UNIVERSITÀ DEGLI STUDI DI MILANO

Department of Food, Environmental and Nutritional Sciences

PhD School in Food Systems – XXXIII Cycle

The role of microorganisms in the effect of food on human health: two case studies

[AGR 16]

ALESSANDRO DALLA VIA
R11869 - R34

Tutor: Prof. Simone Guglielmetti
Ph.D. Dean: Prof. Ella Pagliarini

2019/2020

Index

1. Abstract

This PhD project aims to investigate the interactions that microorganisms may have with some food components. These interactions can be exerted on different levels and can have positive or negative consequences.

In the first study presented we tried to develop a protocol to identify and quantify the presence of enzymatic activity, in the gut, linked to the conversion of choline into trimethylamine (TMA), the precursor of a pro-atherosclerotic agent, trimethylamine oxide (TMAO). First we designed two pairs of primers, cut-Kp and cut-Dd, on the *cutC* gene sequence, identified as the gene capable of encoding for the enzyme responsible for converting choline to TMA. Then we proceeded to select 16 healthy volunteers who provided fecal and urine samples once for three weeks. A qPCR was performed on the fecal samples in order to quantify the *cutC* gene and a 16S rRNA gene profiling. The urine samples were used to quantify TMAO via UPLC-MS. The results obtained from this developed qPCR protocol allowed us to find an association between the levels of TMAO in urine, some taxonomic groups in the gut and a subgroup of the *cutC* gene present in Enterobacteriaceae.

In the second work we used a dairy lactic acid bacteria inoculum for the fermentation of soy drink. Then we characterized and assessed the pro-estrogenic activity of these bacteria. We started inoculated kefir grains into soy drink. We isolated the four best adapted strains from fermented soy: *Lactococcus lactis* K03, *Leuconostoc pseutomesenteroides* K05, *Leuconostoc mesenteroides* K09 and *Lentilactobacillus kefiri* K10. We tested the viability of the four strains at the end of fermentation, their ability to produce a texture similar to a commercial product and their resistance to antibiotics, and we evaluated their ability to use the oligosaccharides naturally present in soybeans. Finally, we were able to evaluate, through

UPLC-MS and the use of a hormonal biosensor in *Saccharomyces cerevisiae*, the ability of the strains to increase the bioavailability of phytoestrogens present in soybeans.

All strains have been shown to ferment sugars and produce a creamy texture. In addition two *Leuconostoc* belonging genus strains, shows the complete fermentation of soy oligosaccharides and the increase of the estrogenic activity of soy drink.

These two works have enabled us to investigate two bacterial enzymatic effects that are very different from each other, which can provide the basis for the development of new practices for human health intervention using bacteria.

2. Introduction

2.1. The human microbiota

An ecosystem is a complex of living organisms interacting with each other and with the environment around them. Our planet has many distinct ecosystems depending on the different biome involved. In each ecosystem, it is possible to find a wide variety of living beings that live in complete connection with the surrounding environment, generating a *do ut des* relationship. This mutual connection between living organisms and the surrounding environment is beneficial for both parties. Organisms receive everything they need to survive, and the environment is modified by their presence. The human body can also be considered as a set of ecosystems. It is, in fact, colonized along its whole surface, during the entire duration of the existence by a wide variety of microorganisms. It has been possible to estimate that the number of bacterial cells in humans is almost equal to the number of cells of which the individual human body is composed (Sender et al., 2016).

The skin (Chen et al., 2018), the oral cavity (Burcham et al., 2020), the vaginal mucosa (Ravel et al., 2011) and the gastro-intestinal tract (Thursby and Juge, 2017) have their own typical microbial populations that are able to interact with the district they live in.

The complex composition of the human microbiota starts to be defined from the first instants after birth. The delivery mode itself, cesarean or natural, has a substantial influence on this assessment. Infants born with natural delivery, in fact, are colonized with microorganisms that can be ascribed to the fecal and vaginal taxa of the mother (Rodríguez et al., 2015), in contrast to infants born by caesarean section. The influence that the delivery has on the composition of the individual microbiota has repercussions throughout the subject's life (Goedert et al., 2014).

The human microbiota is not static and can be influenced by many factors, both intrinsic and extrinsic, which can affect inter and intra-subject variability. Several studies have been performed to this end. A relevant study in this context was carried out by Flores et al. (2014), who reported the variation of the microbiota composition in gut, forehead, palm and tongue of 85 subjects for 3 months. It was observed that the temporal variability of some subjects is considerably higher than others, especially regarding the skin surface, in particular the palm. This study has also shown that the variability of the microbial composition of one surface considered does not predict that of the others.

Among the elements that influence the human microbial ecosystem, external factors such as the experiences of the individual and internal factors such as actions/reactions of the individual, seems to be involved. The study by David et al. (2014) represents concrete example. In this study, the bacterial composition of gut and oral microbiota of two subjects was examined for a whole year. The first subject studied moved from a developed country to a developing country. This radical change in hygienic and dietary habits led to a change in the composition of the predominant bacterial genera, which, however, has been restored to its original status when the subject returned to the original location. The second subject suffered an intestinal infection that led to an irreversible change in the intestinal microbiota; the bacteria lost during the infection have been replaced by other competitors. This study, therefore, supports the hypothesis that microbial communities are generally stable in relation to events that could change their composition; however, some events are irreversible and consequently the microbiota does not return to its original structure.

We will now focus on a specific microbial ecosystem, the intestinal microbiota, to better understand the scientific implications of the two studies presented in this thesis work.

2.2. Effects of intestinal microbiota on humans

Every human ecosystem is characterized by a balance of the interactions between the components of the ecosystem, i.e. the host, the microorganisms and the environment. The interaction between bacteria and human organism can produce effects that impact on the human health both positively and negatively. This is particularly true for the intestinal tract, where the number of microorganisms is the highest and the crosstalk with the host is the most complex and strict.

The wide extension of the intestine, about 300 m^2, includes different sections characterized by different environmental conditions, which result in a high variability of microbial populations specific to each site. This has been confirmed by a study in which almost 300000 16S rRNA gene sequences obtained by pyrosequencing from mouse gut were analyzed (Gu et al., 2013). The result of this study showed the wide biodiversity and the different concentration of bacteria in the small intestine compared to the large. The small intestine presents a high quantity of Lactobacillaceae, whereas anaerobic bacteria such as Bacteroidaceae, Prevotellaceae, Rikenellaceae, Lachnospiraceae and Ruminococcaceae have been found in greater amount in colon and fecal samples.

The abundant bacteria in the gut exert several activities that are relevant for host's health. In fact, the gut microbiota can create a barrier against the wide range of pathogenic bacteria through mechanisms of competitive exclusion on the competition for nutrients and/or adhesion sites (Fatima and Berry, 2017). Another strategy of competition against pathogens involves the production of antimicrobial substances, such as hydrogen peroxide (Hertzberg et al., 2014) or bacteriocins (Zheng et al., 2014).

Since the intestine is main site of contact with the outside world and consequently with external agents, this tract is characterized by a first

defensive barrier, i.e. the lymphocytes located along the epithelium and in organized tissues, which compose the gut-associated lymphoid tissue (GALT). Dendritic cells and macrophages respond actively to the microbiota and its metabolic products, modulating the priming of regulatory T cells (Rescigno, 2014). The intestinal microbiota, in fact, performs an important immunomodulation action through the production of metabolites such as butyrate, which is generated from the fermentation of dietary fiber (Smith et al., 2013) but also through cell-wall components, such as the TgaA protein of *Bifidobacterium bifidum* MIMBb75, able to promote the differentiation of T lymphocytes through the modulation of dendritic cells maturation (Guglielmetti et al., 2014).

2.2.1. The interaction between microbiota and food in the gut

The factor that has the most substantial influence on environmental conditions in the gut is the diet. The diet introduces a very large number of compounds, which differ from each other according to their composition, the time spent in the gastrointestinal tract and the quantity of consumption. The bacteria in the human gut have adapted to live by extracting the source of carbon needed for their growth from the different substrates of food origin, attributing great importance to the variety of food passing through this environment.

One example is dietary fiber, which consists of all edible carbohydrates that are resistant to the action of digestive enzymes and are not absorbed into the small intestine. In addition, bacteria also use all food molecules to perform functions related to secondary metabolism. A study performed by Zheng et al. (2011) estimated the presence of more than 200 metabolites in fecal samples and urine, whose origin was found to be attributable to the intervention of the intestinal microbiota.

These metabolites, which can also only be waste products for the bacteria that produce them, once absorbed by the enterocytes, can lead to effects in the host that can be either beneficial or harmful.

There is therefore a bidirectionality in this link: food choices affect the composition of the microbiota, which in turn affects the health of the host by modifying food molecules.

An example of what has just been discussed is provided by observing the difference between two types of diet: a high-fiber diet, can lead to a prevalence of Bacteroidetes, especially members of the genus *Prevotella*, and an impoverishment of Firmicutes (De Filippo et al, 2010). In contrast, a diet with a high content of animal fats and proteins, typical of Western countries, leads to a prevalence of Firmicutes and Proteobacteria and a concomitant decrease in Bacteroidetes (Hildebrandt et al., 2009). Such diet also leads to less diversity compared to a diet rich in fiber (Sonnenburg et al., 2016).

Similarly, populations from different parts of the planet possess different types of bacteria characterized by enzymatic activities linked to local eating habits, such as the high presence of porphyranase of the intestinal microbiota metagenome of the Japanese population (Hehemann et al., 2010), an enzyme essential for the digestion of nori seaweed, typical of the local diet.

2.2.1.1. Positive health effects

The various beneficial effects of the intestinal microbiota on humans have been studied for decades and there is a wide literature on the subject.

The most observed and evaluated effects are those deriving from the interaction between microbiota and dietary fiber, resulting in the production of short-chain fatty acids (SCFAs) such as acetate, propionate and butyrate (Sittipo et al., 2019).

SCFAs exert many beneficial functions on the host, including the modulation of immune responses and the maintenance of intestinal homeostasis (Smith et al., 2013). Acetate and propionate are mostly produced by Bacteroidetes such as *Bacteroides* spp. and *Prevotella* spp., while butyrate by Firmicutes, such as *Roseburia* spp., *Faecalibacterium* spp. and *Clostridium* spp. (Sittipo et al., 2019).

SCFAs are also very effective in the primary defense line of the intestine by stimulating the production of intestinal mucus (Wrzosek et al., 2013). In addition, the oxidation of butyrate by colonocytes creates an anaerobic layer (Byndloss et al., 2017) that allows the growth of bacteria capable of producing the same butyrate but prevents pathogen proliferation.

The bacteria in the intestinal microbiota are also capable of synthesizing a long series of vitamins, particularly group B vitamins (Magnúsdóttir et al., 2007). The portion of B vitamins produced is added to the portion of vitamins assumed with the diet or is used by other microbial commensals.

Other substrates subject to interaction with the intestinal microbiota are amino acids.

This is the case of tryptophan, whose indolic derivatives produced by bacterial metabolism are able to modulate immune responses and stimulate the secretion of antimicrobial peptides (Gao et al., 2018).

The microbiota can also have effects on micronutrients in diet, such as polyphenols.

As an example, two bacteria isolated in human feces, *Slackia isoflavonconvertens* (Matthies et al., 2009) and *Adlercreutzia equolifaciens* (Maruo et al., 2008) are able to convert daizein and genistein, two phytoestrogens with anticarcinogenic effects present in large quantities in Fabaceae plants, in equol, a compound that exhibits greater activity than its precursors (Kelly et al., 1993).

Another example are berries such as blueberry, raspberry and blackberry, whose polyphenols are not only able to modulate the composition of the

microbiota itself (Guglielmetti et al., 2003) but can also be metabolically converted into urolithin (Selma et al., 2014), a compound with anti-inflammatory and chemopreventive activity (Larrosa et al., 2010).

2.2.1.2. Negative health effects

The intestinal bacterial populations are very heterogeneous and can experience several perturbations, some of which lead to the selection of a bacterial populations with metabolic activities harmful to human health.
Several studies have shown that there is a link between disturbances in the composition of the microbiota and inflammatory bowel diseases (IBD), such as Chron's disease and ulcerative colitis (Albenberg et al., 2012).
These perturbations may derive from external events that lead to the proliferation of non-native species that are competitive with the host organism, but also from our choices, for example, turning towards a diet that presents the typical profile of the Western diet.
Western diet is characterized by a low fiber intake and an increase in complex fatty acids and proteins. The low intake of fiber can lead to the selection of a microbiota that uses intestinal mucus as a source of nutrients (Desai et al., 2016), thus leading to the erosion of the mucus itself, which we remember is one of the first difense barriers of our intestines, thus increasing the concentration of pathogens that promote the incidence of colitis.
In the same way, protein fermentation in the distal part of the colon leads to the production of ammonia, amines and phenols (Windey et al., 2012), implicated in many intestinal diseases such as leaky gut, colorectal cancer and IBD.
High consumption of milk fat leads to the production of taurocholic acid, which causes an increase in the presence of *Bilophila wadsworthia* in the

intestine, a member of the class Deltaproteobacteria associated with the incidence of colitis (Devkota et al., 2012; 2015).

The high consumption of processed meat can lead to high intake of choline and carnitine. These compounds can be converted by the intestinal microbiota into trimethylamine (Craciun et al., 2012) which, in turn, is converted by flavin-monooxygenases in the liver into trimethylamine oxide, a molecule that has been demonstrated to promote atherogenesis (Tang et al. 2013).

Another example is the production of hydrogen sulfite in the intestine starting from several diet components including the amino acid cysteine. The production of this compound can have a positive effect when produced in low quantities or a negative effect when produced in high quantities, as observed in certain intestinal pathological conditions (Blachier et al., 2019).

2.3. Fermented products

Now a question arises, is it possible to modulate the intestinal microbial ecosystem through the ingestion of food-associated bacteria?

This idea has been circulating in the world of microbiology since its beginnings. In the early twentieth century one of the pioneers of this area, Elie Metchnikoff (1910), hypothesized that yoghurt could have a beneficial effect on human health, thanks to the bacteria it contains.

Fermented foods in general are thought to contain microorganisms potentially beneficial for human health. Fermentation is one of the oldest systems adopted for the preservation of food. Interestingly, the many variables in the production of fermented foods, the environmental conditions, the type of food matrix and the type of bacteria used, result in a very heterogeneous amount and quality of fermented products.

However, the progress of food sciences and the industrialization of production processes led to a drastic decline in the consumption of a wide range of traditional fermented foods. Nonetheless, nowadays, the increased attention on the health-promoting properties of food is promoting the rediscovery and diffusion of traditional fermented foods such as kefir, kumis, tempeh, and miso.

2.3.1. Effects of fermentation on human health

Microbial fermentation often affects the potential effects of food on the human health. In this context, the simplest example is represented by fermented milk products such as traditional yoghurt and kefir, which may be well tolerated by lactose intolerant individuals due to the elimination of lactose during fermentation deriving from the microbial β-galactosidase activity (Savaiano, 2014; Hertzler et al., 2003).

Fermented soy foods, such as miso and natto, contain a high amount of isoflavones compared to their unfermented counterparts, soy milk and tofu (Fukutake et al., 1996). The high consumption of these foods in Asian populations has been associated to the low incidence of colon and prostate cancer due to the effect of these phytoestrogens (Adlercreutz et al., 1995).

Bacteria can also produce compounds that can directly affect human health. Lactic acid, for example, can reduce the secretion of pro-inflammatory cytokines (Iraporda et al., 2015); in addition, B vitamins (Russo et al., 2014) and exopolysaccharides, which some bacteria can produce in certain food matrices, have an antioxidant effect and may act as prebiotics (Li et al., 2014; Salazar et al., 2015). Another example is represented by nattokinase, a subtilisin produced by *Bacillus subtilis* var. natto, the main strain used in the preparation of Natto, an Asian food made from cooked yellow soybean. This peptide has been shown to have

antimicrobial activity against pathogens such as *Streptococcus pneuomiae* (Kitagawa et al., 2017) as well as antithrombotic and antihypertensive activity, as demonstrated in a clinical study by Jensen et al. (2016).

Fermented foods can also influence the composition of the human microbiota (Jeong et al., 2017; Yilmaz et al., 2019). For instance, tempeh, an Asian food consisting of boiled soya fermented from lactic acid bacteria and fungi, was shown to modulate the intestinal microbiota, leading to an increase in the intestinal abundance of *Lactobacillus*, *Bifidobacterium* and *Enterococcus* (Kuligowski et al., 2013).

2.4. References

Adlercreutz, C. H., Goldin, B. R., Gorbach, S. L., Höckerstedt, K. A., Watanabe, S., Hämäläinen, E. K., Markkanen, M. H., Mäkelä, T. H., Wähälä, K. T., & Adlercreutz, T. (1995). Soybean phytoestrogen intake and cancer risk. *The Journal of nutrition*, *125*(3 Suppl), 757S–770S.

Albenberg, L. G., Lewis, J. D., & Wu, G. D. (2012). Food and the gut microbiota in inflammatory bowel diseases: a critical connection. *Current opinion in gastroenterology*, *28*(4), 314–320.

Allegretti, J. R., Mullish, B. H., Kelly, C., & Fischer, M. (2019). The evolution of the use of faecal microbiota transplantation and emerging therapeutic indications. *Lancet (London, England)*, *394*(10196), 420–431.

Bauer, M. A., Kainz, K., Carmona-Gutierrez, D., & Madeo, F. (2018). Microbial wars: Competition in ecological niches and within the microbiome. *Microbial cell (Graz, Austria)*, *5*(5), 215–219.

Blachier, F., Beaumont, M., & Kim, E. (2019). Cysteine-derived hydrogen sulfide and gut health: a matter of endogenous or bacterial origin. *Current opinion in clinical nutrition and metabolic care*, *22*(1), 68–75.

Byndloss, M. X., Olsan, E. E., Rivera-Chávez, F., Tiffany, C. R., Cevallos, S. A., Lokken, K. L., Torres, T. P., Byndloss, A. J., Faber, F., Gao, Y., Litvak, Y., Lopez, C. A., Xu, G., Napoli, E., Giulivi, C., Tsolis, R. M., Revzin, A., Lebrilla, C. B., & Bäumler, A. J. (2017). Microbiota-activated PPAR-γ signaling inhibits dysbiotic Enterobacteriaceae expansion. *Science (New York, N.Y.)*, *357*(6351), 570–575.

Craciun, S., & Balskus, E. P. (2012). Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), 21307–21312.

David, L. A., Materna, A. C., Friedman, J., Campos-Baptista, M. I., Blackburn, M. C., Perrotta, A., Erdman, S. E., & Alm, E. J. (2014). Host lifestyle affects human microbiota on daily timescales. *Genome biology*, *15*(7), R89.

De Filippo, C., Cavalieri, D., Di Paola, M., Ramazzotti, M., Poullet, J. B., Massart, S., Collini, S., Pieraccini, G., & Lionetti, P. (2010). Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(33), 14691–14696.

Desai, M. S., Seekatz, A. M., Koropatkin, N. M., Kamada, N., Hickey, C. A., Wolter, M., Pudlo, N. A., Kitamoto, S., Terrapon, N., Muller, A., Young, V. B., Henrissat, B., Wilmes, P., Stappenbeck, T. S., Núñez, G., & Martens,

E. C. (2016). A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility. *Cell*, *167*(5), 1339–1353.e21.

Devkota, S., Wang, Y., Musch, M. W., Leone, V., Fehlner-Peach, H., Nadimpalli, A., Antonopoulos, D. A., Jabri, B., & Chang, E. B. (2012). Dietary-fat-induced taurocholic acid promotes pathobiont expansion and colitis in Il10-/- mice. *Nature*, *487*(7405), 104–108.

