyses. PLoS One **10**:e0124360.
20. **Segata N**, **Waldron L**, **Ballarini A**, **Narasimhan V**, **Jousson O**, **Huttenhower C**. 2012. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods **9**:811–4.
21. **Castellano P**, **Aristoy MC**, **Sentandreu MA**, **Vignolo G**, **Toldrá F**. 2012. *Lactobacillus sakei* CRL1862 improves safety and protein hydrolysis in meat systems. J Appl Microbiol **113**:1407–1416.
22. **Jung JY**, **Lee SH**, **Jin HM**, **Hahn Y**, **Madsen EL**, **Jeon CO**. 2013. Metatranscriptomic analysis of lactic acid bacterial gene expression during kimchi fermentation. Int J Food Microbiol **163**:171–179.

519 23. **Lessard MH**, **Viel C**, **Boyle B**, **St-Gelais D**, **Labrie S**. 2014. Metatranscriptome analysis of
520 fungal strains Penicillium camemberti and Geotrichum candidum reveal cheese matrix
521 breakdown and potential development of sensory properties of ripened Camembert-type
522 cheese. BMC Genomics **15**:235.
523 24. **Chaillou S**, **Champomier-Vergès MC**, **Cornet M**, **Crutz-Le Coq AM**, **Dudez AM**,
524 **Martin V**, **Beaufils S**, **Darbon-Rongère E**, **Bossy R**, **Loux V**, **Zagorec M**. 2005. The
525 complete genome sequence of the meat-borne lactic acid bacterium *Lactobacillus sakei* 23K.
526 Nat Biotechnol **23**:1527–1533.
527 25. **Ardö Y**. 2006. Flavour formation by amino acid catabolism. Biotechnol Adv **24**:238–242.
528 26. **Leroy S**, **Vermassen A**, **Ras G**, **Talon R**. 2017. Insight into the genome of *Staphylococcus*
529 *xylosus*, a ubiquitous species well adapted to meat products. Microorganisms **5**:52.
530 27. **Dainty RH**, **Edwards RA**, **Hibbard CM**. 1989. Spoilage of vacuum-packed beef by
531 aclostridium sp. J Sci Food Agric **49**:473–486.
532 28. **Xu H**, **Gao L**, **Jiang Y**, **Tian Y**, **Peng J**, **Xa Q**, **Chen Y**. 2015. Transcriptome response of
533 *Lactobacillus sakei* to meat protein environment. J Basic Microbiol **55**:490–499.
534 29. **Freiding S**, **Gutsche KA**, **Ehrmann MA**, **Vogel RF**. 2011. Genetic screening of
535 *Lactobacillus sakei* and *Lactobacillus curvatus* strains for their peptidolytic system and
536 amino acid metabolism, and comparison of their volatilomes in a model system. Syst Appl
537 Microbiol **34**:311–20.
538 30. **Talon R**, **Chastagnac C**, **Vergnais L**, **Montel MC**, **Berdagué JL**. 1998. Production of
539 esters by Staphylococci. Int J Food Microbiol **45**:143–150.
540 31. **Lee SB**, **Rhee YK**, **Gu EJ**, **Kim DW**, **Jang GJ**, **Song SH**, **Lee JI**, **Kim BM**, **Lee HJ**, **Hong**
541 **HD**, **Cho CW**, **Kim HJ**. 2017. Mass-based metabolomic analysis of *Lactobacillus sakei* and
542 its growth media at different growth phases. J Microbiol Biotechnol **27**:925–932.
543 32. **Stentz R**, **Cornet M**, **Chaillou S**, **Zagorec M**. 2001. Adaptation of *Lactobacillus sakei* to
544 meat: a new regulatory mechanism of ribose utilization? Lait **81**:131–138.
545 33. **Champomier-Vergès MC**, **Maguin E**, **Mistou MY**, **Anglade P**, **Chich JF**. 2002. Lactic
546 acid bacteria and proteomics: Current knowledge and perspectives. J Chromatogr B Anal
547 Technol Biomed Life Sci **771**:329–342.
548 34. Rimaux T, Vrancken G, Vuylsteke B, De Vuyst L, Leroy F. 2011. The pentose moiety of
549 adenosine and inosine is an important energy source for the fermented-meat starter culture
550 *Lactobacillus sakei* CTC 494. Appl Environ Microbiol **77**:6539–6550.
551 35. **Sunesen LO**, **Dorigoni V**, **Zanardi E**, **Stahnke L**. 2001. Volatile compounds released
552 during ripening in Italian dried sausage. Meat Sci **58**:93–97.