Devkota, S., & Chang, E. B. (2015). Interactions between Diet, Bile Acid Metabolism, Gut Microbiota, and Inflammatory Bowel Diseases. *Digestive diseases (Basel, Switzerland)*, *33*(3), 351–356.

Dimidi, E., Christodoulides, S., Scott, S. M., & Whelan, K. (2017). Mechanisms of Action of Probiotics and the Gastrointestinal Microbiota on Gut Motility and Constipation. *Advances in nutrition (Bethesda, Md.)*, *8*(3), 484–494.

Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW, Vázquez-Baeza Y, Gonzalez A, Knight R, Dunn RR, Fierer N. Temporal variability is a personalized feature of the human microbiome. Genome Biol. 2014 Dec 3;15(12):531.

Fukutake, M., Takahashi, M., Ishida, K., Kawamura, H., Sugimura, T., & Wakabayashi, K. (1996). Quantification of genistein and genistin in soybeans and soybean products. *Food and chemical toxicology: an international journal published for the British Industrial Biological Research Association*, *34*(5), 457–461.

Gao, J., Xu, K., Liu, H., Liu, G., Bai, M., Peng, C., Li, T., & Yin, Y. (2018). Impact of the Gut Microbiota on Intestinal Immunity Mediated by Tryptophan Metabolism. *Frontiers in cellular and infection microbiology*, *8*, 13.

Goedert, J. J., Hua, X., Yu, G., & Shi, J. (2014). Diversity and composition of the adult fecal microbiome associated with history of cesarean birth or appendectomy: Analysis of the American Gut Project. *EBioMedicine*, *1*(2-3), 167–172.

Gu, S., Chen, D., Zhang, J. N., Lv, X., Wang, K., Duan, L. P., Nie, Y., & Wu, X. L. (2013). Bacterial community mapping of the mouse gastrointestinal tract. *PloS one*, *8*(10), e74957.

Guglielmetti, S., Fracassetti, D., Taverniti, V., Del Bo', C., Vendrame, S., Klimis-Zacas, D., Arioli, S., Riso, P., & Porrini, M. (2013). Differential modulation of human intestinal bifidobacterium populations after consumption of a wild blueberry (Vaccinium angustifolium) drink. *Journal of agricultural and food chemistry*, *61*(34), 8134–8140.

Guglielmetti, S., Zanoni, I., Balzaretti, S., Miriani, M., Taverniti, V., De Noni, I., Presti, I., Stuknyte, M., Scarafoni, A., Arioli, S., Iametti, S., Bonomi, F., Mora, D., Karp, M., & Granucci, F. (2014). Murein lytic enzyme TgaA of Bifidobacterium bifidum MIMBb75 modulates dendritic cell maturation through its cysteine- and histidine-dependent amidohydrolase/peptidase (CHAP) amidase domain. *Applied and environmental microbiology*, *80*(17), 5170–5177.

Hehemann, J. H., Correc, G., Barbeyron, T., Helbert, W., Czjzek, M., & Michel, G. (2010). Transfer of carbohydrate-active enzymes from marine bacteria to Japanese gut microbiota. *Nature*, *464*(7290), 908–912.

Hertzberger, R., Arents, J., Dekker, H. L., Pridmore, R. D., Gysler, C., Kleerebezem, M., & de Mattos, M. J. (2014). H(2)O(2) production in species of the Lactobacillus acidophilus group: a central role for a novel NADH-dependent flavin reductase. *Applied and environmental microbiology*, *80*(7), 2229–2239.

Hertzler, S. R., & Clancy, S. M. (2003). Kefir improves lactose digestion and tolerance in adults with lactose maldigestion. *Journal of the American Dietetic Association*, *103*(5), 582–587.

Hildebrandt, M. A., Hoffmann, C., Sherrill-Mix, S. A., Keilbaugh, S. A., Hamady, M., Chen, Y. Y., Knight, R., Ahima, R. S., Bushman, F., & Wu, G. D. (2009). High-fat diet determines the composition of the murine gut microbiome independently of obesity. *Gastroenterology*, *137*(5), 1716–24.e242.

Iraporda, C., Errea, A., Romanin, D. E., Cayet, D., Pereyra, E., Pignataro, O., Sirard, J. C., Garrote, G. L., Abraham, A. G., & Rumbo, M. (2015). Lactate and short chain fatty acids produced by microbial fermentation downregulate proinflammatory responses in intestinal epithelial cells and myeloid cells. *Immunobiology*, *220*(10), 1161–1169.

Jensen, G. S., Lenninger, M., Ero, M. P., & Benson, K. F. (2016). Consumption of nattokinase is associated with reduced blood pressure and von Willebrand factor, a cardiovascular risk marker: results from a

randomized, double-blind, placebo-controlled, multicenter North American clinical trial. *Integrated blood pressure control*, *9*, 95–104.

Jeong, D., Kim, D. H., Kang, I. B., Kim, H., Song, K. Y., Kim, H. S., & Seo, K. H. (2017). Modulation of gut microbiota and increase in fecal water content in mice induced by administration of Lactobacillus kefiranofaciens DN1. *Food & function*, *8*(2), 680–686.

Kelly, G. E., Nelson, C., Waring, M. A., Joannou, G. E., & Reeder, A. Y. (1993). Metabolites of dietary (soya) isoflavones in human urine. *Clinica chimica acta; international journal of clinical chemistry*, *223*(1-2), 9–22

Kitagawa, M., Shiraishi, T., Yamamoto, S., Kutomi, R., Ohkoshi, Y., Sato, T., Wakui, H., Itoh, H., Miyamoto, A., & Yokota, S. I. (2017). Novel antimicrobial activities of a peptide derived from a Japanese soybean fermented food, Natto, against *Streptococcus pneumoniae* and *Bacillus subtili*s group strains. *AMB Express*, *7*(1), 127.

Kuligowski, M., Jasińska-Kuligowska, I., & Nowak, J. (2013). Evaluation of bean and soy tempeh influence on intestinal bacteria and estimation of antibacterial properties of bean tempeh. *Polish journal of microbiology*, *62*(2), 189–194.

Larrosa, M., González-Sarrías, A., Yáñez-Gascón, M. J., Selma, M. V., Azorín-Ortuño, M., Toti, S., Tomás-Barberán, F., Dolara, P., & Espín, J. C. (2010). Anti-inflammatory properties of a pomegranate extract and its metabolite urolithin-A in a colitis rat model and the effect of colon inflammation on phenolic metabolism. *The Journal of nutritional biochemistry*, *21*(8), 717–725.

Li, W., Ji, J., Chen, X., Jiang, M., Rui, X., & Dong, M. (2014). Structural elucidation and antioxidant activities of exopolysaccharides from Lactobacillus helveticus MB2-1. *Carbohydrate polymers*, *102*, 351–359.

Magnúsdóttir, S., Ravcheev, D., de Crécy-Lagard, V., & Thiele, I. (2015). Systematic genome assessment of B-vitamin biosynthesis suggests co-operation among gut microbes. *Frontiers in genetics*, *6*, 148.

Maruo, T., Sakamoto, M., Ito, C., Toda, T., & Benno, Y. (2008). Adlercreutzia equolifaciens gen. nov., sp. nov., an equol-producing bacterium isolated from human faeces, and emended description of the genus Eggerthella. *International journal of systematic and evolutionary microbiology*, *58*(Pt 5), 1221–1227.

Matthies, A., Blaut, M., & Braune, A. (2009). Isolation of a human intestinal bacterium capable of daidzein and genistein conversion. *Applied and environmental microbiology*, *75*(6), 1740–1744.

Metchnikoff, E. (1910) The Prolongation of Life. Optimistic Studies. *New York: G P Putnam's Sons*. 96 p.

Osawa, N., & Mitsuhashi, S. (1964). Infection of germfree mice with *Shingella flexneri* 3A. *The Japanese journal of experimental medicine*, *34*, 77–80.

Pereira, F. C., & Berry, D. (2017). Microbial nutrient niches in the gut. *Environmental microbiology*, *19*(4), 1366–1378.

Ravel, J., Gajer, P., Abdo, Z., Schneider, G. M., Koenig, S. S., McCulle, S. L., Karlebach, S., Gorle, R., Russell, J., Tacket, C. O., Brotman, R. M.,

Davis, C. C., Ault, K., Peralta, L., & Forney, L. J. (2011). Vaginal microbiome of reproductive-age women. *Proceedings of the National Academy of Sciences of the United States of America*, *108 Suppl 1*(Suppl 1), 4680–4687.

Rescigno M. (2014). Intestinal microbiota and its effects on the immune system. *Cellular microbiology*, *16*(7), 1004–1013.

Rodríguez, J. M., Murphy, K., Stanton, C., Ross, R. P., Kober, O. I., Juge, N., Avershina, E., Rudi, K., Narbad, A., Jenmalm, M. C., Marchesi, J. R., & Collado, M. C. (2015). The composition of the gut microbiota throughout life, with an emphasis on early life. *Microbial ecology in health and disease*, *26*, 26050.

Russo, P., Capozzi, V., Arena, M. P., Spadaccino, G., Dueñas, M. T., López, P., Fiocco, D., & Spano, G. (2014). Riboflavin-overproducing strains of Lactobacillus fermentum for riboflavin-enriched bread. *Applied microbiology and biotechnology*, *98*(8), 3691–3700.

Salazar, N., Gueimonde, M., de Los Reyes-Gavilán, C. G., & Ruas-Madiedo, P. (2016). Exopolysaccharides Produced by Lactic Acid Bacteria and Bifidobacteria as Fermentable Substrates by the Intestinal Microbiota. *Critical reviews in food science and nutrition*, *56*(9), 1440–1453.

Santos, A., San Mauro, M., Sanchez, A., Torres, J. M., & Marquina, D. (2003). The antimicrobial properties of different strains of Lactobacillus spp. isolated from kefir. *Systematic and applied microbiology*, *26*(3), 434–437.

Savaiano D. A. (2014). Lactose digestion from yogurt: mechanism and relevance. *The American journal of clinical nutrition*, *99*(5 Suppl), 1251S–5S.

Selma, M. V., Beltrán, D., García-Villalba, R., Espín, J. C., & Tomás-Barberán, F. A. (2014). Description of urolithin production capacity from ellagic acid of two human intestinal Gordonibacter species. *Food & function*, *5*(8), 1779–1784.

Sender, R., Fuchs, S., & Milo, R. (2016). Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS biology*, *14*(8), e1002533.

Sittipo, P., Shim, J. W., & Lee, Y. K. (2019). Microbial Metabolites Determine Host Health and the Status of Some Diseases. *International journal of molecular sciences*, *20*(21), 5296.

Smith, P. M., Howitt, M. R., Panikov, N., Michaud, M., Gallini, C. A., Bohlooly-Y, M., Glickman, J. N., & Garrett, W. S. (2013). The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis. *Science (New York, N.Y.)*, *341*(6145), 569–573.

Sommer, F., Anderson, J. M., Bharti, R., Raes, J., & Rosenstiel, P. (2017). The resilience of the intestinal microbiota influences health and disease. *Nature reviews. Microbiology*, *15*(10), 630–638.

Sonnenburg, E. D., Smits, S. A., Tikhonov, M., Higginbottom, S. K., Wingreen, N. S., & Sonnenburg, J. L. (2016). Diet-induced extinctions in the gut microbiota compound over generations. *Nature*, *529*(7585), 212–215.

Strachan D. P. (1989). Hay fever, hygiene, and household size. *BMJ (Clinical research ed.)*, *299*(6710), 1259–1260.

Tang, W. H., Wang, Z., Levison, B. S., Koeth, R. A., Britt, E. B., Fu, X., Wu, Y., & Hazen, S. L. (2013). Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *The New England journal of medicine*, *368*(17), 1575–1584.

Windey, K., De Preter, V., & Verbeke, K. (2012). Relevance of protein fermentation to gut health. *Molecular nutrition & food research*, *56*(1), 184–196.

Wrzosek, L., Miquel, S., Noordine, M. L., Bouet, S., Joncquel Chevalier-Curt, M., Robert, V., Philippe, C., Bridonneau, C., Cherbuy, C., Robbe-Masselot, C., Langella, P., & Thomas, M. (2013). Bacteroides thetaiotaomicron and Faecalibacterium prausnitzii influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. *BMC biology*, *11*, 61.

Yılmaz, I., Dolar, M. E., & Özpınar, H. (2019). Effect of administering kefir on the changes in fecal microbiota and symptoms of inflammatory bowel disease: A randomized controlled trial. *The Turkish journal of gastroenterology: the official journal of Turkish Society of Gastroenterology*, *30*(3), 242–253.

Zheng, J., Gänzle, M. G., Lin, X. B., Ruan, L., & Sun, M. (2015). Diversity and dynamics of bacteriocins from human microbiome. *Environmental microbiology*, *17*(6), 2133–2143.

Zheng, X., Xie, G., Zhao, A., Zhao, L., Yao, C., Chiu, N. H., Zhou, Z., Bao, Y., Jia, W., Nicholson, J. K., & Jia, W. (2011). The footprints of gut microbial-mammalian co-metabolism. *Journal of proteome research*, *10*(12), 5512–5522.

3. Aim of the PhD Work

The aim of this thesis project was to investigate the effects that bacteria can have on food components. These effects, as mentioned in the introduction, can be multiple and can have numerous implications for human health.

In this work, we analyzed two types of effects of microorganisms on food molecules: a negative effect performed by bacteria in the intestinal microbiota and a positive effect exerted by microorganisms directly on a food matrix.

In detail, in this project we investigated:

- The role of intestinal bacteria in the levels of trimethylamine N-oxide (TMAO, a pro-atherosclerotic agent) as a consequence of the conversion of choline into trimethylamine (TMA). In specific, the aim of this study was to verify whether urinary levels of TMAO may be associated with the fecal relative abundance of specific bacterial taxa and the bacterial choline trimethylamine-lyase gene *cutC*.
- The ability of selected food-associated lactic acid bacteria to modulate the phytoestrogenic activity of soy isoflavones. In specific, we studied lactic acid bacteria isolated from a milk-based fermented product, kefir, through the progressive adaptation of its microbiota to a commercial soy drink.

4. Results

4.1. Urinary TMAO Levels Are Associated with the Taxonomic Composition of the Gut Microbiota and with the Choline TMA-Lyase Gene (*cutC*) Harbored by Enterobacteriaceae

4.1.1 Abstract

Gut microbiota metabolization of dietary choline may promote atherosclerosis through trimethylamine (TMA), which is rapidly absorbed and converted in the liver to proatherogenic trimethylamine-N-oxide (TMAO). The aim of this study was to verify whether TMAO urinary levels may be associated with the fecal relative abundance of specific bacterial taxa and the bacterial choline TMA-lyase gene *cutC*. The analysis of sequences available in GenBank grouped the *cutC* gene into two main clusters, cut-Dd and cut-Kp. A quantitative real-time polymerase chain reaction (qPCR) protocol was developed to quantify *cutC* and was used with DNA isolated from three fecal samples collected weekly over the course of three consecutive weeks from 16 healthy adults. The same DNA was used for 16S rRNA gene profiling. Concomitantly, urine was used to quantify TMAO by ultra-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS). All samples were positive for *cutC* and TMAO. Correlation analysis showed that the cut-Kp gene cluster was significantly associated with *Enterobacteriaceae*. Linear mixed models revealed that urinary TMAO levels may be predicted by fecal cut-Kp and by 23 operational taxonomic units (OTUs). Most of the OTUs significantly associated with TMAO were also significantly associated with cut-Kp, confirming the possible relationship between these two factors. In conclusion, this preliminary method-development study suggests the existence of a relationship between TMAO excreted in urine, specific fecal

bacterial OTUs, and a *cutC* subgroup ascribable to the choline-TMA conversion enzymes of *Enterobacteriaceae*.

4.1.2. Introduction

From infancy, the microorganisms colonizing the human gastrointestinal tract (GIT), collectively known as GIT microbiota, act as a “hidden” metabolic organ that exerts indispensable functions for the development and physiology of the human organism, such as the production of vitamins, modulation of the immune system, competitive exclusion toward exogenous pathogenic bacteria, xenobiotic detoxification, and production of short-chain fatty acids (Mohajeri et al., 2018). Nonetheless, detrimental activities have also been associated with gut commensal microorganisms, such as the production of carcinogens by the bacterial nitroreductases and azoreductases (Claus et al., 2016), or the conversion of primary bile acids to toxic compounds by the microbiota-associated enzyme cholesterol dehydrogenase and 7-dehydroxylase (Ridlon et al., 2016). In addition, it was proposed that the intestinal bacterial enzymatic activities that produce trimethylamine (TMA) may promote atherosclerosis. TMA, in fact, is readily absorbed from the intestinal tract and, once in the liver, is converted into trimethylamine-N-oxide (TMAO) (Tang et al., 2013), whose plasma level has been identified as a metabolite strongly associated with atherosclerosis in a large case-control cohort for cardiovascular disease (Wang et al., 2011). In particular, TMAO was proposed to promote atherogenesis by increasing cholesterol in macrophages and enhancing the accumulation of foam cells in artery walls (Tang et al., 2013; Wang et al., 2011). Nonetheless, the literature has contradicted the role of TMAO, and recent studies have questioned its deleterious role in the cardiovascular system (Nowinski et al., 2018), suggesting, on the contrary,

that TMAO could have protective functions (Huc et al., 2018; Collins et al., 2016).

Reportedly, a dominant contribution to the production of TMA in the gut comes from the microbial metabolism of diet-derived substrates such as carnitine- and choline-containing molecules (Tang et al., 2013; Wang et al., 2011; Koeth et al., 2013). Choline is an essential nutrient that is used by cells to synthesize membrane phospholipids. Furthermore, choline is the precursor of the neurotransmitter acetylcholine and a major source for methyl groups via its metabolite, trimethylglycine (betaine) (Zeisel et al., 1994). The main dietary sources of the choline moiety, which is mostly present in food as lecithin (i.e., phosphatidylcholine), were reported to be eggs, liver, soybeans, and pork (Zeisel et al., 2003). Although they are also present in numerous other foods (Wiedeman et al., 2018), recent surveys in the USA indicated that choline may be underconsumed in specific populations (e.g., pregnant women and vegans) (Wallace et al., 2018). Based on the average observed choline intake in healthy European populations, a panel of the European Food Safety Authority set the adequate intake of choline at 400 mg/day (EFSA, 2016).

Recent literature has suggested that the enhanced abundance of choline utilization genes in the intestinal microbiome is associated with increased TMA levels in the gut and, subsequently, with a higher hepatic production of TMAO. Proof of the importance of choline-derived TMA in the context of TMAO toxicity was recently provided by the study of Craciun and Balskus, in which the specific inhibition in mouse intestine of the microbial choline TMA-lyase (the primary enzymatic activity involved in the production of TMA from choline) (Craciun and Balskus, 2012) resulted in a significant reduction in plasma TMAO levels and recovery from dietary-induced platelet aggregation and thrombus formation (Roberts et al., 2018). Choline TMA-lyase is discontinuously distributed in bacterial taxa. Consequently, it was speculated that the phylogenetic composition of the

microbiota is plausibly a poor predictor of the intestinal potential to convert choline into TMA (Craciun and Balskus, 2012; Rath et al., 2017; Martinez-del Campo et al., 2015). However, in another study, the taxonomic structure of the gut microbiota was used to predict genes involved in choline metabolism (Xu et al., 2017) by means of PICRUSt, a bioinformatic tool used to infer the functional profiles of the microbial communities from 16S rRNA gene profiling data (Langille et al., 2013). Although the toxicity of TMAO has been extensively investigated in the last 10 years, the association potentially existing among host TMAO levels, gut microbiota composition, and the intestinal microbial metabolization of choline has been only marginally considered. In this context, we developed a molecular protocol for the targeted quantification in the fecal microbiome of the bacterial gene *cutC* coding for the glycyl radical enzyme homolog choline TMA-lyase (Craciun and Balskus, 2012; Craciun et al., 2014). This protocol was applied to quantify the *cutC* gene abundance in the fecal samples collected at different time points from a group of healthy adults. Then, the obtained results were analyzed in comparison with the bacterial taxonomic composition and the urinary levels of TMAO concomitantly determined in the same population to deduce the potential association of excreted TMAO with gut microbial taxa and/or specific choline TMA-lyase enzymes.