553 36. **Olesen PT**, **Stahnke LH**. 2004. The influence of environmental parameters on the
554 catabolism of branched-chain amino acids by *Staphylococcus xylosus* and *Staphylococcus*
555 *carnosus*. Food Microbiol **21**:43–50.
556 37. **Kenneally PM**, **Fransen NG**, **Grau H**, **O'Neill EE**, **Arendt EK**. 1999. Effects of
557 environmental conditions on microbial proteolysis in a pork myofibril model system. J Appl
558 Microbiol **87**:794–803.
559 38. **Peterson RJ**, **Chang SS**. 1982. Identification of volatile flavor compounds of fresh, frozen
560 beef stew and a comparison of these with those of canned beef stew. J Food Sci **47**:1444–
561 1448.
562 39. **Toldra´ F**. 1998. Proteolysis and lipolysis in flavour development of dry-cured meat
563 products. Meat Sci **49**:S101–S110.
564 40. **Bautista-Gallego J**, **Alessandria V**, **Fontana M**, **Bisotti S**, **Taricco S**, **Dolci P**, **Cocolin L**,
565 **Rantsiou K**. 2014. Diversity and functional characterization of *Lactobacillus* spp. isolated
566 throughout the ripening of a hard cheese. Int J Food Microbiol **181**:60–6.
567 41. **Iacumin L**, **Comi G**, **Cantoni C**, **Cocolin L**. 2006. Molecular and technological
568 characterization of *Staphylococcus xylosus* isolated from naturally fermented Italian sausages
569 by RAPD, Rep-PCR and Sau-PCR analysis. Meat Sci **74**:281–288.
570 42. **Vauterin L**, **Vauterin P**. 1992. Computer-aided objective comparison of electrophoresis
571 patterns for grouping and identification of microorganisms. Eur J ClinMicrobiol **1**:37–41.
572 43. **Weisburg WG**, **Barns SM**, **Pelletier DA**, **Lane DJ**. 1991. 16S Ribosomal DNA
573 amplification for phylogenetic study. J Bacteriol **173**:697–703.
574 44. **Ercolini D**, **Russo F**, **Ferrocino I**, **Villani F**. 2009. Molecular identification of mesophilic
575 and psychrotrophic bacteria from raw cow's milk. Food Microbiol **26**:228–231.
576 45. **Cox MP**, **Peterson DA**, **Biggs PJ**. 2010. SolexaQA: At-a-glance quality assessment of
577 Illumina second-generation sequencing data. BMC Bioinformatics **11**:485.
578 46. **Schmieder R**, **Edwards R**. 2011. Quality control and preprocessing of metagenomic
579 datasets. Bioinformatics **27**:863–864.
580 47. **Langmead B**, **Salzberg SL**. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods
581 **9**:357–359.
582 48. **Zerbino DR**, **Birney E**. 2008. Velvet: Algorithms for de novo short read assembly using de
583 Bruijn graphs. Genome Res **18**:821–829.
584 49. **Zhu W**, **Lomsadze A**, **Borodovsky M**. 2010. Ab initio gene identification in metagenomic
585 sequences. Nucleic Acids Res **38**:1–15.
586 50. **Edgar RC**. 2010. Search and clustering orders of magnitude faster than BLAST.

Bioinformatics **26**:2460–2461.

51. **Davis C**. 2013. mBLAST: Keeping up with the sequencing explosion for (meta) genome analysis. J Data Mining Genomics Proteomics **4:** 135.

52. **Kanehisa M**, **Goto S**, **Sato Y**, **Kawashima M**, **Furumichi M**, **Tanabe M**. 2014. Data, information, knowledge and principle: Back to metabolism in KEGG. Nucleic Acids Res **42**:199–205.

53. **Mitra S**, **Stärk M**, **Huson DH**. 2011. Analysis of 16S rRNA environmental sequences using MEGAN. BMC Genomics **12 3**:S17.

54. **Love M**, **Huber W, Anders S**. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **15**(12):1

598 **Figure legends**

599 **Fig 1** Taxonomy analysis of fermented sausages. The plot shows the distribution of taxa during the
600 ripening time. Only taxa with an incidence above 0.5% in at least 2 samples are shown. Samples are
601 labeled according to time (0, 3, 7, and 40 days) and type (S; spontaneous; I; inoculated)

602

603 **Fig 2** Functional classification of fermented sausages during ripening. Functional classes were
604 determined according to the first level of the KEGG annotations. Samples are color coded
605 according to time (0, 3, 7, and 40 days) and type (S; spontaneous; I; inoculated)

606

607 **Fig 3** Relationships between metabolic pathways and samples. KEGG network summarizing the
608 relationships between metabolic pathways related to carbohydrates (red), aminoacids (yellow) and
609 lipids (blue) and samples (Cyan: Inoculated and green: Spontaneous). Metabolic pathways and
610 samples are connected with a line (i.e. edge) and its thickness is made proportional to the
611 abundance of that pathway in the connected sample.

612

613 **Fig 4** Abundance of VOCs during ripening. Acetic acid, acetoin and ethyl acetate concentration
614 (µg/kg) over time (0, 3, 7 and 40 days) and for fermentation conditions (Red: Inoculated; Blue:
615 Spontaneous). Boxes represent the interquartile range (IQR) between the first and third quartiles,
616 and the line inside represents the median (2nd quartile). Whiskers denote the lowest and the highest
617 values within IQR from the first and third quartiles, respectively. Circles represent outliers beyond
618 the whiskers.

619

620 **Fig 5** Correlation patterns between VOCs and samples. Correlation between the abundance of
621 VOCs and Spontaneous (S) and Inoculated (I) samples. Rows and columns are clustered by Ward
622 linkage hierarchical clustering. The intensity of the colors represents the degree of correlation
623 between the samples and VOCs as measured by the Spearman's correlations.

Fig 6 Correlation between taxa, ripening-related metabolic pathways and volatilome data. Correlation network showing significant (false-discovery rate [FDR], < 0.05) Spearman's correlations between KEGG genes belonging to amino acid and lipid metabolism, VOCs, and taxa. Node size was made proportional to the number of significant correlations. Edge color indicates negative (blue) or positive (pink) correlations.

Fig 7. Liking test. Plot A: Radar graph displaying the liking of appearance, odor, taste, flavor, texture and overall liking expressed by consumers for the spontaneous and inoculated fermentation. Plot B: Distribution of the liking scores of Flavour and Odour ($P < 0.05$) for fermentation conditions (Red: Inoculated; Blue: Spontaneous fermentation). Boxes represent the interquartile range (IQR) between the first and third quartiles, and the line inside represents the median (2nd quartile). Whiskers denote the lowest and the highest values within IQR from the first and third quartiles, respectively. Circles represent outliers beyond the whiskers.

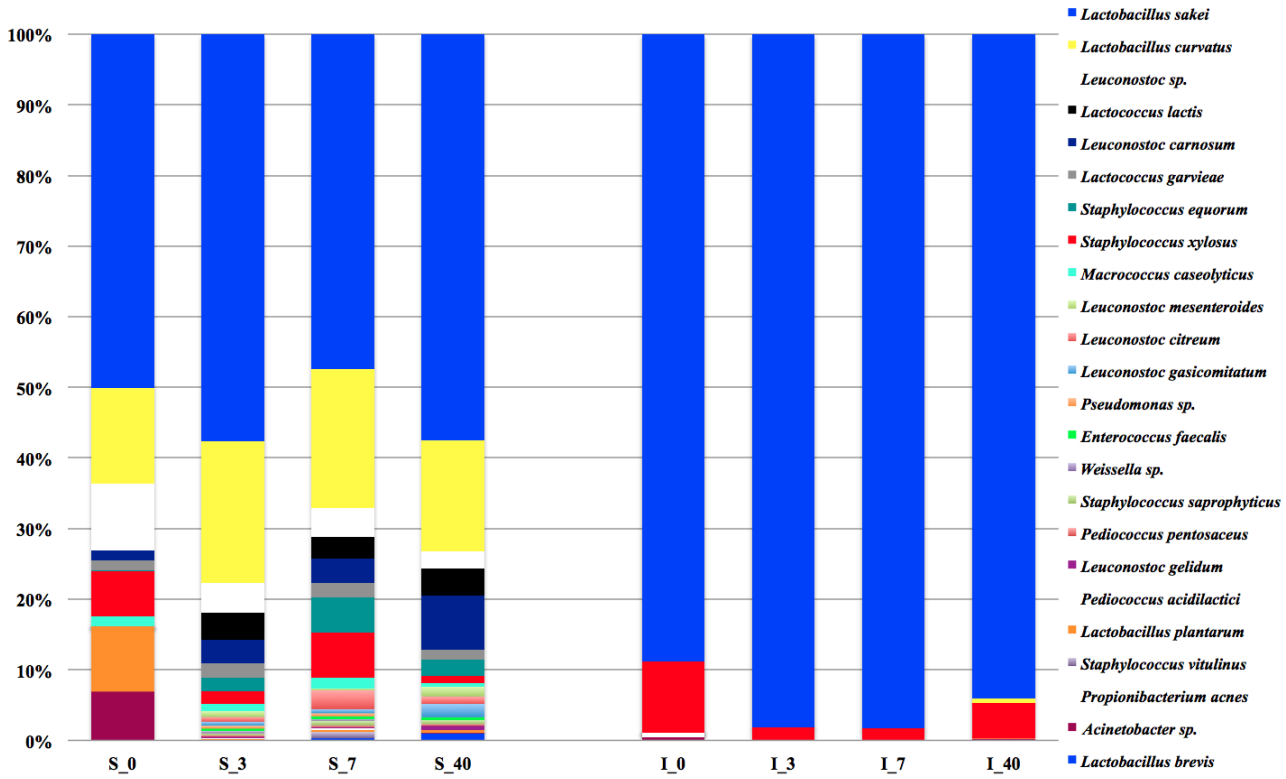

100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
S_0
S_3
S_7
S_40
I_0
I_3
I_7
I_40
Lactobacillus sakei
Lactobacillus curvatus
Leuconostoc sp.
Lactococcus lactis
Leuconostoc carnosum
Lactococcus garvieae
Staphylococcus equorum
Staphylococcus xylosus
Macrococcus caseolyticus
Leuconostoc mesenteroides
Leuconostoc citreum
Leuconostoc gasicomitatum
Pseudomonas sp.
Enterococcus faecalis
Weissella sp.
Staphylococcus saprophyticus
Pediococcus pentosaceus
Leuconostoc gelidum
Pediococcus acidilactici
Lactobacillus plantarum
Staphylococcus vitulinus
Propionibacterium acnes
Acinetobacter sp.
Lactobacillus brevis