4.1.3. Materials and methods

4.1.3.1. Design and use of primers targeting the cutC *gene*

The primers used in polymerase chain reaction (PCR) for the amplification of the *cutC* gene were designed as follows. The GenBank database and Conserved Domain Database (CDD) at the National Center for Biotechnology Information (NCBI) were queried to select 52 nonredundant representative bacterial proteins of the choline trimethylamine-lyase

protein family TIGR04394 (choline_CutC; EC Number 4.3.99.4), including the *cutC* enzymes of *Desulfovibrio desulfuricans* (Tang et al., 2013), and *Klebsiella pneumoniae* (Kalnins et al., 2015). Then, the corresponding CDS nucleotide sequences of selected proteins were used to build a UPMGA tree upon ClustalW multiple alignments. According to the obtained dendrogram, sequences were clustered in two groups: One including the *cutC* sequence of *K. pneumoniae*, named cut-Kp, and one including the *cutC* sequence of *D. desulfuricans*, named cut-Dd (Supplementary **Figure S1**). Finally, a pair of primers was designed in the most conserved regions of each group of sequences: cut-Dd-F, 5'-CGTGTTGACCAGTACATGTA-3' and cut-Dd-R 5'-GCTGGTAACCTGCGAAGAA-3' (expected amplicon of 185 bp); cut-Kp-F, 5'-GATCTGACCTATCTGATTATGG-3', and cut-Kp-R, 5'-TTGTGGAGCATCATCTTGAT-3' (expected amplicon of 190 bp).

4.1.3.2. PCR detection of cutC *gene in single strains*

The two primer pairs designed as described above were used in endpoint PCR with the genomic DNA extracted from 64 bacterial strains (**Table S1**). Reaction mix was prepared in 25 Mm μl, including 0.5 units of DreamTaq Polymerase (ThermoFisher, Fermentas, Waltham, MA, USA), 1× concentration of DreamTaq Polymerase Buffer (ThermoFisher, Fermentas,), 0.25 μM of each primer, 200 μM of deoxyribonucleotide triphosphate (dNTPs), and 0.5 mM of MgCl2. The PCR cycle program used was the following: Initial denaturation at 95 °C for 2 min, followed by 35 cycles of denaturation at 94 °C for 45 s, annealing at 58 °C for 45 s for the cut-Dd couple and 56 °C for 45 s for the cut-Kp couple, and extension at 72 °C for 20 s. A final extension of 7 min at 72 °C was then applied.

4.1.3.3. Detection of choline-utilization activity in single strains

Bacterial strains were grown in the respective culture medium (reported in **Table S1**) for 48 h. Afterward, the biomasses were collected by centrifugation at 9500 g for 10 min. The cell pellets were then washed with sterile PBS and resuspended in fresh medium with the addition of 0.2% filter-sterilized choline. Bacteria were incubated at 37 °C for 48 h in glass tubes with screw cap. Afterward, supernatants were collected and used for mass spectrometry (MS) and nuclear magnetic resonance (NMR) analyses. The MS analyzes were performed by directly injecting 5 μl of diluted broth cultures after the removal of the bacterial cells by centrifugation and subsequent filtration with a 0.45-μm syringe filter. In detail, the broth cultures were analyzed in full scan in the range from 50 u to 400 u on an HR-MS Orbitrap model Exactive with a HESI-II probe for electrospray ionization (Thermo Scientific, San Jose, CA, USA). The resolution, gain control, mass tolerance, and maximum ion injection time was set to 50 K, 1E6, 2 ppm, and 100 ms, respectively. The MS data were processed using Xcalibur software (Thermo Scientific). Choline and TMA were used as reference standard. Choline and TMA were also directly detected in broth cultures by 1H-NMR with a 60 MHz benchtop NMR spectrometer Spinsolve 60 Carbon Ultra, Magritek GmbH (Aachen, Germany).

4.1.3.4. Study population

Study participants were recruited within the University of Milan campus. In total, four females and 12 males aged 21–45 (mean: 29.8 years) were enrolled (**Table S2**). The inclusion criteria were as follows: Healthy adult volunteers of both sexes who provided signed informed consent of their participation in the study. The exclusion criteria were as follows: Antibiotic consumption in the month preceding the start of the study, consumption of

antacids or prokinetic gastrointestinal drugs, episodes of viral or bacterial enteritis in the two months prior to the study, episodes of gastric or duodenal ulcers in the previous five years, pregnancy or breastfeeding, recent history of alcohol abuse or suspected drug use, and any severe disease that may interfere with treatment. Ethical permission was granted by the University of Milan Ethics Committee (ref: opinion no. 37/16, 15 December 2016).

4.1.3.5. Collection of fecal and urine samples

Three fecal sample were collected weekly over the course of three consecutive weeks from each volunteer. All the participants were asked to follow their regular diet during the three weeks. Concomitantly to the fecal sample, the volunteers provided 24-h urine collection.

Urine samples were collected over 24 h in sterile tanks and on the same days that fecal samples were been collected. The volume of collected urine was recorded in order to calculate the daily excretion of trimethylamine oxide (TMAO). Immediately after delivery, part of the urine samples was transferred in 10-ml sterile tubes and stored at -80 °C until analysis.

4.1.3.6. Analysis of cutC *gene by quantitative real-time PCR*

The *cutC* gene was quantified in fecal DNA with quantitative real-time PCR (qPCR) with both primer pairs, cut-Dd and cut-Kp. To this aim, DNA was extracted from feces using the kit PowerLyzer® PowerFecal® DNA Isolation Kit (MO BIO Laboratories, Inc.), starting from 0.25 ± 0.02 mg of sample according to the manufacturer's instructions. Primer pairs were tested with a gradient qPCR in a range of eight temperatures in order to find the most efficient annealing temperature using DNA of *Streptococcus*

dysgalactiae 485 and *Klebsiella* sp. A1.2 as reference DNA. In addition, the amplification efficiency of the two pairs of primers was tested in qPCR experiments with six serial 1:3 dilutions of genomic DNA isolated from *Streptococcus dysgalactiae* 485, *Klebsiella* sp. A1.2, and human fecal metagenomic DNA. All DNA (bacterial and metagenomic) serial dilutions were tested with primer concentrations of 0.5 μM, 0.4 μM, and 0.3 μM. Efficiency curves were obtained with Bio-Rad software by setting samples as "standard" and obtaining a curve with efficiency (E) parameter and R2 value. Based on the results of these setup experiments, primers were then used at a final concentration of 0.5 μM, as with this concentration, we obtained an R2 value of 0.98. In addition, two randomly selected fecal DNA samples were tested at the different concentration by adding 70 ng, 50 ng, 25 ng, and 10 ng in qPCR reactions. Based on Ct value comparison between the different DNA concentrations, the *cutC* gene quantification was subsequently performed using 50 ng of total DNA. The reaction mix contained the SsoFast TM Eva-SuperGreen Supermix 2× (Bio-Rad Laboratories), deionized Milli-Q water (Millipore), and primers. All DNA samples (5 μL in each well) were tested in technical duplicate. The qPCR cycles employed were the following: Initial denaturation at 95 °C for 3 min, followed by 44 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C (for cut-Dd primers) or 58.5 °C (for cut-Kp primers) for 30 s, and elongation at 72 °C for 5 s. A final denaturation ramp between 65 °C and 95 °C for 5 s was performed for the melting curve analysis. Moreover, specificity of qPCR reaction was confirmed by checking the presence of only one amplification and of the expected size in electrophoresis on a 2% agarose gel. A total of 48 fecal samples were analyzed. Each sample was analyzed with each primer set in duplicate. The $2^{-\Delta\Delta Ct}$ method was used for the relative quantification of *cutC* gene, using the EUB panbacterial primers [Muyzer] targeting the 16S rRNA gene as reference. Data were reported

as relative increase of *cutC* copy number compared to the level of the sample that showed the highest significant Ct in qPCR set as 1.

4.1.3.7. Analysis of the bacterial taxonomic composition of fecal samples

The bacterial community structure of the fecal microbiota was analyzed as described elsewhere (Cattaneo et al., 2019; Gargari et al., 2018), with DNA extracted from feces as described in Section 4.2.3.2. In brief, extracted DNA was analyzed through 16S rRNA gene profiling. Sequencing reads were generated at the Institute for Genome Sciences (University of Maryland, School of Medicine, Baltimore, MD, USA) with Illumina HiSeq 2500 rapid run sequencing of the V3–V4 variable region. Sequencing reads were equally distributed among the samples. Sequences were filtered and trimmed based on their quality. We obtained a sequence length of 301 bp for both R1 and R2 sequences with an average quality score (Phred score) higher than 35. Sequencing reads were rarefied at 5000 per sample. Subsequently, sequence reads were analyzed through the bioinformatic pipeline Quantitative Insights into Microbial Ecology (QIIME) version 1.9.1 (Caporaso et al., 2010) with the GreenGenes database updated to version 13.5. The relative abundance of bacteria in each fecal sample was reported at the taxonomic levels of phylum, class, order, family, genus, and operational taxonomic units (OTUs). Sequences were deposited in the European Nucleotide Archive (ENA) of the European Bioinformatics Institute under accession code PRJEB34169.

4.1.3.8. TMAO quantification in urine samples

TMAO levels in urine samples were determined by ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS) (Waters Acquity UPLC system). The analysis method involved

the use of a totally porous column with stationary C8 stable bond (Agilent Poroshell C8-SB) and a mobile phase consisting of a gradient acetonitrile and formate buffer (3mM of ammonium formate and 0.1% formic acid). The UPLC system was equipped with a triple quadrupole detector, which allowed the development of a "multiple reaction monitoring" (MRM) method for the analysis of TMAO. In detail, once thawed at room temperature and after centrifugation at 6000 rpm for 5 min, 25 μl of urine sample was diluted in 950 μl of UPLC mobile phase (1/1 (*v*/*v*) acetonitrile/ultra-pure sterile water + 0.025% of formic acid), and 25 μl of deuterated internal standard solution (1 ppm, TMAO-d9, Spectra 2000) was used for the normalization of results (Johnson, 2008). The UPLC samples were prepared mixing 950 μl of mobile phase (1/1 (*v*/*v*) acetonitrile/ultra-pure sterile water + 0.025% formic acid), 25 μl of urine sample, and 25 μl of deuterium-labeled methyl d9-TMAO solution (1 ppm; Spectra 2000 S.r.l., Roma, Italy). Mobile phase: 1/1 (*v*/*v*) acetonitrile/ultra-pure sterile water + 0.025% of formic acid. The run time per sample was 8 min. Sample freezing and thawing or their prolonged storage at room temperature did not have an impact on the TMAO quantification. A triple set of working standards of TMAO (trimethylamine N-oxide dihydrate, Fluka) at concentrations of 5 ppm, 50 ppm, 100 ppm was prepared according to the method described above, replacing the 25 μl of urine sample with 25 μl of standard solution. The average response factor was used for calculation.

4.1.3.9. Statistical analysis

Statistical analyses of data were carried out using R statistic software (version 3.4.2). Concerning *cutC* gene and TMAO data, intrasubject variability was defined "high" when variance among the three replicates results were higher than twice the median of all variances. Correlation

analyses were performed using the Kendall and Spearman formula with the items specified in the text as predictors and dependent variables. Significance was set at $p \leq 0.05$, and mean differences in the range $0.05 < p < 0.10$ were accepted as trends. To find associations among TMAO levels, bacterial taxa relative abundance, and *cutC* gene abundance, the machine learning supervised linear mixed model (LMM) algorithm was used. In brief, the LMM was performed using "lmer" function in the "lme4" library (Bates et al., 2015). All samples were used in the LMM analysis (n = 48), considering that three measurements were available for each subject. The Akaike's Information Criterion (AIC) was used to test the goodness of fit of the LMM. The AIC index/value depends on the ANOVA test results between two models: The model that considered the effect of the predictors and the null model.

4.1.4. Results

4.1.4.1. Distribution of the cutC gene among bacterial taxa

According to the literature, the ability of intestinal bacteria to convert the choline moiety to TMA is primarily associated with a recently discovered choline utilization (cut) genetic region harboring the *cutC* gene, which encodes a glycyl radical enzyme catalyzing C–N bond cleavage (Craciun and Balskus, 2012; Martinez-del Campo, 2015). For this reason, we designed primers specifically targeting the *cutC* gene. These primers were intended for quantitative PCR (qPCR) experiments, and we avoided the use of degenerations in their sequence. In contrast, to target all putative *cutC* sequences identified in GenBank, we clustered the putative *cutC* genes into two groups (named Dd and Kp) according to sequence similarity (Supplementary **Figure S1**) and designed a pair of primers for each group in the most conserved sequence regions. Group Dd included putative *cutC* genes from *Firmicutes* (*Anaerococcus*, *Clostridium*,

Enterococcus, *Streptococcus*), *Proteobacteria* (*Desulfotalea*, *Desulfovibrio*, *Enterobacter*), and *Actinobacteria* (*Olsenella*). Group Kp comprised putative *cutC* gene sequences from *Proteobacteria* (*Aeromonas*, *Enterobacter*, *Erwinia*, *Escherichia*, *Klebsiella*, *Pectobacterium*, *Pelobacter*, *Proteus*, *Providencia*, *Raoultella*, *Serratia*) and *Firmicutes* (*Desulfosporosinus*, *Enterococcus*). Subsequently, the two primer sets were used in endpoint PCR reactions to test the presence of putative *cutC* genes within the genomic DNA isolated from the pure cultures of 64 bacterial strains. We obtained an amplicon of the expected size from seven strains. Specifically, strains *Streptococcus dysgalactiae* 485, *S. dysgalactiae* 486, and *S. dysgalactiae* A1.3 gave a band of the expected size with primers cut-Dd. In addition, strains *Enterococcus gilvus* MD179, *Enterococcus hirae* MD160, *Klebsiella oxytoca* MIMgr, and *Klebsiella* sp. MIMgr were positive with primers cut-Kp (**Figure 1A, B**). MS and NMR analyses revealed the ability to metabolize choline and produce TMA only for the same seven strains that resulted in positive PCR experiments (**Figure 1C** and Supplementary **Figure S2**).

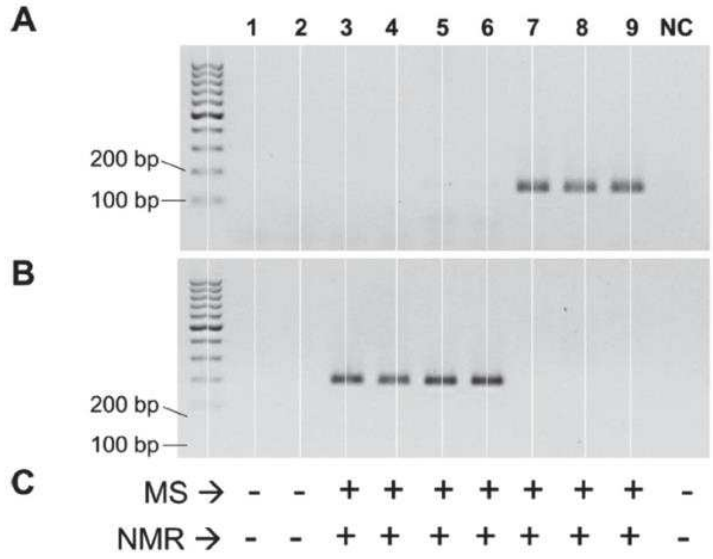


Figure 1. Detection of the choline-utilization activity in pure bacterial cultures. Panels (**A, B**) represent agarose gel resulting from end-point PCR with primers cut-Dd (**A**) and cut-Kp (**B**). Panel (**C**) summarizes the detection of TMA in cell-free broth by mass spectrometry (MS) and nuclear magnetic resonance (NMR); +, TMA detected; -, TMA not detected. Lanes: **1**, *Escherichia coli* 3.1; **2**, *Lactococcus garvieae* FMBgr; **3**, *Enterococcus gilvus* MD160; **4**, *Enterococcus hirae* MD179; **5**, *Klebsiella oxytoca* MIMgr; **6***, Klebsiella* sp. A1.2; **7**, *Streptococcus dysgalactiae* 485; **8**, *Streptococcus dysgalactiae* 486; **9**, *Streptococcus dysgalactiae* A 1.2; **NC**, negative control (i.e., M17 broth incubated without bacteria).

4.1.4.2. Bacterial taxonomic structure of the fecal microbiota

The metagenomic DNA isolated from the feces collected at three time points from 16 healthy adults (n = 48) was used in 16S rRNA gene profiling experiments. A total of 12,588,795 filtered high-quality sequence reads were generated with an average of 13,340 ± 8677 (mean ± standard deviation; max-min 11,594–4570) per sample. We failed to stratify samples according to the 16S rRNA gene profiling data, indicating that fecal bacterial community structure was homogeneous among samples and among subjects (Supplementary **Figure S3**). In addition, we also observed that the overall composition of the fecal microbiota in each subject remained mostly stable over the three collection time points

(Supplementary **Figure S3**). Globally, 182 bacterial genera were estimated, with a minimum of 36 and a maximum of 98 genera per fecal sample. *Bacteroides* was the most prevalent genus, followed by four genera of the order *Clostridiales* (undefined *Ruminococcaceae*, undefined *Lachnospiraceae*, *Ruminococcus*, and *Faecalibacterium*) (Supplementary **Figure S4A**). At the family level, most of the reads were ascribed to only three families, i.e., *Ruminococcaceae*, *Bacteroidaceae*, and *Lachnospiraceae* (Supplementary **Figure S4B**).

4.1.4.3. Putative cutC *genes in human fecal metagenomic DNA*

In order to investigate the presence of *cutC* genes in the human gut microbiome, the cut-Dd and cut-Kp primer sets were used in qPCR experiments using the same fecal metagenomic DNA as a template from healthy adults used for microbiota profiling. All analyzed fecal samples gave a positive signal in qPCR with both primer pairs (**Figure 2**). In general, cut-Kp was detected at a higher relative concentration than cut-Dd (median $\Delta\Delta$Ct of 5.33 and 0.85 for cut-Kp and cut-Dd, respectively) (**Figure 2A, B**). In addition, with both cut-Kp and cut-Dd, six volunteers out of 16 showed a variance among the three replicates that was higher than twice the median of all variances, indicating a higher intrasubject variability (**Figure 2A, B**).