Fig 1

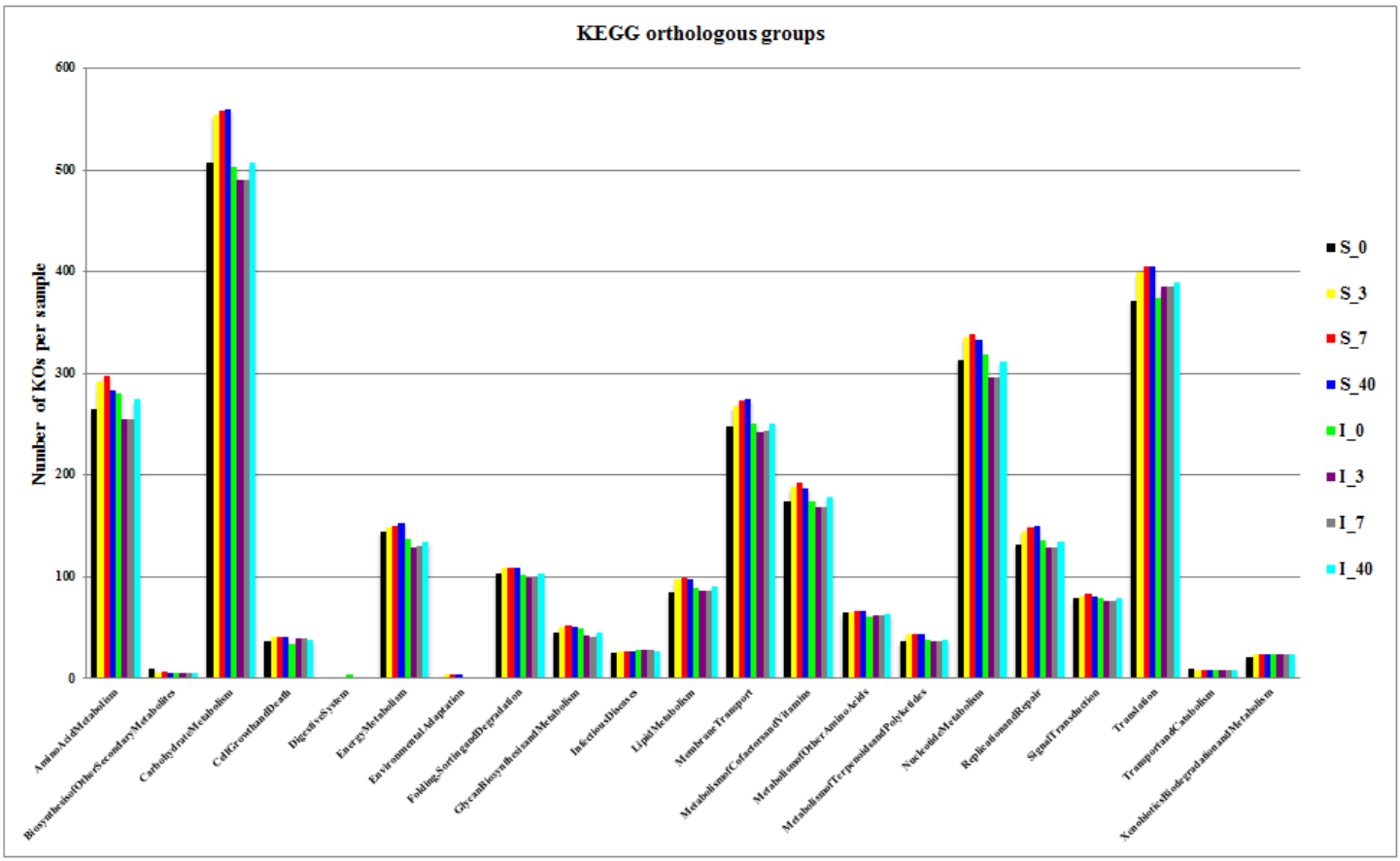

KEGG orthologous groups
Number of KOs per sample
600
500
400
300
200
100
0
S_0
S_3
S_7
S_40
I_0
I_3
I_7
I_40
AminoAcidMetabolism
BiosynthesisofOtherSecondaryMetabolites
CarbohydrateMetabolism
CellGrowthandDeath
DigestiveSystem
EnergyMetabolism
EnvironmentalAdaptation
Folding,SortingandDegradation
GlycanBiosynthesisandMetabolism
InfectiousDiseases
LipidMetabolism
MembraneTransport
MetabolismofCofactorsandVitamins
MetabolismofOtherAminoAcids
MetabolismofTerpenoidsandPolyketides
NucleotideMetabolism
ReplicationandRepair
SignalTransduction
Translation
TransportandCatabolism
XenobioticsBiodegradationandMetabolism

Fig. 2

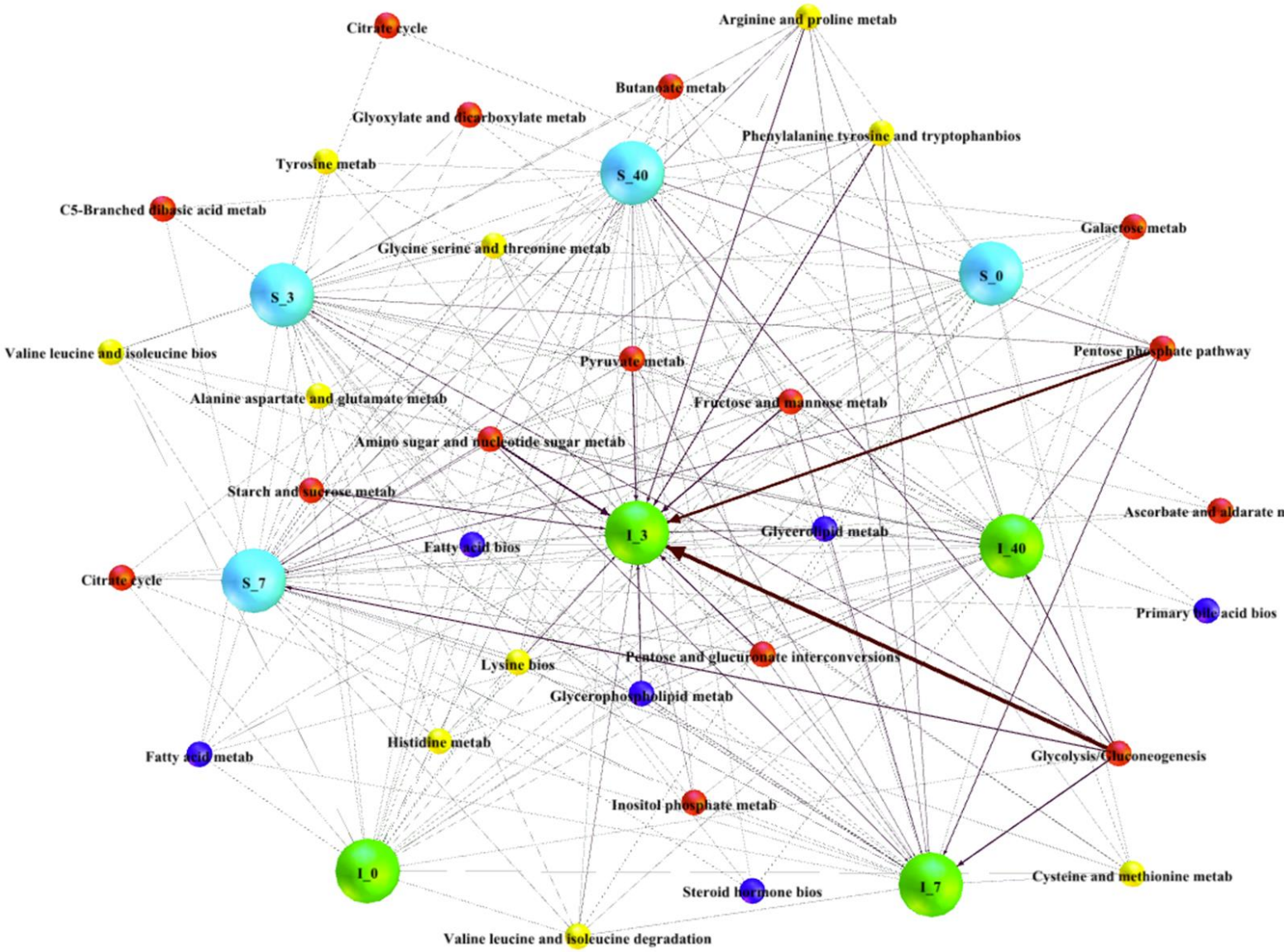

Citrate cycle
Arginine and proline metab
Butanoate metab
Glyoxylate and dicarboxylate metab
Phenylalanine tyrosine and tryptophanbios
Tyrosine metab
S_40
C5-Branched dibasic acid metab
Galactose metab
Glycine serine and threonine metab
S_0
S_3
Valine leucine and isoleucine bios
Pentose phosphate pathway
Pyruvate metab
Alanine aspartate and glutamate metab
Fructose and mannose metab
Amino sugar and nucleotide sugar metab
Starch and sucrose metab
Ascorbate and aldarate me
I_3
Glycerolipid metab
Fatty acid bios
I_40
Citrate cycle
S_7
Primary bile acid bios
Pentose and glucuronate interconversions
Lysine bios
Glycerophospholipid metab
Histidine metab
Fatty acid metab
Glycolysis/Gluconeogenesis
Inositol phosphate metab
I_0
I_7
Cysteine and methionine metab
Steroid hormone bios
Valine leucine and isoleucine degradation