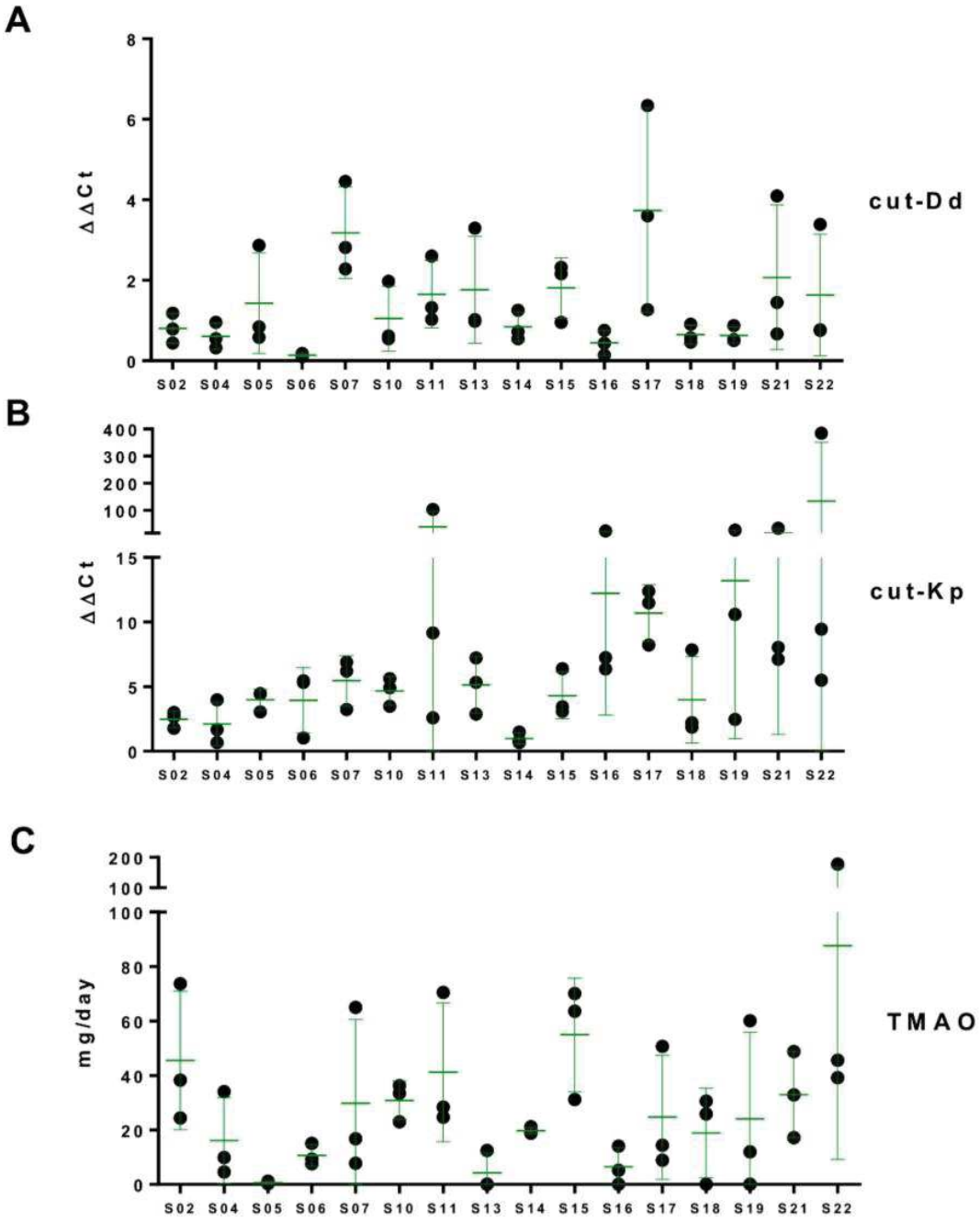


Figure 2. Fecal levels of the *cutC* gene and daily urinary excretion of trimethylamine-N-oxide (TMAO). The relative abundance of *cutC* was determined by quantitative real-time polymerase chain reaction (qPCR) with the primer pair cut-Dd-F/R (panel **A**) and cut-Kp-F/R (**B**). The TMAO concentration was determined by ultra-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS) in urine collected over 24 h (**C**). Green bars represent the mean ± standard deviation of three measurements per subject.

Subsequently, we performed correlation analyses between the *cutC* abundances determined with qPCR and the 16S rRNA gene profiling data to find potential relationships between the choline TMA-lyase genes and specific bacterial taxa of the fecal microbiota. To this end, we used the median relative abundance of bacterial taxa in fecal samples as predictors, whereas the dependent variables considered were the median abundances of cut-Dd and cut-Kp determined by qPCR per subject.

We found that cut-Dd was positively correlated with taxa belonging to the phylum Firmicutes, including an undefined *Mogibacteriaceae* genus, *Oscillospira*, and the family *Christensenellaceae*. On the contrary, cut-Dd was negatively correlated with the Firmicutes order *Bacillales*, the *Firmicutes* genus *Streptococcus*, and the *Proteobacteria* genus *Haemophilus* (Supplementary **Figure S5**). Conversely, cut-Kp was positively

associated with *Proteobacteria*. In particular, inside this phylum, a significant correlation was found with the family *Enterobacteriaceae* (Supplementary **Figure S5**).

4.1.4.4. Daily urinary excretion of TMAO

Subjects were asked to collect 24-h urine specimens the same days when the fecal samples were taken. Then, the levels of TMAO were quantified by UPLC-MS in all urine samples, revealing wide variability among the investigated healthy adults, with levels of urinary TMAO excretion ranging from less than 1 mg to more than 175 mg per day (**Figure 2**). We also observed an evident intrasubject variability in five volunteers whose TMAO excretion showed a variance among the three replicates that was higher than twice the median of all variances (**Figure 2C**). In particular, four out of the five volunteers with wide intrasubject variability (i.e., S07, S11, S19,

and S22) were found to possess high intrasubject variability for *cutC* gene levels determined in qPCR experiments (**Figure 2**).

4.1.4.5. Associations among urinary TMAO, fecal cutC*, and fecal bacterial taxa*

A linear mixed model was used to infer potential significant relationships among the datasets collected from volunteers at the three time points considered (**Figure 3**). TMAO was significantly associated with the cut-Kp/cut-Dd synergy ($p < 0.001$). Furthermore, studying the association of the single *cut* gene types, we observed that the relationship with TMAO was mainly determined by cut-Kp (**Figure 3**). In addition, we found a significant association between TMAO and 23 operational taxonomic units (OTUs). Conversely, cut-Kp and cut-Dd were significantly associated with 18 and eight OTUs, respectively. Notably, most of the OTUs that were significantly associated with cut-Kp (i.e., 15 out of 18) were also associated with TMAO, confirming the relationship between these two variables. Nine of the identified OTUs belonged to the phylum *Bacteroidetes*, while the remaining 21 were ascribed to *Firmicutes*. In addition, 80% of the OTUs (n = 24) belonged to only three families: *Bacteroidaceae*, *Lachnospiraceae*, and *Ruminococcaceae*. In particular, the most significant association (i.e., $p < 0.001$) referred to *Bacteroides caccae*, an undefined *Lachnospiraceae* genus, and several undefined *Ruminoccaceae* species (for TMAO and cut-Kp), *Bacteroides fragilis*, and an undefined *Clostridiales* species (for cut-Kp only) and an *Oscillospira* species (for cut-Dd) (**Figure 3**).

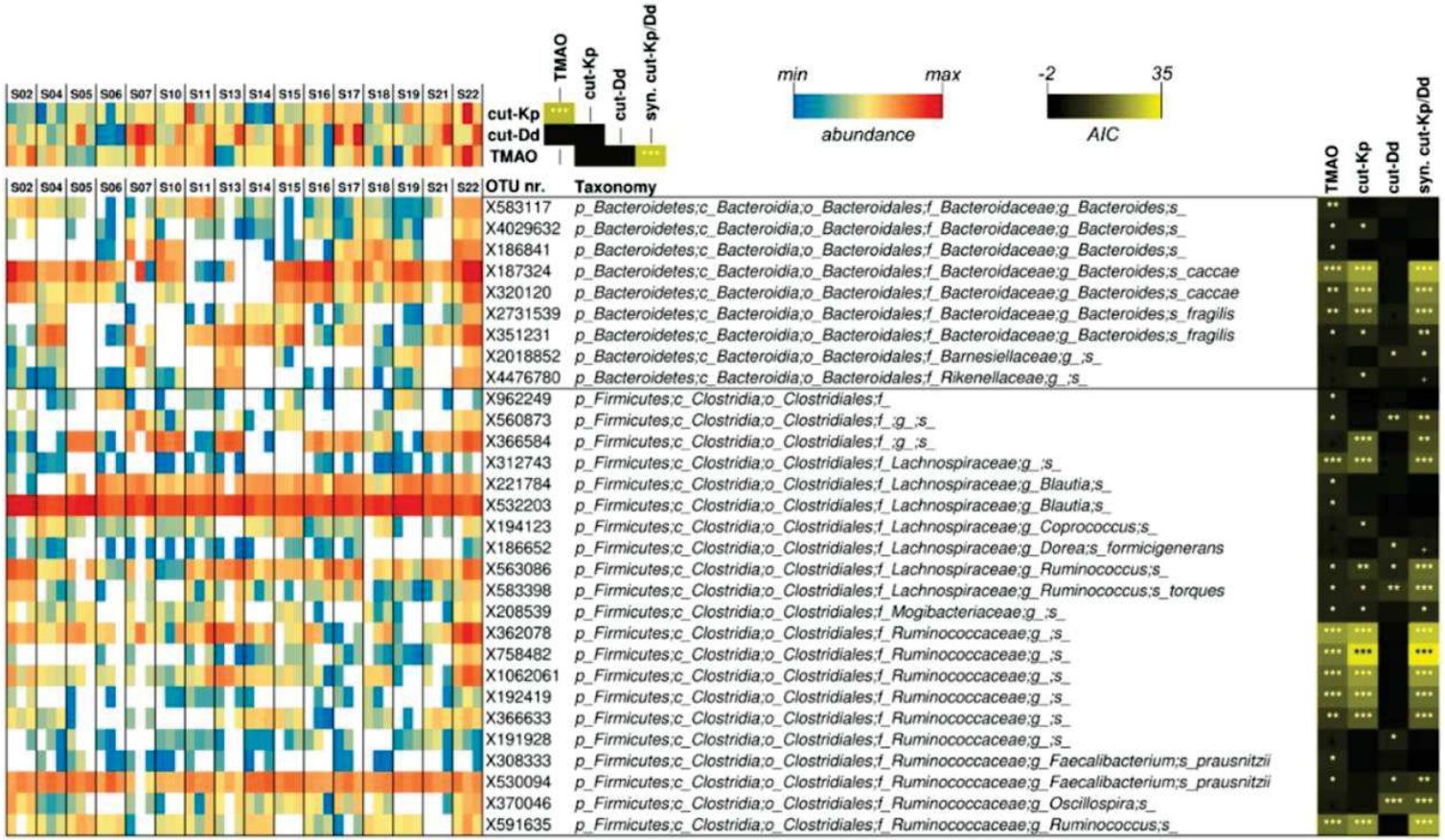


Figure 3. Analysis of the associations among fecal *cutC* gene abundances, fecal bacterial operational taxonomic units (OTUs), and urinary excreted TMAO carried out through a linear mixed model (LMM). Only OTUs that showed a significant association with *cutC* or TMAO are reported. The heatmap on the right represents TMAO levels, and *cutC* gene and OTU relative abundances. White boxes in the blue-yellow-red heatmap indicate that the OTU was not detected in that specific sample. The taxonomic lineage of each taxon is shown: p, phylum; c, class; o, order; f, family; g, genus; s, species. The black-yellow heatmap represents the Akaike's information criterion (AIC) values of the LMM analysis. Asterisks indicate significant associations: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; $^{+}$, $p < 0.1$. syn. cut-Kp/Dd = synergy between cut-Kp and cut-Dd in LMM analysis.

4.1.5. Discussion

A growing number of studies have linked host TMAO levels to different diseases or prepathological metabolic states (Chhibber-Goel et al., 2017; Janeiro et al, 2018). Conversely, TMAO has also been proposed as a beneficial factor that may promote protein stabilization and protect cells from osmotic and hydrostatic stresses according to a compensatory response mechanism (Ufnal and Nowinski, 2019). The biological role of TMAO is therefore still debated. Nonetheless, a growing number of scientific studies have suggested that this molecule may play an important role in health and diseases (Nowinski and Ufnal, 2018).

It has been suggested that an important contribution to the hepatic production of TMAO is given by the TMA produced in the gut by microbial degradation of TMA-containing dietary molecules (Zarzycki et al., 2015). In particular, TMAO levels and their physiological consequences were shown to be significantly affected by the TMA derived from choline (Craciun and Balskus, 2012). In this context, we studied the levels of TMAO excreted daily with urine, the composition of the intestinal microbiota, and the abundance of the choline TMA-lyase gene *cutC* in a group of healthy adult subjects with an Italian dietary pattern. The aim of this observational study was to verify whether TMAO levels excreted with the urine might be associated with the relative abundance of specific bacterial taxa and the bacterial gene *cutC* in feces. Literature focusing on the relationship among these three elements, particularly in non-diseased populations, is limited and partially contradicting (Craciun and Balskus, 2012; Rath et al., 2017; Martinez-del Campo et al., 2015, Xu et al., 2017).

The gene *cutC*, encoding the lyase enzyme essential for the conversion of choline into TMA (Jameson et al., 2016a), is not evenly distributed across bacterial taxa due to gene loss and horizontal gene transfer events that differently involve strains within the same species (Craciun and Balskus, 2012; Martinez-del Campo 2015; Falony et al., 2015).

Therefore, predicting the choline degradation potential of a microbial ecosystem solely based on the taxonomic composition has many intrinsic limitations. The use of primers selectively targeting a specific enzymatic conserved domain may overcome this problem, permitting the selective quantification of the abundance of a gene coding for a specific enzymatic activity in the metagenomic DNA. A similar approach was used by Martinez-Del Campo et al., (2015) who designed degenerate primers for the PCR amplification of the *cutC* gene from fecal metagenomic DNA and single strains. The use of degenerate primers was necessitated by the fact that the CutC protein possesses sequence heterogeneity. In particular, Martinez-Del Campo et al. showed that the amino acid sequences deduced from the predicted bacterial *cutC* genes can be clustered into two groups (clades 1 and 2) (Martinez-del Campo et al., 2015), which correspond to the CutC types I and II identified by Jameson et al. within a neighbor-joining phylogenetic tree constructed from amino acid sequences of glycyl radical enzymes (Jameson et al., 2016a). The same result was found in our study by generating a distance tree based on the nucleotide sequences of putative *cutC* genes (Supplementary Figure S1). In particular, cluster cut-Dd corresponded to clade 1 and CutC type I, whereas cut-Kp included sequences coding for putative proteins found in clade 2 and CutC type II reported by the authors Martinez-del Campo et al. (2015) and Jameson et al. (2016a), respectively.

For this reason, we developed two nondegenerate primer pairs located at the level of the catalytic site of the encoded enzyme that were useful for the amplification in (q)PCR experiments of the two clusters of the gene *cutC*.

When the two primer sets were used with the DNA of single strains, the only positive amplification signals were obtained with the bacteria that demonstrated the ability to metabolize choline in the biotransformation assay and produce TMA, confirming the suitability of these molecular

probes to target choline-TMA-converting bacteria. Specifically, the bacterial strains identified here as able to degrade choline to TMA include species previously confirmed to exert this conversion, such as *Streptococcus dysgalactiae* (Martinez-del Campo et al., 2015). In addition, we found *Klebsiella oxytoca*, which was reported to harbor a putative cut gene cluster (Zarzycki et al., 2015), but has never been confirmed phenotypically. We also identified two positive *Enterococcus* strains. Reportedly, TMA production from choline has also been described for some enterococci, but not for the species *E. gilvus*, which is often isolated from food matrices, including meat, milk, and cheeses (Francalanzza et al., 2007; Colombo et al., 2010), and for the zoonotic pathogen *E. hirae* (de Jong et al., 2019).

The qPCR experiments conducted showed that putative bacterial *cutC* genes were present in the fecal samples of all healthy adult subjects investigated. The high prevalence of this bacterial gene in the human gut microbiome was reported in a previous study, in which the presence of *cutC* homologs was observed in 96.6% of the assembled stool metagenomes of healthy individuals from the Human Microbiome Project (HMP) (Martinez-del Campo et al., 2015).

Reportedly, most of the TMA produced in the gut is absorbed into the portal circulation by passive diffusion (Bennett et al., 2013). Then, approximately 95% of the absorbed TMA is oxidized in the liver by flavin monooxygenases and excreted in the urine within 24 h (Zhang et al., 1995; Zeisel et al., 2017). Therefore, in this study, we performed a quantification of TMAO levels in urine samples obtained by 24-h collection.

The data presented here revealed a marked variability of both *cutC* and TMAO levels over the three time points considered in approximately 40% of volunteers. This instability was plausibly due to the variability of the daily food consumption of each subject. In this study, volunteers were free to follow their usual diet. Therefore, the analysis of multiple time

points at approximately one-week intervals was useful to address the observed temporal instability of these parameters. To the best of our knowledge, this is the first work to report the stability of intestinal *cutC* and urinary TMAO levels over time.

This study has several limitations:

1. First, we quantified the abundance of a gene of the intestinal microbiome without considering if and how much this gene was expressed. This could therefore limit the possibility of associating the abundance of this gene with its product.
2. Furthermore, the production of TMA, in addition to the presence of the bacterial gene that allows its production (*cutC*), depends on the availability of the choline substrate, which mainly comes from the diet.
3. Nonetheless, the contribution to the TMA produced in the intestine and, consequently, to the TMAO generated in the liver, derives from different chemical moieties (mainly choline, betaine, and carnitine) and includes different carnitine monooxygenase CntAB and the glycine betaine reductase GrdH, in addition to the choline TMA-lyase *cutC* (Jameson et al., 2018).
4. In addition, TMAO urinary levels may also depend on host factors that may largely vary from subject to subject, such as (i) the gut-to-blood barrier permeability to TMA (Ufnal and Pham, 2017), (ii) the oxidation of TMA in the liver by flavin monooxygenase (Wang et al., 2011), and the kidney function (Johnson et al., 2018).
5. Finally, TMAO can also be ingested directly from foods such as fish and seafood, which are naturally rich in this molecule (Seibel and Walsh, 2002).

However, despite the limitations described above, this study showed that changes in urine TMAO levels are associated with changes in the fecal

abundance of the *cutC* gene and variations in the relative abundance of several bacterial taxonomic units of the fecal microbiota. In particular, TMAO was significantly associated with the levels of a specific subcategory of the *cutC* gene, which we named cut-Kp here. This result could be explained by the relative abundance of cut-Kp, which, by qPCR results, was approximately six-times higher than that of cut-Dd. According to correlation analysis, the most important contribution to cut-Kp gene abundance is provided by *Proteobacteria*, particularly by *Enterobacteriaceae*. This result is supported by the fact that cut-Kp has been quantified with primers designed on a cluster of gene sequences having the *cutC* of the *Enterobacteriaceae* species *K. pneumoniae* as a reference. Reportedly, the analysis of human gut metagenomes revealed a high proportion of the genera *Klebsiella* and *Escherichia*, which harbor three potential TMA-producing pathways, suggesting the importance of these bacteria for TMA cycling in the human gut (Jameson et al., 2016b). Most OTUs that were found to be significantly associated with TMAO also had cut-Kp, confirming the relationship between TMAO and cut-Kp levels. A few OTUs were also associated with cut-Dd. All the taxonomic units associated with TMAO and *cutC* belong to only two taxonomic orders, *Bacteroidales* and *Clostridiales*. In particular, almost all the OTUs are attributable to only three families: *Bacteroidaceae*, *Lachnospiraceae*, and *Ruminococcaceae*. Notably, these families have been identified as the most metabolically active bacteria of the human microbiota and play a dominant role in the colonic fermentation of dietary fibers (Gosalbes et al., 2011; Wilson, 2018). Reportedly, many of these bacteria do not display choline-utilization activities (e.g., cut genes have never been identified in *Bacteroidetes* and *Faecalibacterium*). Nonetheless, we can hypothesize an indirect association of these bacteria with *cutC* and TMAO based on the speculation that the higher presence of these bacteria might determine a greater utilization of the available nutritional

sources in the colon, reducing substrates for the remaining bacterial communities. The latter may then receive selective pressure for the expansion of the activities related to the metabolization of the residual energy and carbon sources such as choline, resulting in increased TMA production.