Fig. 3

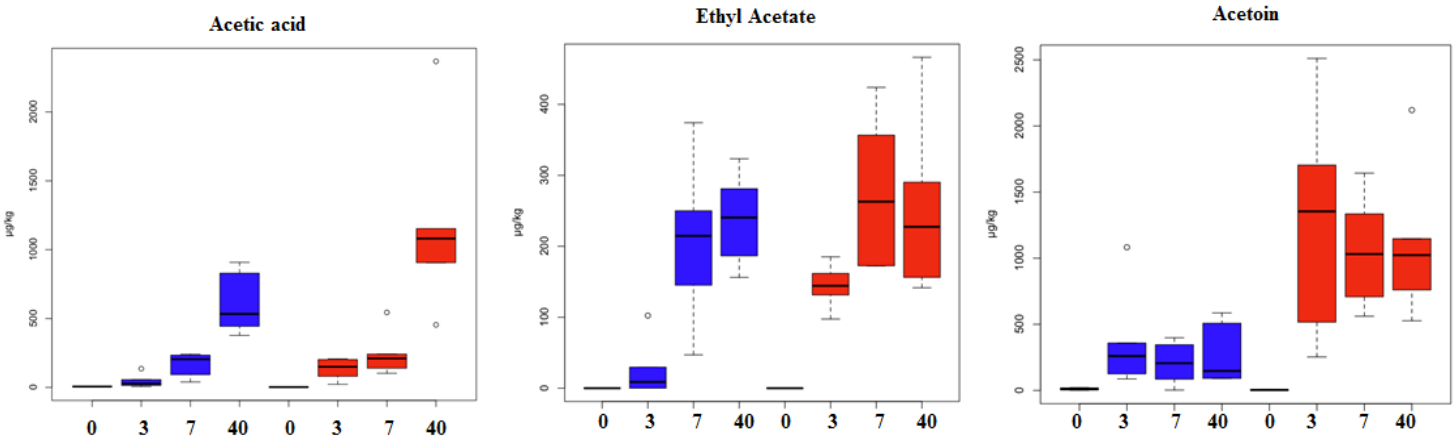


Fig. 4

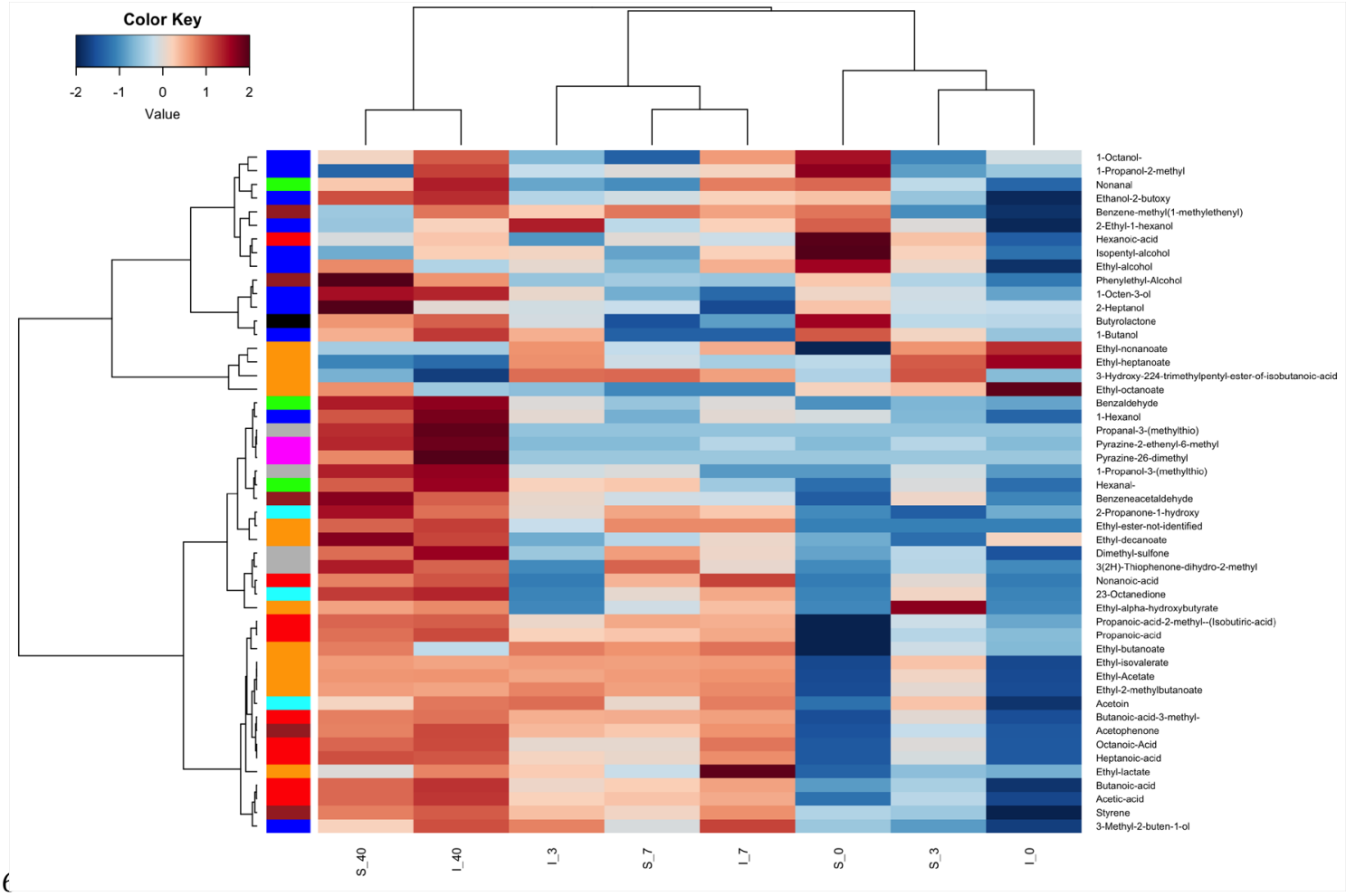

Color Key
-2
-1
0
1
2
Value
1-Octanol-
1-Propanol-2-methyl
Nonanal
Ethanol-2-butoxy
Benzene-methyl(1-methylethenyl)
2-Ethyl-1-hexanol
Hexanoic-acid
Isopentyl-alcohol
Ethyl-alcohol
Phenylethyl-Alcohol
1-Octen-3-ol
2-Heptanol
Butyrolactone
1-Butanol
Ethyl-nonanoate
Ethyl-heptanoate
3-Hydroxy-224-trimethylpentyl-ester-of-isobutanoic-acid
Ethyl-octanoate
Benzaldehyde
1-Hexanol
Propanal-3-(methylthio)
Pyrazine-2-ethenyl-6-methyl
Pyrazine-26-dimethyl
1-Propanol-3-(methylthio)
Hexanal-
Benzeneacetaldehyde
2-Propanone-1-hydroxy
Ethyl-ester-not-identified
Ethyl-decanoate
Dimethyl-sulfone
3(2H)-Thiophenone-dihydro-2-methyl
Nonanoic-acid
23-Octanedione
Ethyl-alpha-hydroxybutyrate
Propanoic-acid-2-methyl--(Isobutiric-acid)
Propanoic-acid
Ethyl-butanoate
Ethyl-isovalerate
Ethyl-Acetate
Ethyl-2-methylbutanoate
Acetoin
Butanoic-acid-3-methyl-
Acetophenone
Octanoic-Acid
Heptanoic-acid
Ethyl-lactate
Butanoic-acid
Acetic-acid
Styrene
3-Methyl-2-buten-1-ol
S_40
I_40
I_3
S_7
I_7
S_0
S_3
I_0

Fig. 5

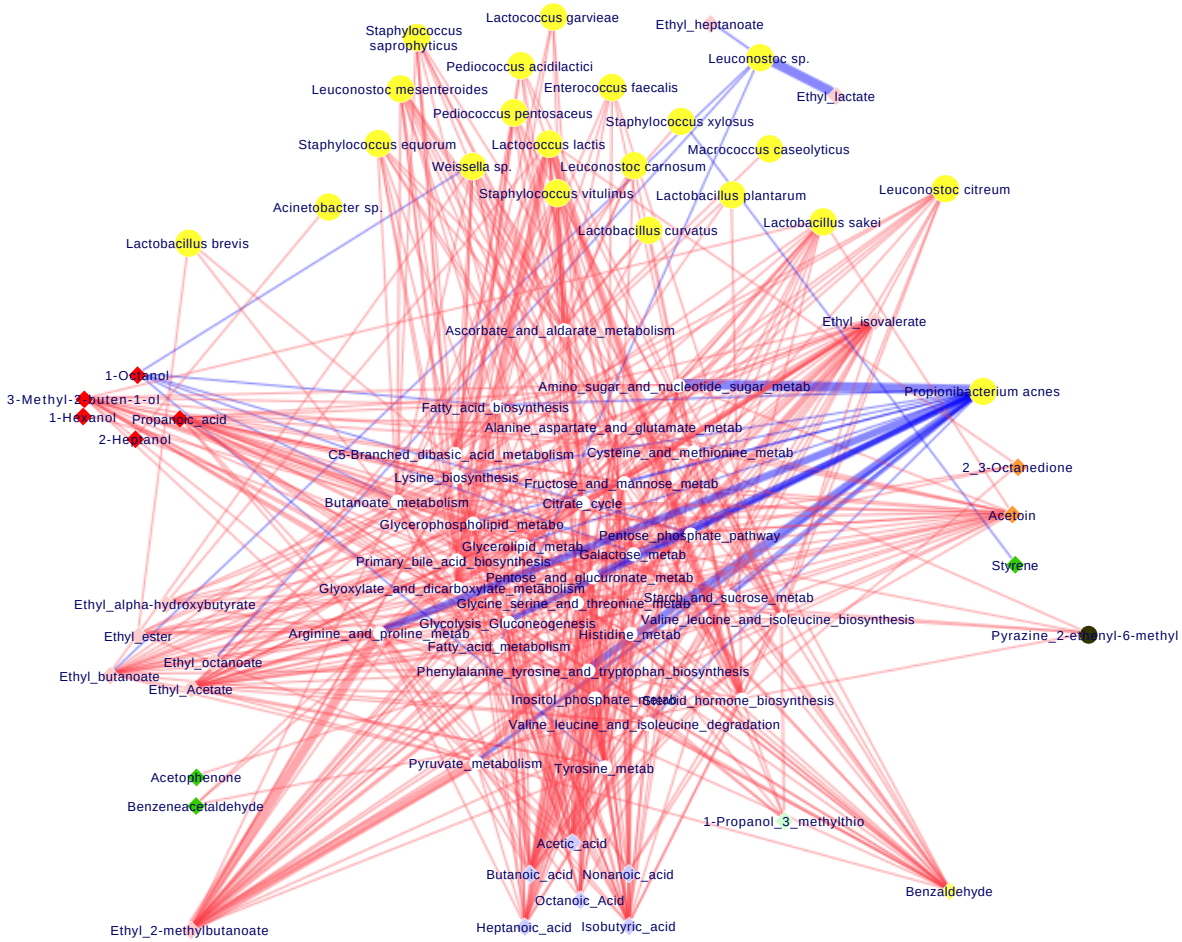
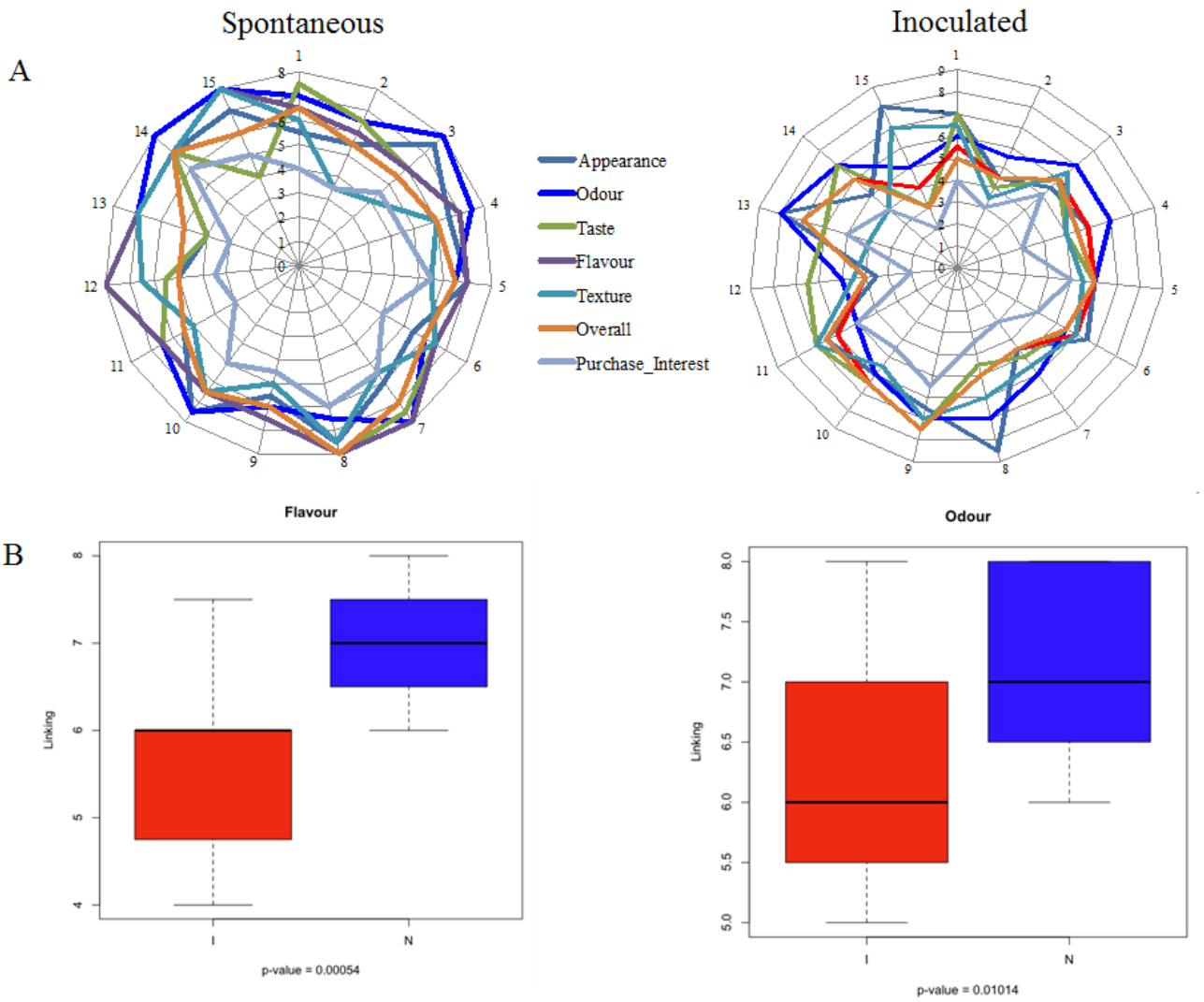