4.1.6. Supplementary material

Table S1. Bacterial strains used for the screening of the choline utilization activity. All strains were available at the culture collection of the division of Food Microbiology and Bioprocesses, Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan. Bacterial strains that displayed the choline-utilization activity are reported in red. MRS, DeMan-Rogosa-Sharpe broth; RCM, Reinforced Clostridial Medium; LB, Luria Bertani broth; cMRS, MRS supplemented with 0.05 % cysteine-HCl. MRS, M17 and reagents for LB were from Difco (Difco Laboratories Inc., Detroit, MI); RCM was from Oxoid (Basingstoke, UK).

Phylum		Bacterial strain	Growth medium	Incubation temperature
Firmicutes	1	*Carnobacterium divergens* 3b-5ba	MRS	37°C
	2	*Carnobacterium divergens* ML1-94		
	3	*Carnobacterium divergens* N14		
	4	*Carnobacterium divergens* N20		
	5	*Carnobacterium divergens* Ovb-3		
	6	*Carnobacterium maltaromaticum* F29-1		
	7	*Carnobacterium maltaromaticum* F46-1		
	8	*Carnobacterium maltaromaticum* FM-C4		
	9	*Carnobacterium maltaromaticum* ML1-95		
	10	*Carnobacterium maltaromaticum* ML1-97		
	11	*Carnobacterium maltaromaticum* N1		
	12	*Lactobacillus harbinensis* 95		
	13	*Lactobacillus helveticus* 103		
	14	*Lactobacillus acidophilus* LA5		
	15	*Lactobacillus acidophilus* NCFM		
	16	*Lactobacillus brevis* 92		
	17	*Lactobacillus casei* LMG		
	18	*Lactobacillus coryniformis* 94		
	19	*Lactobacillus delbrueckii* subsp. b*ulgaricus* MIM-Y		
	20	*Lactobacillus delbrueckii* subsp. *lactis* MIM-F		
	21	*Lactobacillus fermentum* 2		
	22	*Lactobacillus helveticus* MIMLh5		
	23	*Lactobacillus johnsonii* DSM 10533		
	24	*Lactobacillus parabuchneri* 58		
	25	*Lactobacillus paracasei* 134		
	26	*Lactobacillus paracasei* DG		
	27	*Lactobacillus paracasei* S01		
	28	*Lactobacillus paracasei* Shirota		
	29	*Lactobacillus plantarum* 93		
	30	*Lactobacillus reuteri* DSM 17938		
	31	*Lactobacillus rhamnosus* 13		
	32	*Lactobacillus rhamnosus* GG		

Phylum		Bacterial strain	Growth medium	Incubation temperature
Firmicutes	33	*Lactococcus garvieae* MIMGr	M17	37°C
	34	***Enterococcus gilvus* MD179**		
	35	***Enterococcus hirae* MD160**		
	36	*Leuconostoc mesenteroides* To 3.4		
	37	*Streptococcus agalactiae* A1.9		
	38	***Streptococcus dysgalactiae* 485**		
	39	***Streptococcus dysgalactiae* 486**		
	40	***Streptococcus dysgalactiae* A1.3**		
	41	*Weissella cibaria* CR23		
	42	*Weissella confusa* CR55		
	43	*Clostridium butyricum* DSM 10702	RCM	
	44	*Clostridium tyrobutyricum* DSM 2637		
Actinobacteria	45	*Bifidobacterium animalis* subsp. *lactis* BB12	cMRS	
	46	*Bifidobacterium bifidum* MIMBb23sg		
Proteobacteria	47	*Escherichia coli* 1.1	LB	
	48	*Escherichia coli* 1.2		
	49	*Escherichia coli* 1.3		
	50	*Escherichia coli* 2.1		
	51	*Escherichia coli* 2.2		
	52	*Escherichia coli* 2.2		
	53	*Escherichia coli* 3.1		
	54	*Escherichia coli* DSM 1003		
	55	*Escherichia coli* DSM 682		
	56	*Enterobacter agglomerans* 1.1	M17	30°C
	57	*Enterobacter agglomerans* 1.2		
	58	*Enterobacter agglomerans* 1.4		
	59	*Enterobacter agglomerans* 1.6		
	60	*Enterobacter cloacae* 1.1		
	61	***Klebsiella oxytoca* MIMgr**		
	62	***Klebsiella* sp. A1.2**		
	63	*Serratia marcescens* 1.2		
	64	*Serratia marcescens* 1.3		

Table S2. Basic characteristics of study participants.

Subject (n=16)	Sex (4F/12M)	Age (21-45)
S02	M	33
S04	F	34
S05	M	25
S06	M	28
S07	M	27
S10	M	33
S11	M	40
S13	F	26
S14	M	26
S15	M	45
S16	F	33
S17	M	30
S18	M	24
S19	F	28
S21	M	21
S22	M	24

Figure S1. UPGMA hierarchical clustering based on ClustalW alignment of amino acid sequences of the choline trimethylamine lyase CutC. Sequences have been selected as described in material and methods. The GenBank accession number of the nucleotide sequence corresponding to each item is reported on the right of the tree. +, *cutC* gene of *Klebsiella pneumoniae* Amm1 (Kalnins et al 2015). ++, *cutC* gene of *Desulfovibrio desulfuricans* ATCC 27774 (Craciun and Balskus 2012). *Forward* and *Reverse* refer to the primer pairs designed to amplify all sequences in the corresponding cluster.

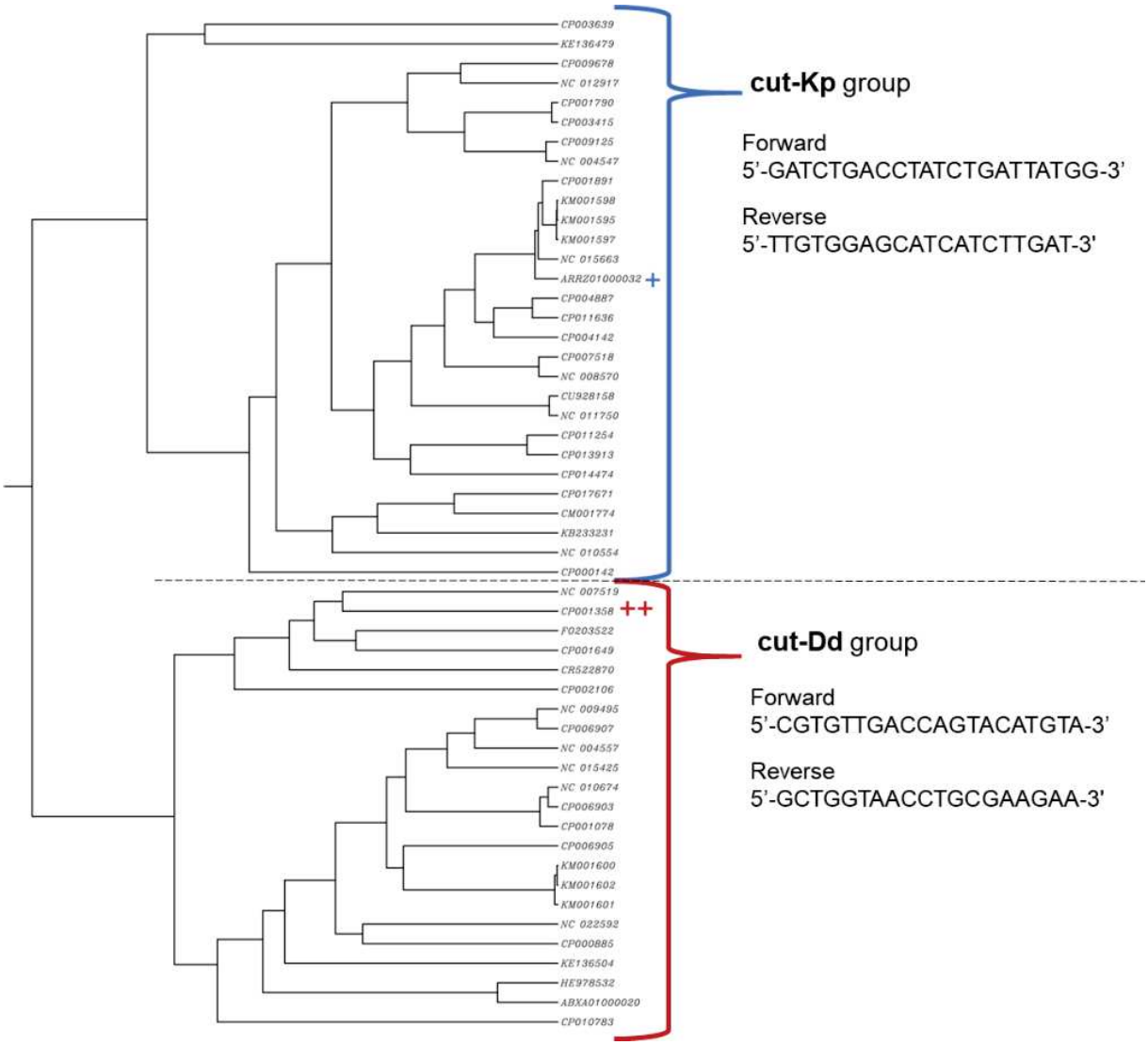


Figure S2. Verification of choline utilization and TMA production by single bacterial strains. Typical results by negative (A) and positive (B) representative strains are shown. Nuclear magnetic resonance (^{1}H-NMR) and mass spectrometry (MS) spectra are reported on the left and right, respectively. A, M17 broth supplemented with choline after incubation with *Lactococcus garvieae* MIMGr A; B, M17 broth supplemented with choline after incubation with *Klebsiella oxytoca* MIMgr; C, choline and TMA standards (50 mM in 0.1 M phosphate buffer, H_2O, pH = 6.7).

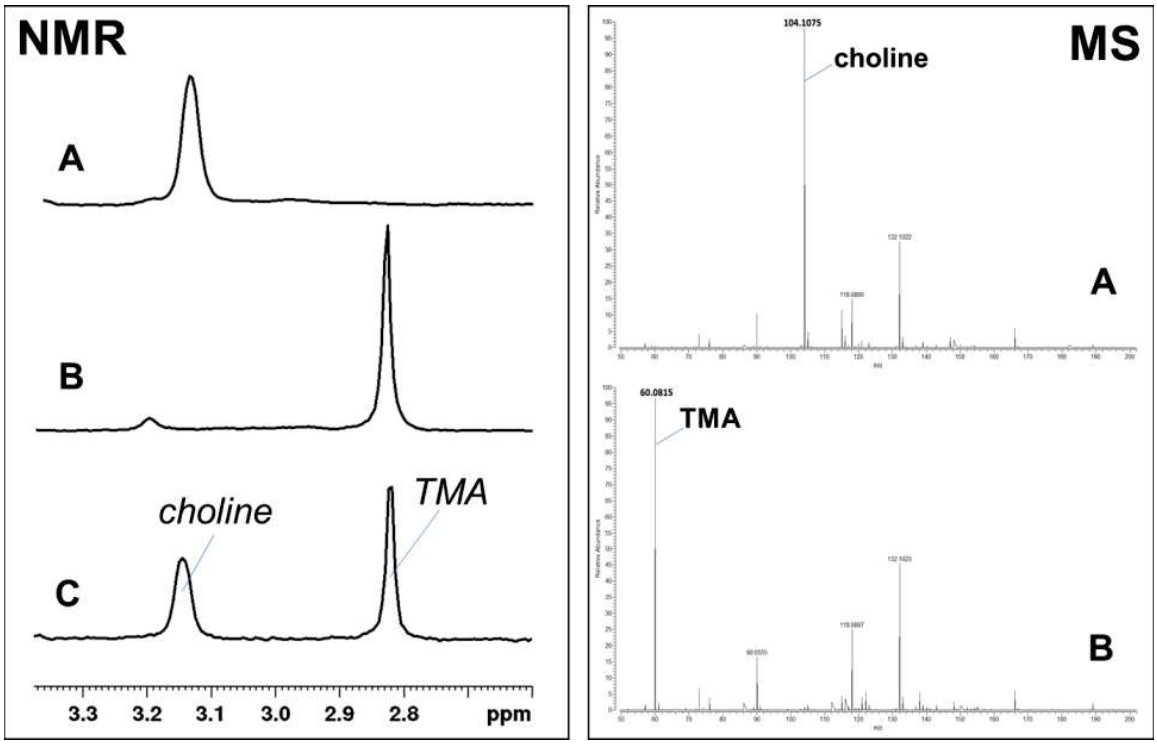
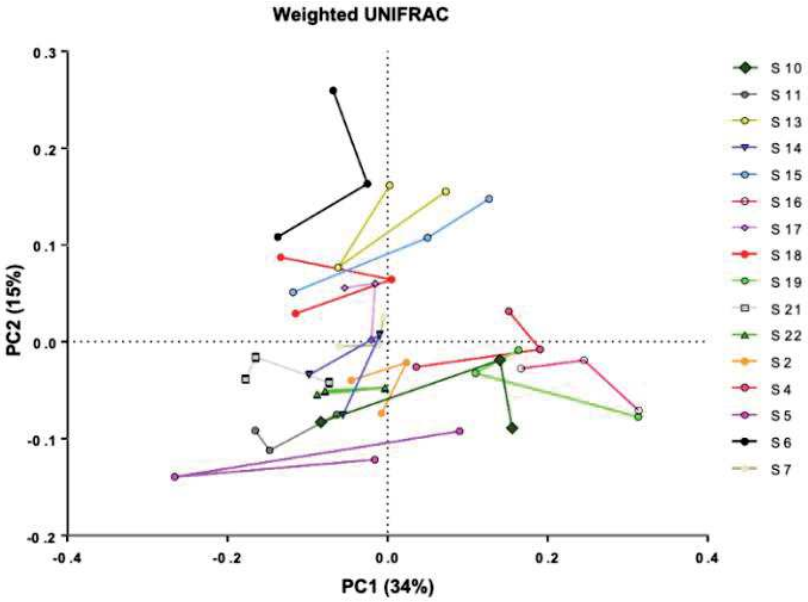


Figure S3. Bacterial community structure of fecal samples. **A**, principal coordinates analysis of weighted Unifrac distances based on 16S rRNA gene profiling data; lines connect samples from the same subject; the percentage of variance of the coordinates are explained in brackets. **B**, stacked histograms of bacterial genera in each fecal sample. The 14 most abundant bacterial genera are shown; other genera are shown in greyscale color. *und.*, undefined.

A



B

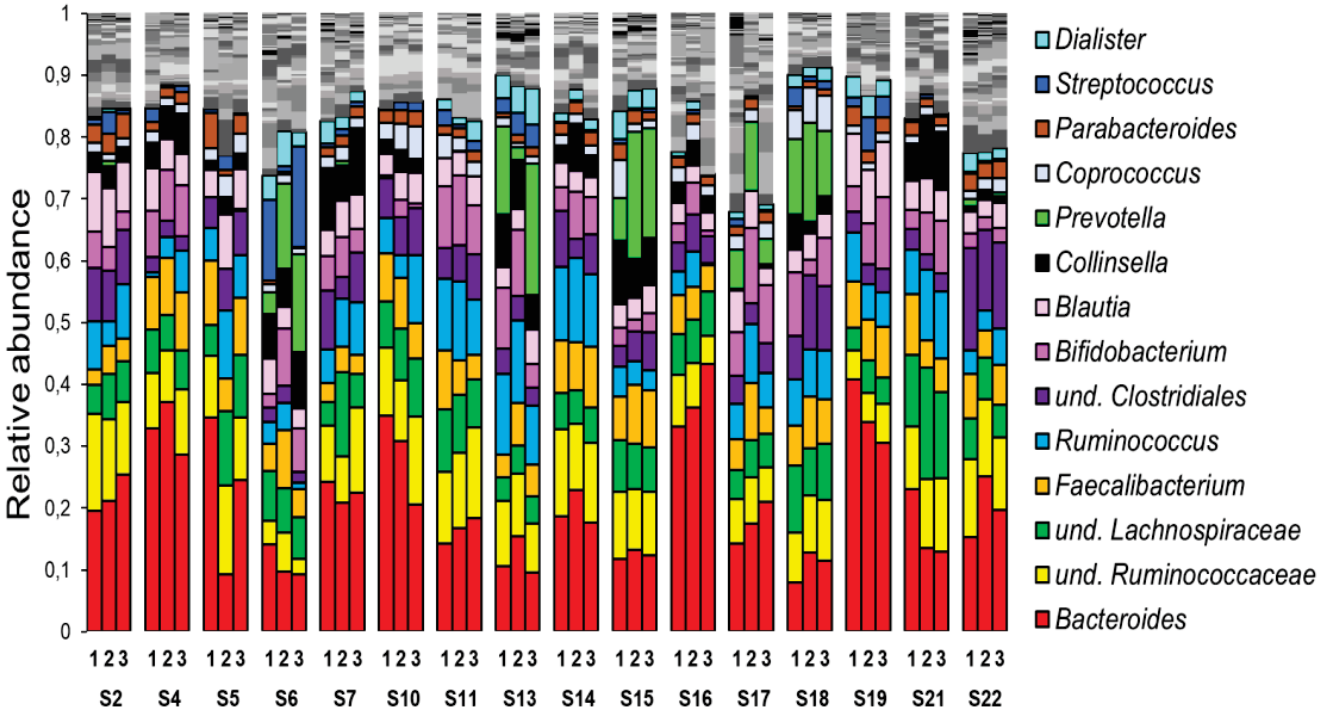


Figure S4. Tukey box and whiskers plots representing the most abundant genera (**A**) and families (**B**) detected by 16S rRNA gene profiling in fecal samples collected from the adult volunteers participating to this study.

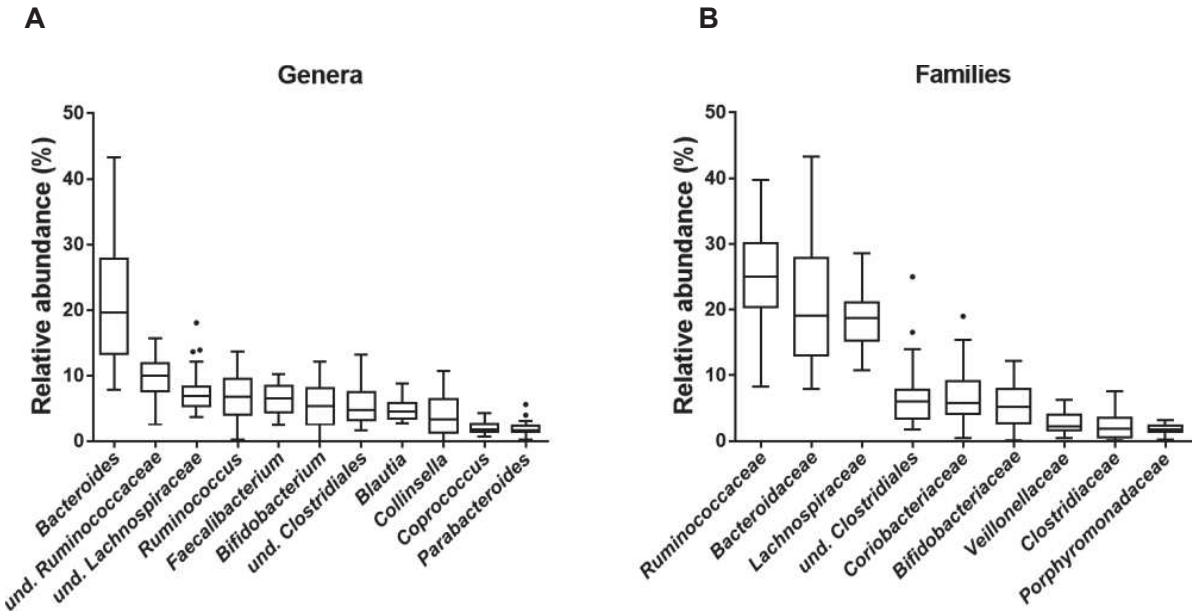


Figure S5. Correlations among the fecal relative abundances of choline TMA-lyase gene *cutC* and bacterial taxa. *cutC* abundances were determined by qPCR with primer pairs cut-Dd and cut-Kp; 16S rRNA gene profiling data were used to determine the relative abundance of bacterial taxa at the taxonomic levels of phylum (*p_*), class (*c_*), order (*o_*), family (*f_*) and genus (*g_*). The analysis was performed using median data of three measurements per subject. The heatmap represents the R value of Spearman's correlation (minimum to maximum values are indicated in heatmap legend). Asterisks indicate the Kendall rank correlation: *P < 0.05; **P < 0.01; ***P < 0.001.