Lactococcus garvieae
Ethyl_heptanoate
Staphylococcus saprophyticus
Leuconostoc sp.
Pediococcus acidilactici
Leuconostoc mesenteroides
Enterococcus faecalis
Ethyl_lactate
Pediococcus pentosaceus
Staphylococcus xylosus
Staphylococcus equorum
Lactococcus lactis
Macrococcus caseolyticus
Weissella sp.
Leuconostoc carnosum
Staphylococcus vitulinus
Lactobacillus plantarum
Leuconostoc citreum
Acinetobacter sp.
Lactobacillus sakei
Lactobacillus curvatus
Lactobacillus brevis
Ethyl_isovalerate
Ascorbate_and_aldarate_metabolism
1-Octanol
Amino_sugar_and_nucleotide_sugar_metab
Propionibacterium acnes
3-Methyl-2-buten-1-ol
Fatty_acid_biosynthesis
1-Heptanol
Propanoic_acid
Alanine_aspartate_and_glutamate_metab
2-Heptanol
C5-Branched_dibasic_acid_metabolism
Cysteine_and_methionine_metab
2_3-Octanedione
Lysine_biosynthesis
Fructose_and_mannose_metab
Butanoate_metabolism
Citrate_cycle
Acetoin
Glycerophospholipid_metabo
Pentose_phosphate_pathway
Glycerolipid_metab
Galactose_metab
Styrene
Primary_bile_acid_biosynthesis
Pentose_and_glucuronate_metab
Glyoxylate_and_dicarboxylate_metabolism
Ethyl_alpha-hydroxybutyrate
Glycine_serine_and_threonine_metab
Starch_and_sucrose_metab
Glycolysis_Gluconeogenesis
Valine_leucine_and_isoleucine_biosynthesis
Ethyl_ester
Arginine_and_proline_metab
Histidine_metab
Pyrazine_2-ethyl-6-methyl
Fatty_acid_metabolism
Ethyl_octanoate
Ethyl_butanoate
Phenylalanine_tyrosine_and_tryptophan_biosynthesis
Ethyl_Acetate
Inositol_phosphate_metab
Steroid_hormone_biosynthesis
Valine_leucine_and_isoleucine_degradation
Pyruvate_metabolism
Tyrosine_metab
Acetophenone
Benzeneacetaldehyde
1-Propanol_3_methylthio
Acetic_acid
Butanoic_acid
Nonanoic_acid
Benzaldehyde
Octanoic_Acid
Ethyl_2-methylbutanoate
Heptanoic_acid
Isobutyric_acid

Fig 6


A
Spontaneous
Inoculated
Appearance
Odour
Taste
Flavour
Texture
Overall
Purchase_Interest
B
Flavour
Linking
p-value = 0.00054
Odour
Linking
p-value = 0.01014

6’
673

Fig 7