Taxonomy	cut-Dd	cut-Kp
p_Firmicutes;c_Bacilli;o_Bacillales	**	
p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g_Streptococcus	*	
p_Firmicutes;c_Clostridia;o_Clostridiales;f_Mogibacteriaceae	***	
*p_Firmicutes;c_Clostridia;o_Clostridiales;f_Mogibacteriaceae;g_*	***	
p_Firmicutes;c_Clostridia;o_Clostridiales;f_Christensenellaceae	*	
p_Firmicutes;c_Clostridia;o_Clostridiales;f_Ruminococcaceae;g_Oscillospira	*	
p_Firmicutes;c_Erysipelotrichi;o_Erysipelotrichales;f_Erysipelotrichaceae;g_cc_115	*	
p_Proteobacteria		**
p_Proteobacteria;c_Gammaproteobacteria		*
p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales		*
p_Proteobacteria;c_Gammaproteobacteria;o_Enterobacteriales;f_Enterobacteriaceae		*
p_Proteobacteria;c_Gammaproteobacteria;o_Pasteurellales	**	
p_Proteobacteria;c_Gammaproteobacteria;o_Pasteurellales;f_Pasteurellaceae	**	
p_Proteobacteria;c_Gammaproteobacteria;o_Pasteurellales;f_Pasteurellaceae;g_Haemophilus	**	

-0.68 0.75

4.1.7. References

Bates, D., Mächler, M., Bolker, B., Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of statistical software, 67,* 1-48.

Bennett, B. J., de Aguiar Vallim, T. Q., Wang, Z., Shih, D. M., Meng, Y., Gregory, J., Allayee, H., Lee, R., Graham, M., Crooke, R., Edwards, P. A., Hazen, S. L., & Lusis, A. J. (2013). Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell metabolism*, *17*(1), 49–60.

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., Fierer, N., Peña, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights, D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M., Reeder, J., … Knight, R. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, *7*(5), 335–336.

Cattaneo, C., Gargari, G., Koirala, R., Laureati, M., Riso, P., Guglielmetti, S., & Pagliarini, E. (2019). New insights into the relationship between taste perception and oral microbiota composition. *Scientific reports*, *9*(1), 3549.

Chhibber-Goel, J., Singhal, V., Parakh, N., Bhargava, B., & Sharma, A. (2017). The Metabolite Trimethylamine-N-Oxide is an Emergent Biomarker of Human Health. *Current medicinal chemistry*, *24*(36), 3942–3953.

Claus, S. P., Guillou, H., & Ellero-Simatos, S. (2016). The gut microbiota: a major player in the toxicity of environmental pollutants?. *NPJ biofilms and microbiomes*, *2*, 16003.

Collins, H. L., Drazul-Schrader, D., Sulpizio, A. C., Koster, P. D., Williamson, Y., Adelman, S. J., Owen, K., Sanli, T., & Bellamine, A. (2016). L-Carnitine intake and high trimethylamine N-oxide plasma levels correlate with low aortic lesions in ApoE(-/-) transgenic mice expressing CETP. *Atherosclerosis*, *244*, 29–37.

Colombo, E., Franzetti, L., Frusca, M., & Scarpellini, M. (2010). Phenotypic and genotypic characterization of lactic acid bacteria isolated from Artisanal Italian goat cheese. *Journal of food protection*, *73*(4), 657–662.

Craciun, S., & Balskus, E. P. (2012). Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(52), 21307–21312.

Craciun, S., Marks, J. A., & Balskus, E. P. (2014). Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. *ACS chemical biology*, *9*(7), 1408–1413.

de Jong, A., Simjee, S., Rose, M., Moyaert, H., El Garch, F., Youala, M., & EASSA Study Group (2019). Antimicrobial resistance monitoring in commensal enterococci from healthy cattle, pigs and chickens across Europe during 2004-14 (EASSA Study). *The Journal of antimicrobial chemotherapy*, *74*(4), 921–930.

EFSA panel on dietetic products, nutritional and allergies (NDA). (2016). Dietary reference values for choline. *EFSA journal, 14,* 70.

Falony, G., Vieira-Silva, S., & Raes, J. (2015). Microbiology Meets Big Data: The Case of Gut Microbiota-Derived Trimethylamine. *Annual review of microbiology*, *69*, 305–321.

Fracalanzza, S. A., Scheidegger, E. M., Santos, P. F., Leite, P. C., & Teixeira, L. M. (2007). Antimicrobial resistance profiles of enterococci isolated from poultry meat and pasteurized milk in Rio de Janeiro, Brazil. *Memorias do Instituto Oswaldo Cruz*, *102*(7), 853–859.

Gargari, G., Taverniti, V., Gardana, C., Cremon, C., Canducci, F., Pagano, I., Barbaro, M. R., Bellacosa, L., Castellazzi, A. M., Valsecchi, C., Tagliacarne, S. C., Bellini, M., Bertani, L., Gambaccini, D., Marchi, S., Cicala, M., Germanà, B., Dal Pont, E., Vecchi, M., Ogliari, C., … Guglielmetti, S. (2018). Fecal Clostridiales distribution and short-chain fatty acids reflect bowel habits in irritable bowel syndrome. *Environmental microbiology*, *20*(9), 3201–3213.

Gosalbes, M. J., Durbán, A., Pignatelli, M., Abellan, J. J., Jiménez-Hernández, N., Pérez-Cobas, A. E., Latorre, A., & Moya, A. (2011). Metatranscriptomic approach to analyze the functional human gut microbiota. *PloS one*, *6*(3), e17447.

Huc, T., Drapala, A., Gawrys, M., Konop, M., Bielinska, K., Zaorska, E., Samborowska, E., Wyczalkowska-Tomasik, A., Pączek, L., Dadlez, M., & Ufnal, M. (2018). Chronic, low-dose TMAO treatment reduces diastolic dysfunction and heart fibrosis in hypertensive rats. *American journal of physiology. Heart and circulatory physiology*, *315*(6), H1805–H1820.

Jameson, E., Fu, T., Brown, I. R., Paszkiewicz, K., Purdy, K. J., Frank, S., & Chen, Y. (2016a). Anaerobic choline metabolism in microcompartments

promotes growth and swarming of Proteus mirabilis. *Environmental microbiology*, *18*(9), 2886–2898.

Jameson, E., Doxey, A. C., Airs, R., Purdy, K. J., Murrell, J. C., & Chen, Y. (2016b). Metagenomic data-mining reveals contrasting microbial populations responsible for trimethylamine formation in human gut and marine ecosystems. *Microbial genomics*, *2*(9), e000080.

Jameson, E., Quareshy, M., & Chen, Y. (2018). Methodological considerations for the identification of choline and carnitine-degrading bacteria in the gut. *Methods (San Diego, Calif.)*, *149*, 42–48.

Janeiro, M. H., Ramírez, M. J., Milagro, F. I., Martínez, J. A., & Solas, M. (2018). Implication of Trimethylamine N-Oxide (TMAO) in Disease: Potential Biomarker or New Therapeutic Target. *Nutrients*, *10*(10), 1398.

Johnson D. W. (2008). A flow injection electrospray ionization tandem mass spectrometric method for the simultaneous measurement of trimethylamine and trimethylamine N-oxide in urine. *Journal of mass spectrometry : JMS*, *43*(4), 495–499.

Johnson, C., Prokopienko, A. J., West, R. E., 3rd, Nolin, T. D., & Stubbs, J. R. (2018). Decreased Kidney Function Is Associated with Enhanced Hepatic Flavin Monooxygenase Activity and Increased Circulating Trimethylamine *N*-Oxide Concentrations in Mice. *Drug metabolism and disposition: the biological fate of chemicals*, *46*(9), 1304–1309.

Kalnins, G., Kuka, J., Grinberga, S., Makrecka-Kuka, M., Liepinsh, E., Dambrova, M., & Tars, K. (2015). Structure and Function of CutC Choline

Lyase from Human Microbiota Bacterium Klebsiella pneumoniae. *The Journal of biological chemistry*, *290*(35), 21732–21740.

Koeth, R. A., Wang, Z., Levison, B. S., Buffa, J. A., Org, E., Sheehy, B. T., Britt, E. B., Fu, X., Wu, Y., Li, L., Smith, J. D., DiDonato, J. A., Chen, J., Li, H., Wu, G. D., Lewis, J. D., Warrier, M., Brown, J. M., Krauss, R. M., Tang, W. H., … Hazen, S. L. (2013). Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nature medicine*, *19*(5), 576–585.

Langille, M. G., Zaneveld, J., Caporaso, J. G., McDonald, D., Knights, D., Reyes, J. A., Clemente, J. C., Burkepile, D. E., Vega Thurber, R. L., Knight, R., Beiko, R. G., & Huttenhower, C. (2013). Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology*, *31*(9), 814–821.

Martínez-del Campo, A., Bodea, S., Hamer, H. A., Marks, J. A., Haiser, H. J., Turnbaugh, P. J., & Balskus, E. P. (2015). Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *mBio*, *6*(2), e00042-15.

Mohajeri, M. H., Brummer, R., Rastall, R. A., Weersma, R. K., Harmsen, H., Faas, M., & Eggersdorfer, M. (2018). The role of the microbiome for human health: from basic science to clinical applications. *European journal of nutrition*, *57*(Suppl 1), 1–14.

Nowinski, A., & Ufnal, M. (2018). Trimethylamine N-oxide: A harmful, protective or diagnostic marker in lifestyle diseases?. *Nutrition (Burbank, Los Angeles County, Calif.)*, *46*, 7–12.

Rath, S., Heidrich, B., Pieper, D. H., & Vital, M. (2017). Uncovering the trimethylamine-producing bacteria of the human gut microbiota. *Microbiome*, *5*(1), 54.

Ridlon, J. M., Harris, S. C., Bhowmik, S., Kang, D. J., & Hylemon, P. B. (2016). Consequences of bile salt biotransformations by intestinal bacteria. *Gut microbes*, *7*(1), 22–39.

Roberts, A. B., Gu, X., Buffa, J. A., Hurd, A. G., Wang, Z., Zhu, W., Gupta, N., Skye, S. M., Cody, D. B., Levison, B. S., Barrington, W. T., Russell, M. W., Reed, J. M., Duzan, A., Lang, J. M., Fu, X., Li, L., Myers, A. J., Rachakonda, S., DiDonato, J. A., … Hazen, S. L. (2018). Development of a gut microbe-targeted nonlethal therapeutic to inhibit thrombosis potential. *Nature medicine*, *24*(9), 1407–1417.

Seibel, B. A., & Walsh, P. J. (2002). Trimethylamine oxide accumulation in marine animals: relationship to acylglycerol storage. *The Journal of experimental biology*, *205*(Pt 3), 297–306.

Tang, W. H., Wang, Z., Levison, B. S., Koeth, R. A., Britt, E. B., Fu, X., Wu, Y., & Hazen, S. L. (2013). Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *The New England journal of medicine*, *368*(17), 1575–1584.

Ufnal, M., & Nowiński, A. (2019). Is increased plasma TMAO a compensatory response to hydrostatic and osmotic stress in cardiovascular diseases?. *Medical hypotheses*, *130*, 109271.

Ufnal, M., & Pham, K. (2017). The gut-blood barrier permeability - A new marker in cardiovascular and metabolic diseases?. *Medical hypotheses*, *98*, 35–37.

Wallace, T. C., Blusztajn, J. K., Caudill, M. A., Klatt, K. C., Natker, E., Zeisel, S. H., & Zelman, K. M. (2018). Choline: The Underconsumed and Underappreciated Essential Nutrient. *Nutrition today*, *53*(6), 240–253.

Wang, Z., Klipfell, E., Bennett, B. J., Koeth, R., Levison, B. S., Dugar, B., Feldstein, A. E., Britt, E. B., Fu, X., Chung, Y. M., Wu, Y., Schauer, P., Smith, J. D., Allayee, H., Tang, W. H., DiDonato, J. A., Lusis, A. J., & Hazen, S. L. (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*, *472*(7341), 57–63.

Wiedeman, A. M., Barr, S. I., Green, T. J., Xu, Z., Innis, S. M., & Kitts, D. D. (2018). Dietary Choline Intake: Current State of Knowledge Across the Life Cycle. *Nutrients*, *10*(10), 1513.

Wilson, M. (2018). The indigenous microbiota of the gastrointestinal tract. *The human microbiota in health and disease: An ecological and community-based approach. 1st ed.; CRC Press, Taylor & Francis Group:Boca Raton FL, USA,* 427 p.

Xu, K. Y., Xia, G. H., Lu, J. Q., Chen, M. X., Zhen, X., Wang, S., You, C., Nie, J., Zhou, H. W., & Yin, J. (2017). Impaired renal function and dysbiosis of gut microbiota contribute to increased trimethylamine-N-oxide in chronic kidney disease patients. *Scientific reports*, *7*(1), 1445.

Zarzycki, J., Erbilgin, O., & Kerfeld, C. A. (2015). Bioinformatic characterization of glycyl radical enzyme-associated bacterial

microcompartments. *Applied and environmental microbiology*, *81*(24), 8315–8329.

Zeisel, S. H., & Blusztajn, J. K. (1994). Choline and human nutrition. *Annual review of nutrition*, *14*, 269–296.

Zeisel, S. H., Mar, M. H., Howe, J. C., & Holden, J. M. (2003). Concentrations of choline-containing compounds and betaine in common foods. *The Journal of nutrition*, *133*(5), 1302–1307.

Zeisel, S. H., & Warrier, M. (2017). Trimethylamine N-Oxide, the Microbiome, and Heart and Kidney Disease. *Annual review of nutrition*, *37*, 157–181.

Zhang, A. Q., Mitchell, S., & Smith, R. (1995). Fish odour syndrome: verification of carrier detection test. *Journal of inherited metabolic disease*, *18*(6), 669–674.

4.1.8. Authors contribution

Dalla Via, A. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Gargari, G. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Taverniti, V. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Rondini, G. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Velardi, V. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.

Gambaro, V. – Department of Pharmaceutical Sciences, University of Milan, 20122 Milan, Italy.
Visconti, G. L. – Department of Pharmaceutical Sciences, University of Milan, 20122 Milan, Italy.
De Vitis, V. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Gardana, C. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Ragg, E. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Pinto, A. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Riso, P. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Guglielmetti, S. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.

4.2. Use of kefir-derived soy-adapted lactic acid bacteria for the preparation of a fermented soy drink with increased estrogenic activity

4.2.1. Abstract

Phytoestrogens are a group of molecules deeply studied for their positive effects on human health. These compounds are naturally present in plants belonging to the *Fabaceae* family, in particular soybeans. The aim of this study was to select safe food microorganisms with the ability to ferment a soy drink resulting in increased estrogenic activity and improved functional properties. We used milk kefir grains, a dairy source of microorganisms with proven health-promoting properties, as starting inoculum for a soymilk. After 14 passages of daily inoculum in fresh soy drink, we isolated four lactic acid bacterial strains: *Lactotoccus lactis* K03, *Leuconostc pseudomesenteroides* K05, *Leuconostc mesenteroides* K09 and *Lentilactobacillus kefiri* K10. Isolated strains were proven to be safe for human consumption according to the assessment of their antibiotic resistance profile and comparative genomics. Furthermore, functional characterization of the bacterial strains demonstrated their ability to ferment the sugars naturally present in soybeans and to produce a creamy texture. In addition, we demonstrated by means of a yeast-based bioluminescence reporter system that the two strains belonging to the genus *Leuconostoc* increased the estrogenic activity of the soybean drink. In conclusion, in this study, we presented lactic acid bacterial strains that can be used for the preparation of a fermented soy-based yogurt-like product with enhanced estrogenic properties.

4.2.2. Introduction

According to the “biodiversity hypothesis” (Haahtela, 2019), during urbanization, increased antibiotic use, westernization of diet, pollution, and improved hygiene practices strongly reduced the contact between humans and microorganisms, resulting in the taxonomic impoverishment of the microbiotas associated to the human body. According to this theory, the reduced biodiversity in human-associated microbial ecosystems promoted improper immune system functioning, with consequential increased incidence of autoimmune, allergic and, in general, noncommunicable diseases (Blaser and Falkow, 2009; Haahtela, 2019). In this context, in recent years, fermented foods received growing attention as a source of live microbial cells that can positively modulate the composition of the intestinal microbiota and host physiology (Rezac et al., 2018). In addition, growing scientific evidence demonstrates that the fermentation of food, besides improving shelf-life, safety and sensory characteristics, may also enhance nutritional and health-promoting properties (Marco et al., 2016; Rezac et al., 2018). In fact, during fermentation, the microorganisms in food may produce vitamins and bioactive molecules and increase the bioavailability of food constituents (Sanlier er al., 2019). Fermentation is one of the oldest techniques adopted for the preservation and modification of food. Numerous fermented food products were demonstrated to confer health benefits, such as sauerkraut (Yu et al., 2013), kombucha (Aloulou et al., 2012), but also new products created using bacteria with proven beneficial action but never used for food fermentation (Plé et al., 2016).

A category of food molecules that attracted increasing research attention are the phytoestrogens (PEs), bioactive compounds naturally present in several vegetal foods whose structure and shape are similar to estrogens, such as estradiol. These compounds have been studied for their potential

effects in numerous hormone-associated conditions such as breast cancer (Martin et al, 1978; Cohen et al., 2000), prostate cancer (Adlercreutz et al., 1995; Setchell, 1998), cardiovascular disorders (Baum et al., 1998; Potter et al., 1998) and menopausal symptoms (Kurzer et al., 2000; Potter et al., 1998; Jayachandran and Xu, 2019). A rich source of PEs are plants belonging to the *Fabaceae* family, including soy, green peas and red clover, which contain isoflavones, which are among the first PEs discovered (Rossiter and Beck, 1966). In these plants, isoflavones are mostly present in the form of β-glycosides, acetyl glycosides and malonyl glycosides, which are much less estrogenic than their respective aglycones (Landete et al., 2016; Křížová et al., 2019).

Once ingested, isoflavones glycosides can be deglycosylated by β-glycosidases and further modified by intestinal bacteria, producing equol, dihydrodaizein and o-desmethylangolensin (Lampe et al., 2009; Wang and Murphy, 1998). The β-glucosidasic activity can also be exerted by food-associated microorganisms, such as those involved in production of tempeh, where the fermentation of soybeans takes place through the inoculum of *Rhizopus oryzae* and *Lacticaseibacillus casei* (Matsuda et al., 1994), or fuyu, tofu fermented through the use of *Mucor racemosus* (Borgia and Sypherd, 1977).

The goal of this study was the selection of safe food microorganisms with the ability to ferment a soy infusion resulting in increased estrogenic activity and improved functional properties. To this aim, we used milk kefir grains as the initial source of microorganisms, because this fermented product possesses a complex consortium of microorganisms including lactic acid bacteria, yeasts and acetic bacteria (Prado et al., 2015; Garofalo et al., 2014) with demonstrated health-promoting properties (Silva et al., 2009; de Oliveira Leite et al., 2013; Hertzler and Clancy, 2003; Turan et al., 2014; Merenstein et al., 2009; Jeong et al., 2017; Yilmaz et al., 2019). We inoculated kefir grains in soy drink until a stable microbial

consortium was obtained. Subsequently, four lactic acid bacterial strains were isolated and characterized for their potential use in the preparation of a multi-functional fermented soy drink product with enhanced estrogenic properties.

4.2.3. Materials and methods

4.2.3.1. Adaptation of kefir bacterial consortium to soy drink

5 g of granules derived from a domestic (noncommercial) kefir widespread in the north of Piedmont (Bogogno, NO, Italy) were inoculated in 50 ml commercial soy drink (water, salt, decorticated soybeans 7%) and placed at 25°C. After 24 h, the fermented product was homogenized and part of it used to inoculate other 50 ml of fresh soy drink. This operation was repeated daily for two weeks.

4.2.3.2. Isolation and identification of microbial strains from fermented soy drink

Aliquots of the fermented product were sown by spread-plating on two agarized culture media: deMan Rogosa Sharpe (MRS), pH 5.5, (Difco) and M17 (Difco) supplemented with 1% glucose and 1% lactose (Sigma-Aldrich). Then, Petri dishes were incubated at 25°C for 48 h.
Several colonies including all potentially different morphologies observed on plates were selected and transferred by streak-plating to fresh agarized media. This passage was repeated five times to obtain pure cultures. Then, bacterial isolates were grouped by molecular fingerprinting through BOX-PCR with BOXA1 primer as in (Guglielmetti et al., 2010), obtaining four genotyping groups. One representative strain from each group were taxonomically identified by sequencing the 16S rRNA gene. In brief, after cultivation in liquid medium, genomic DNA was extracted using DNEasy®

Ultraclean® Microbial® Kit (Qiagen, Hilden, Germany). The 16S rRNA gene was amplified through PCR with panbacterial primers (Suzuki and Giovannoni, 1996) and the amplicon was sequenced. Finally, BLASTN program was used to search sequence similarity within the "16S ribosomal RNA sequences (Bacteria and Archea)" database within GenBank. The identification of the subspecies within the specie *Lactococcus lactis* was performed via PCR with primers targeting the *his* operon (Corroler *et al.*, 1999).

4.2.3.3. Antibiotics resistance profile

The bacterial strains isolated from fermented soy were tested for their resistance to a panel of nine antibiotics as suggested by EFSA (2012). In details, the Minimum Inhibitory Concentration (MIC) values were evaluated for each antibiotic within different ranges, as follow: ampicillin (from 0.5 to 16 µg/ml), vancomycin (from 1 to 32 µg/ml), gentamicin (from 8 to 256 µg/ml), kanamycin (from 16 to 512 µg/ ml), streptomycin (from 8 to 256 µg/ml), erythromycin (from 0.25 to 8 µg/ml), clindamycin (from 0.25 to 8 µg/ml), tetracycline (from 1 to 32 µg/ml) and chloramphenicol (from 2 to 64 µg/ml). All antibiotics were purchased from Merck. The MICs were determined by micro-dilution method, using ISO sensitest broth (Oxoid) 90% and MRS (Difco) 10% and MIC tests were performed in a 384-well plates, filled with an automatic liquid handling system (EpMotion, Eppendorf) to a final volume of 80 µl. Bacterial cells were diluted to a concentration of $1x10^6$ cell/µl, as determined by flow cytometry (BD AccuriTM C6 Plus Flow Cytometer, BD Biosciences). Then bacteria were incubated in absence (control) and in presence of each antibiotic at six different concentrations. *Lacticaseibacillus paracasei* LMG12586 was used as reference strain according to ISO10932. Each strain was exposed, in duplicate, to each antimicrobial concentration, starting from

overnight cultures. The 384-well plates were incubated at 30°C for the four isolated strains, and 37°C for the reference strain for 24 h and the cell density evaluated by OD measuring, using a spectrophotometer (MicroWave RS2, Biotek, USA) and the Gene5 software (Biotek, USA). The MIC was determined as the lowest anti-microbial concentration that inhibited bacterial growth and the results were interpreted according to the EFSA Guidance on the assessment of bacterial antimicrobial susceptibility.

4.2.3.4. Genome sequencing, annotation and comparative analyses

The genome of the four bacterial strains isolated from fermented soy was determined by the use of an Illumina Hiseq 2500 system with paired-end and shotgun libraries. From each strain, we obtained reads length of 151 nucleotides for both R1 and R2. The number of high-quality paired-end reads (quality Phred score > 30) obtained per strain was: K03 = 5'949'531; K05 = 6'532'359; K09 = 5'331'816; K10 = 5'426'117. The SPAdes version 3.14.1 (Bankevich et al., 2012) algorithm was used for assembling reads into contigs and then in Scaffolds. The success of the assembly was tested with Bandage version 0.8 (Wick et al., 2015). General information on the obtained draft genomes is shown in **Table 1**. The virulence factor database VFDB (Chen et al., 2005) was investigated with the BLASTn algorithm using the nucleotide sequence of scaffolds as input. Using the same scaffolds as input, the antibiotic resistance sequences were also searched using the AMR-Finder algorithm on the "National Database of Antibiotics Resistant Organisms" which contains all the sequences attributable to antibiotic resistance updated to December 2020.

ID	Size	GC %	N50	L50	N°scaffold	N°CDS
***L. lactis* K03**	2'474'618	35.1	114'671	7	363	2684
***L. pseudomesenteroides* K05**	2'257'707	38.9	118'283	7	332	2420
***L. mesenteroides* K09**	2'051'111	37.6	1'183'347	1	185	2161
***L. kefiri* K10**	2'479'373	41.8	54'552	15	447	2592

Table 1. of the scaffolds for each strain isolated from fermented soy drink. CDS, coding sequence. GC%, percent of guanosine + cytosine in the genome. L50, number of nucleotides in the smallest scaffold that with all the larger ones can comprise 50% of the genome.

Sequencing data were deposited in the European Nucleotide Archive of the European Bioinformatics Institute under accession code PRJEB42630.

4.2.3.5. Fermentation of soy drink with the isolated bacterial strains

An inoculum of each individual strain was performed in the respective growth medium: M17 + 2% sucrose for strain *Lactococcus lactis* subsp. *lactis* K03, and MRS for strains *Leuconostoc pseudomesenteroides* K05, *Leuconostoc mesenteroides* subsp. *mesenteroides* K09 and *Lentilactobacillus kefiri* K10. After inoculum, broth cultures were incubated at 30°C. After 24 h, cell viability was evaluated by flow cytometry in accordance with the ISO 19344 protocol. In brief, the fluorescent dyes SYTO24 (Thermo Fisher Scientific Inc., Monza, Italy) and propidium iodide (Sigma-Aldrich, Steinheim, Germany) were added to a diluted cell suspension in saline solution (0.9% NaCl) at a final concentration of 1 μM and 2 μM respectively. The samples were then incubated at 37°C for 15 min before fluorescence measurements in the flow cytometer (BD AccuriTM C6 Plus Flow Cytometer, BD Biosciences). The bacterial strains were then individually inoculated at a concentration of $1x10^6$ viable cells/ml in a commercial soy drink (*Bevanda Biologica di Soia*, Coop; infusion of 7% decorticated soybeans in water, 1.25 g/l NaCl, pH 6.4). An inoculum

composed of the combination of the four strains at a final concentration of $1x10^6$ viable cells/ml for each strain was also used to ferment the soy drink. In addition, a strain of *Streptococcus thermophilus* was isolated from a commercial fermented soy-based product (soy yogurt) containing the same amount of decorticated soybeans (7%). This strain was cultivated in M17 medium + 2% sucrose and then inoculated in soy drink. Each inoculum was incubated in triplicate at 30°C for 24 or 48 h.

4.2.3.6. Viable bacterial count in the fermented product

The viability of bacterial cells in the fermented soy drink was determined after 24 and 48 h of incubation by serial dilution and plating in triplicate. In brief, 10-fold serial dilutions were prepared in saline, and subsequently spread on homofermentative-heterofermentative differential (HHD) agar medium (McDonald et al., 1987). Simultaneously, the pH of the fermented samples was measured at both 24 and 48 h. The same experiment was performed after 48 h of incubation for the fermented product containing the combination of the four strains.

4.2.3.7. Ultraperformance liquid chromatography - tandem mass spectrometry (UPLC-MS/MS) analyses

UPLC-MS/MS was used to identify and quantify sugars and isoflavones in soymilk products. To this aim, samples were diluted 1:1 in methanol for the analysis isoflavones, and 10:1 in acetonitrile for sugars. Then, samples were stirred at 10000 rpm for 30 s with a bench vortex and centrifuged at 11200 rcf for 5 min; then, the supernatant was taken. The standard molecules (Sigma-Aldrich) of daidzein, genistein, equol and dihydrodaidzein were solubilized in methanol, while the standards of sucrose, verbascose, raffinose and stachyose were solubilized in a solution H_2O:acetonitrile 25:75; all standards were prepared at a

concentration of 1 mg/ml. Isoflavones were analyzed using an Alliance chromatographic system mod. 2695 (Waters, Milford, MA, USA) with a diode array detector mod.2996 (Waters) and a triple quadripole mass spectrometer mod. Quattromicro (Micromass, Beverly, MA, USA). A 3.5-µm C18 X-Bridge column (150×2.1 mm, Waters) was used for the isoflavone analysis, at a flow rate of 250 µl/min. The eluents employed were: (A) a solution of ammonium acetate pH 5.2 (Sigma -Aldrich) and (B) acetonitrile (Sigma-Aldrich), these two solutions were used in linear gradient: 25% B for 8 min and 25% to 50%B in 10min. the column was maintained at 30°C, the volume injected into the LC-MS was 5 µl. The capillary voltage was set at 3.0 kV, the conical voltage at 15 V, the origin temperature was 130°C, and the desolvation temperature at 350°C. The mass spectometer operated in ESI-negative mode to detect ions with (m/z-) daidzein 253 Da, (m/z-) equol 241 Da, (m/z-) genistein 271 Da. All data has been acquired through Masslink 4.0 (Micromass).

Sugars have been analyzed by using an UPLC model Acquity (Waters) equipped with a High-Resolution MS model Exactive (Thermo Scientific, San Jose, CA, USA) operating in negative mode. A 1.7 µm BEH Amide column (150x2.1 mm, Waters, Milford, MA, USA) was used for the separation in isocratic mode at a flowrate of 0.25 ml/min. The eluent was 0.02% NH_4OH in acetonitrile: 0.02% NH_4OH in water (65:35, v/v). The column and the sample were maintained at 35 and 20°C, respectively. The Mass conditions were the following: spray voltage -3 kV, sheath gas 35, auxiliary gas 10, capillary temperature 275°C, heather 120°C, capillary voltage -37.5 V, tube lens -80 V, skimmer -16 V. All data were acquired by Xcalibur software (Thermo Scientific). Acquisition was carried out in scan mode in the range 100-600 u. Calibration curves were obtained from glucose and fructose stock solutions prepared by dissolving 20 mg of standard powder in 20 ml of water. The working solutions in

water:acetonitrile (35:65, v/v) were prepared in the range of 2-50 µg/ml. 2 ml was injected in the UPLC system. All solvents were from Sigma-Aldrich.

4.2.3.8. Estrogenic activity assay through the biosensor Saccharomyces cerevisiae *BMAEREluc/ERα*

Estrogenic activity on fermented soymilks and controls was assessed by means of the reporter yeast strain *S. cerevisiae* BMAEREluc/ERα. Reporter yeast cells were grown as previously described by Leskinen *et al.* (2005). In brief, yeast culture was cultivated in synthetic dextrose medium (SD) composed of yest nitrogen base medium (6.7 g/l) (Difco, Detroit, MI, USA) supplemented with ammonium sulfate (5 g/l), d-glucose (20 g/l), adenine (0.1 g/l), L-histidine (0.1 g/l) and L-leucine (0.26 g/l) (Sigma-Aldrich) at 30 °C. After 24 h of incubation, yeast broth culture was diluted to $OD_{600\ nm}$ 0.8 in a 96-well, white, flat-bottomed plate-96 Optiplate-96 culture plate (PerkinElmer Inc.) as described by Välimaa et al. (2009). Subsequently, 10% (vol/vol) of each sample was added to the culture at a final volume of 50 µl. Then, 50 µl of D-luciferin (Sigma-Aldrich) dissolved in 0.1 M citrate buffer pH 5.0 was added to each well. The luminescence produced was measured using a PerkinElmer Wallac VICTOR3 1420 (PerkinElmer, Monza, Italy) luminometer. Measurements were taken every 90 s after 2 s of plate shaking for 1 h, at 30°C. Bioluminescence measurement were carried out in duplicate for each sample. As positive control for estrogenic activity, zearalanone (Sigma) was used at a final concentration of 10^{-5} M. The estrogenic activity was reported as relative activation index (RAI), which was calculated as the slope of the regression line obtained by interpolating the data of the relative luminescence units (RLUs) over time (**Figure 1**). All linear regressions had a R^2>0.9. The system has been validated using daidzein, genistein and equol (Sigma-Aldrich) dissolved in ethanol the range of concentrations 10^{-5}-10^{-9} M). A

two-tailed unpaired t-test was performed to find significant differences (*p* values <0.05 were considered significant).

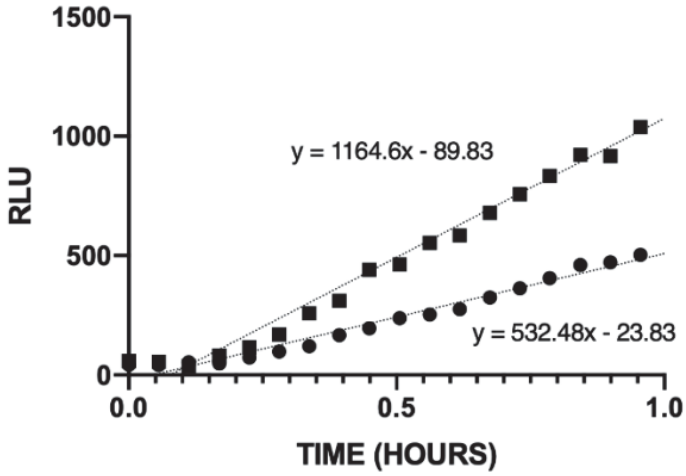


Figure 1. Examples of regression lines used to calculate the relative activation index (RAI) of soy drink samples. Two exemplificative samples are represented: sample indicated with the black square has an induction index higher (RAI=1165) than the sample indicated with the black circle (RAI=532). The bioluminescence emission data, expressed as relative luminescence units (RLUs), were measured during 1 h after the addition of luciferin.

4.2.4. Results

4.2.4.1. Adaptation of kefir grains to soy drink

Artisanal kefir grains were inoculated progressively propagated in the soy drink for 2 weeks. Proceeding with the inoculations, the grains gradually disappeared, and a homogeneous creamy product was obtained (**Figure. 2**).

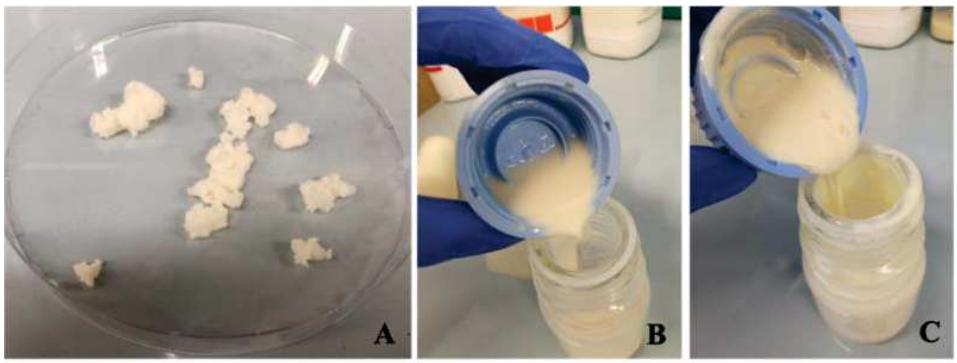


Figure 2. Fermentation of the commercial soy drink with kefir microorganisms. **A**, Kefir grains used for the inoculation of soy drink. **B**, the soy drink before inoculation. **C**, fermented soy drink obtained after two weeks of repeated inoculations. In the last image, it can be observed the absence of kefir grains and increased creaminess of the fermented product.

4.2.4.2. Isolation of soy-adapted bacterial strains

Several colonies were isolated on MRS and M17 agar plates from the fermented product obtained after two weeks of serial cultivations. BOX-PCR fingerprinting clustered the isolates into four genotypic groups, one including the isolates obtained from M17 medium, and the other three groups for MRS isolates (**Figure 3**).

Two representative colonies for each group were chosen to perform the taxonomic assignment of the isolates by 16S rRNA sequencing. Four species were identified: *Lactococcus lactis* (group II), isolated from M17, and *Leuconostoc pseudomesenteroides* (I), *Leuconostoc mesenteroides* (III) and *Lentilactobacillus kefiri* (IV) isolated from MRS. The *his*-PCR experiments confirmed that the *Lactococcus lactis* isolates belonged to the *lactis* subspecies. Then, one representative strain for each gruop was chosen: *Leuconostoc pseudomesenteroides* K05, *Lactococcus lactis* subsp. *lactis* K03, *Leuconostoc mesenteroides* K09, and *Lentilactobacillus kefiri* K10 (**Figure 3**).

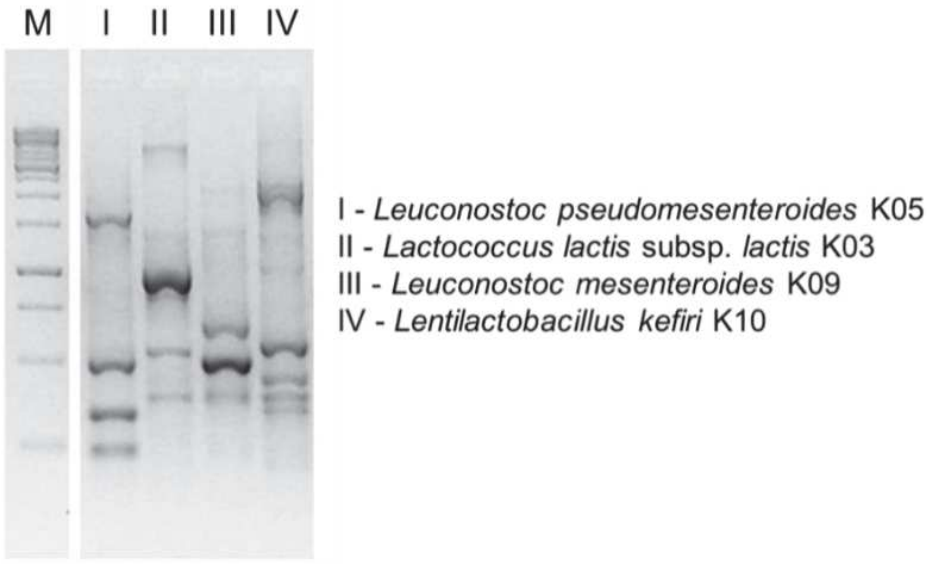


Figure 3. Genetic fingerprinting profiles obtained by BOX-PCR for the bacteria isolated from M17 and MRS agar plates. The taxonomic identification of each genotyping group has been obtained by 16S rRNA gene sequencing and is reported on the right.

4.2.4.3. Bacterial viability in fermented product

We observed that cell viability varied among the four isolated strains after growth for 24 and 48 h in soy drink (**Table 2**). The *L. lactis* K03 strain showed a decrease of one order of magnitude at the two times, from 8.1×10^8 CFU/ml at 24 h to 7×10^7 CFU/ml at 48 h. The two strains of the genus *Leuconostoc*, K05 and K09, did not change with the prolonged incubation, while strain *L. kefiri* K10 showed an increase of one order of magnitude, from 2.6×10^7 CFU/ml at 24 h to 2.7×10^8 CFU/ml at 48 h. In addition, the change of fermented product pH was very limited at the 48 compared to the 24 h incubation time (**Table 2**).

Time	*L.lactis* K03		*L.pseudomesenteroides* K05		*L.mesenteroides* K09		*L.kefiri* K10	
	UFC/ml	pH	UFC/ml	pH	UFC/ml	pH	UFC/ml	pH
24h	8.1×10^8	4.25	4.1×10^8	4.36	2.01×10^9	4.66	2.6×10^7	6.43
48h	7×10^7	4.26	2.33×10^8	4.34	1.18×10^9	4.61	2.7×10^8	6.36

Table 2. Viable counts, expressed in CFU/ml, and pH of the soy fermented products at 24 and 48 h.

Subsequently, the HHD medium was used to distinguish the four strains when used concurrently to ferment the soy drink (**Figure 4**). After 48 h of incubation, we detected 2×10^{7} CFU/ml of *L. lactis* K03, 1.9×10^{8} CFU/ml of *L. pseudomesenteroides* K05, 1.7×10^{8} CFU/ml of *L. mesenteroides* K09, and 3.3×10^{7} CFU/ml of *L. kefiri* K10. The final pH of this sample was 4.23.

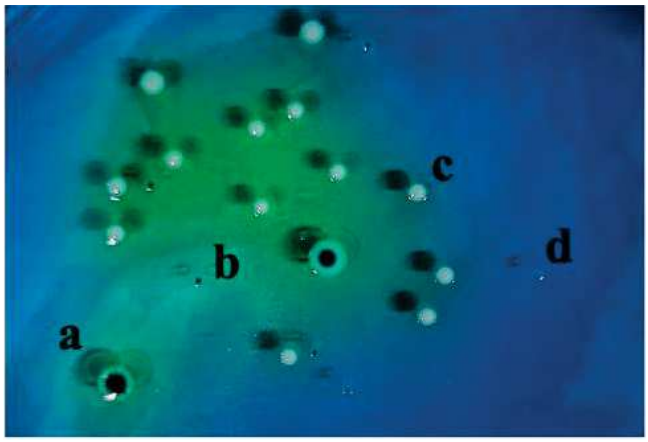


Figure 4. Colony morphology of the four strains on HHD medium, (**a**) *L. lactis* K03 (big white colonies, blue in the middle), (**b**) *L. pseduomesenteroides* K05 (small blue colonies), (**c**) *L. mesenteroides* K09 (medium-size white colonies) and (**d**) *L. kefiri* K10 (small white colonies).

4.2.4.4. Antimicrobial susceptibility testing

The minimum inhibitory concentration (MIC) values of the antibiotics are reported in **Table 3**. The test validation was performed using *L. paracasei* LMG12586, which showed no resistance to all tested antibiotics, according to the ISO10932 control values (**Table 3A**). *L. lactis* K03 showed slightly reduced susceptibility to streptomycin (64 μg/ml instead of 32 μg/ml), *L. pseudomesenteroides* K05 to kanamycin (64 μg/ml instead of 16 μg/ml) and clindamycin (4 μg/ml instead of 1 μg/ml), and *L. mesenteroides* K09 to kanamycin (128 μg/ml instead of 16 μg/ml) and chloramphenicol (8 μg/ml instead of 4 μg/ml). *L. kefiri* K10 presented a very low susceptibility to tetracycline, exceeding the highest concentration of antibiotic tested (>32 μg/ml instead of 8 μg/ml).

A

molecule	ISO10932 (ug/ml)	*L.paracasei* LMG12586 (ug/ml)
ampicillin	0.5-2	1
vancomycin	n.r.	>32
gentamicin	1-4	<8
kanamycin	16-64	64
streptomycin	8-32	32
erythromycin	0.062-0.25	<0.25
clindamycin	0.062-0.25	<0.25
tetracycline	1-4	4
chloramphenicol	4-8	8

B

molecule	*L.lactis* K03 (ug/ml)	*L.pseudomesenteroides* K05 (ug/ml)	*L.mesenteroides* K09 (ug/ml)	*L.kefiri* K10 (ug/ml)
ampicillin	<0.5	1	2	1
vancomycin	<1	>32	>32	>32
gentamicin	<0.8	<8	<8	<8
kanamycin	<16	**64**	**128**	<16
streptomycin	**64**	64	64	<8
erythromycin	<0.25	<0.25	<0.25	<0.25
clindamycin	<0.25	**4**	<0.25	<0.25
tetracycline	<1	16	8	**>32**
chloramphenicol	8	4	**8**	8

Table 3. The values of the minimum inhibitory concentrations (MIC) of the *L. paracasei* strain LMG12586 are shown with reference to the ISO10932 (**A**) values used to validate the test. MICs of the four isolated strains (**B**) are shown. Values exceeding the limits reported by EFSA for the genus tested are shown in bold. The EFSA breakpoint used are respectively for *Lactococcus lactis*, *Leuconostoc* and *Lactobacillus* facultative heterofermentative (EFSA; 2012).

4.2.4.5. Comparative genomics for the safety assessment of the bacterial strains

The draft genome of the four bacterial strains was used in comparative genomics analysis to identify genes putatively coding for virulence factors, toxins or acquired antibiotic resistance genes. This analysis did not reveal any potential antibiotic resistance genes for all tested strains. Regarding the search within the virulence factor database, the only similarities were found with genes involved in stress adaptation, primary metabolism, and capsule synthesis. No genes for toxin production or host invasion were found (**Table 4**).

ID	qseqid	pident	length	evalue	bitscore	Output (Description)	Putative function
L. lactis K03	NODE_1	74.45	1037	3.24E-115	424	(rmlB) DTDP-glucose-4,6-dehydratase, putative [Capsule (CVF186)] [Streptococcus sanguinis SK36]	Involved in biofilm formation (Yang et al., 2014; Yoshida et al., 2006)
	NODE_5	77.58	1048	4.45E-176	625	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus agalactiae 2603V/R]	Involved in secretion and maturation of cysteine protease (Wen et al., 2005)
		77.58	1048	4.45E-176	625	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus agalactiae A909]	
		77.48	1048	2.07E-174	619	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus agalactiae NEM316]	
		74.33	1235	7.88E-134	484	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS8232]	
		74.80	1111	3.67E-132	479	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS6180]	
		74.80	1111	3.67E-132	479	(tig/ropA) trigger factor, ppiase [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS9429]	
		74.80	1111	3.67E-132	479	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS315]	
		74.17	1231	3.67E-132	479	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS5005]	
		74.71	1111	1.71E-130	473	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS10394]	
		74.71	1111	1.71E-130	473	(tig/ropA) trigger factor (prolyl isomerase) [Trigger factor (CVF149)] [Streptococcus pyogenes str. Manfredo]	
		74.71	1111	1.71E-130	473	(tig/ropA) Trigger factor, ppiase [Trigger factor (CVF149)] [Streptococcus pyogenes MGAS10750]	
		74.71	1111	1.71E-130	473	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus pyogenes M1 GAS]	
		74.82	977	3.74E-117	429	(tig/ropA) trigger factor [Trigger factor (CVF149)] [Streptococcus mutans UA159]	
	NODE_6	85.82	1312	0.0E+00	1380	(SSU98_1513) phosphopyruvate hydratase [Fibronectin-binding protein (A2L15)] [Streptococcus suis 98HAH33]	Plasminogen binding factor (Bergmann et al., 2003)
		85.65	1310	0.0E+00	1365	(eno) Enolase, putative [Streptococcal enolase (CVF153)] [Streptococcus sanguinis SK36]	
		85.63	1308	0.0E+00	1362	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae A909]	
		85.43	1311	0.0E+00	1349	(eno) enolase [Streptococcal enolase (CVF153)] [Streptococcus gordonii str. Challis substr. CH1]	
		85.40	1308	0.0E+00	1345	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae 2603V/R]	
		85.32	1308	0.0E+00	1339	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae NEM316]	
		85.28	1311	0.0E+00	1338	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus pneumoniae Hungary19A-6]	
		85.29	1312	0.0E+00	1338	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus pneumoniae D39]	
		85.21	1312	0.0E+00	1332	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus pneumoniae CGSP14]	
		85.20	1311	0.0E+00	1332	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus pneumoniae TIGR4]	
	NODE_6	75.91	1017	2.99E-142	512	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase, type I [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus gordonii str. Challis substr. CH1]	Surface-displayed plasminogen-binding protein (Bergmann et al., 2004)
		75.76	1019	6.46E-139	501	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus agalactiae A909]	
		75.85	1023	6.46E-139	501	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase,plasmin receptor [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes str. Manfredo]	
		75.69	1020	2.32E-138	499	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus mutans UA159]	
		75.66	1019	3.01E-137	496	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus agalactiae NEM316]	
		75.56	1019	1.4E-135	490	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus agalactiae 2603V/R]	
		75.66	1023	1.4E-135	490	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS2096]	
		75.66	1023	1.4E-135	490	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS10750]	
		75.66	1023	1.4E-135	490	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS315]	
		75.56	1019	5.03E-135	488	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase, putative [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus sanguinis SK36]	
	NODE_9	82.18	1016	0.0E+00	865	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus agalactiae A909]	Surface-displayed plasminogen-binding protein (Bergmann et al., 2004)
		82.09	1016	0.0E+00	859	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus agalactiae NEM316]	
		82.00	1016	0.0E+00	854	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus agalactiae 2603V/R]	
		81.69	1032	0.0E+00	854	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS2096]	
		81.51	1033	0.0E+00	843	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS10394]	
		81.50	1032	0.0E+00	843	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS10750]	
		81.49	1032	0.0E+00	843	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS10270]	
		81.39	1032	0.0E+00	837	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Fibronectin-binding protein (A1210)] [Streptococcus pyogenes MGAS6180]	
		81.46	1014	0.0E+00	826	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase,plasmin receptor [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes str. Manfredo]	
		81.36	1014	0.0E+00	821	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes M1 GAS]	
		81.10	1037	0.0E+00	811	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pneumoniae R6]	
		81.16	1014	0.0E+00	809	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS8232]	
		81.16	1014	0.0E+00	809	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pyogenes MGAS315]	
		81.00	1037	0.0E+00	806	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pneumoniae CGSP14]	
		81.27	1020	0.0E+00	806	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase, type I [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pneumoniae Hungary19A-6]	
		81.06	1019	0.0E+00	795	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pneumoniae TIGR4]	
		80.96	1019	0.0E+00	789	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase, type I [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus pneumoniae D39]	
		80.75	1018	0.0E+00	778	(plr/gapA) Glyceraldehyde 3-phosphate dehydrogenase, putative [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus sanguinis SK36]	
		80.33	1022	0.0E+00	756	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus mutans UA159]	
		80.18	1019	0.0E+00	749	(plr/gapA) glyceraldehyde-3-phosphate dehydrogenase, type I [Streptococcal plasmin receptor/GAPDH (CVF123)] [Streptococcus gordonii str. Challis substr. CH1]	
	NODE_13	74.00	1154	2.44E-125	455	(tuf) translation elongation factor Tu [EF-Tu (CVF587)] [Mycoplasma mycoides subsp. mycoides SC str. PG1]	Adhesion to host extracellular matrix components (Harvey et al., 2019)

ID	qseqid	pident	length	evalue	bitscore	Output (Description)	Putative function
L. pseudomes. K05	NODE_1	74.51	1267	1.02E-139	505	(SSU98_1513) phosphopyruvate hydratase [Fibronectin-binding protein (AI215)] [Streptococcus suis 98HAH33]	Plasminogen binding factor (Bergmann et al., 2003)
		74.37	1264	4.76E-138	499	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae 2603V/R]	
		74.29	1264	2.22E-136	494	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae NEM316]	
		74.27	1267	3.71E-134	486	(eno) Enolase, putative [Streptococcal enolase (CVF153)] [Streptococcus sanguinis SK36]	
	NODE_2	75.38	1056	2.05E-131	477	(OOM_0626) elongation factor [EF-Tu (CVF827)] [Francisella noatunensis subsp. orientalis str. Toba 04]	Adhesion to host extracellular matrix components (Harvey et al., 2019)
	NODE_11	70.59	782	9.94E-30	137	(htpB) 60 kDa chaperonin (Protein Cpn60)(groEL protein)(Heat shock protein B). [Hsp60 (CVF347)] [Legionella pneumophila str. Paris]	Invasion factor (Garduño et al., 2011)
		79.08	196	1.29E-28	134	(htpB) Hsp60, 60K heat shock protein HtpB [Hsp60 (CVF347)] [Legionella pneumophila str. Corby]	
		70.36	786	2.78E-25	122	(htpB) Hsp60, 60K heat shock protein HtpB [Hsp60 (VF0159)] [Legionella pneumophila subsp. pneumophila str. Philadelphia 1]	
		77.34	128	2.84E-10	73.1	(htpB) 60 kDa chaperonin (Protein Cpn60)(groEL protein)(Heat shock protein B). [Hsp60 (CVF347)] [Legionella pneumophila str. Lens]	
L. mes. K09	NODE_1	73.99	1269	2.43E-127	466	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae 2603V/R]	Plasminogen binding factor (Bergmann et al., 2003)
		73.92	1269	1.13E-125	460	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae NEM316]	
	NODE_5	72.92	139	8.21E-42	178	(lisR) putative two-component response regulator [LisR/LisK (CVF253)] [Listeria ivanovii subsp. ivanovii PAM 55]	Play role in stress-sensing and resistance to antimicrobial agents (Cotter et al., 1999; 2002)
L. kefiri K10	NODE_16	73.26	1298	3.79E-117	427	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae A909]	Plasminogen binding factor (Bergmann et al., 2003)
		73.05	1298	3.81E-112	411	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae 2603V/R]	
		73.05	1298	3.81E-112	411	(eno) phosphopyruvate hydratase [Streptococcal enolase (CVF153)] [Streptococcus agalactiae NEM316]	
	NODE_724	87.70	299	4.66E-13	76.8	(cap5D) - [Capsule (CVF110)] [Staphylococcus aureus subsp. aureus USA300_TCH1516]	Involved in capsula synthesis (Nanra et al., 2013)
		87.70	1280	4.66E-13	76.8	(capD) polysaccharide biosynthesis protein CapD [Capsule (CVF110)] [Staphylococcus aureus subsp. aureus str. Newman]	
		87.70	1280	4.66E-13	76.8	(SaurJH1_0143) polysaccharide biosynthesis protein CapD [Capsule (CVF110)] [Staphylococcus aureus subsp. aureus JH1]	
		87.70	1268	4.66E-13	76.8	(capD) capsular polysaccharide synthesis enzyme Cap5D [Capsule (CVF110)] [Staphylococcus aureus subsp. aureus N315]	
		87.70	1268	4.66E-13	76.8	(SAOUHSC_00117) capsular polysaccharide biosynthesis protein Cap5D, putative [Capsule (CVF110)] [Staphylococcus aureus subsp. aureus NCTC 8325]	
		87.70	1268	4.66E-13	76.8	(capD) capsular polysaccharide synthesis enzyme CapD [Capsule (CVF110)] [Staphylococcus aureus RF122]	
		87.70	1268	4.66E-13	76.8	(cap8D) capsular polysaccharide synthesis enzyme Cap8D [Capsule (VF0003)] [Staphylococcus aureus subsp. aureus MW2]	
		86.15	1268	2.17E-11	71.3	(capD) capsular polysaccharide synthesis enzyme [Capsule (CVF110)] [Staphylococcus aureus subsp. aureus MRSA252]	

Table 4. The results of the search on the VFBD database are shown. ID, strain analyzed. Qseqid, query sec id. Pident, percent of nucleotides in the high score pair that match. Length, sequence length. Evalue, number of expected hits of similar quality (score) that could be found just by chance. Bitscore, similarity score. Putative function, putative function of the protein encoded by the gene.

4.2.4.6. UPLC-MS analysis of the soy drink

The chromatographic analyses evidenced the presence of daidzin and genistin in all fermented samples tested. However, the samples fermented with *L. pseudomesenteroides* K05 and *L. mesenteroides* K09 showed a decrease of daidzin concentration (104 μg/ml and 151 μg/ml respectively) compared to unfermented soymilk used as a negative control (203 μg/ml), and, consequently, the increase in daidzein concentration (80 μg/ml and 67 μg/ml) compared to the same control (3 μg/ml) (**Table 5**). Similarly, the two strains determined the decrease of genistin concentration (32 μg/ml and 21 μg/ml) compared to the negative control (53 μg/ml) and the increase of genistein (87 μg/ml and 124 μg/ml) compared to control (3 μg/ml) (**Table 5**). Similar results were obtained for the sample obtained with the combination of the four strains (mix), corresponding todaidzine 152 μg/ml, daidzeine 62 μg/ml, genistin 36 μg/ml and genistein 85 μg/ml. In the samples fermented with *L. lactis* K03, *L. kefiri* K10 and the *S. thermophilus* strain, there were no variation of the glycosylated and deglycosylated forms of the isoflavones tested (**Table 5**). Furthermore, dihydro daidzein and equol were not detected in any sample.

Chromatographic analysis also revealed that fermented products contained less sugars than unfermented soy milk. In specific, the samples fermented with *L. pseudomesenteroides* K05, *L. mesenteroides* K09*, L. kefiri* K10 and mix did not contain any of the tested sugars (sucrose, stachyose and raffinose) (**Table 6**). On the contrary, in samples fermented with *L. lactis* K03 and *S. thermophilus*, sucrose concentration was 273 μg/ml and 215 μg/ml respectively, corresponding to almost half of the sucrose concentration found in the unfermented soy drink (470 μg/ml), whereas the concentrations of stachyose and raffinose were similar to the

negative control (**Table 6**). Verbascose was not detected in any of the samples.

Isoflavones µg/ml	*L.lactis* K03	*L.pseudomesenteroides* K05	*L.mesenteroides* K09	*L.kefiri* K10	MIX	*S.thermophilus*	Soymilk
Daidzin	241	104	151	225	152	279	203
Genistin	79	32	21	68	36	89	53
Daidzein	8	80	67	4	62	5	3
Genistein	6	87	124	4	85	5	3
Dihydro Daidzein	0	0	0	0	0	0	0
Equol	0	0	0	0	0	0	0

Table 5. Concentration of isoflavones detected by UPLC-MS in fermented products and total isoflavones.

Sugars µg/ml	*L.lactis* K03	*L.pseudomesenteroides* K05	*L.mesenteroides* K09	*L.kefiri* K10	MIX	*S.thermophilus*	Soymilk
Saccharose	273	0	0	0	0	215	470
Stachyose	93	0	0	0	0	89	103
Raffinose	5	0	0	0	0	4	4.3
Verbascose	0	0	0	0	0	0	0

Table 6. Concentration of sugars detected through UPLC-MS in fermented products and unfermented soymilk.

4.2.4.7. Estrogenic activity of fermented products

The experiment validation allowed to observe that the biosensor was sensitive to the presence of genistein and equol in the same way as zearalanone, at the concentration tested (10^{-5} M), but it was not sensitive to daidzein.

The slope of the regression lines was used to evaluate the relative activation index of each sample, which was analyzed in triplicate (**Figure 4**). It was observed a significant difference between the unfermented soy drink control and all tested samples. The samples fermented with *L. lactis* K03 and *S. thermophilus* had a significantly lower activation index than the unfermented soymilk sample ($p<0.001$). On the contrary, samples fermented with *L. pseudomesenteroides* K05 and *L. mesenteroides* K09 had a higher relative activation index than the unfermented soymilk sample

(p<0.001) but, they were not significant different between them. The sample fermented with *L. kefiri* K10 was significantly different from the control sample of (p<0.05). The internal positive control, zearalenone, did not show any significant difference compared to the sample fermented with *L. mesenteroides* K09, while it was significantly different from the sample fermented by *L. pseudomesenteroides* K05.

mix

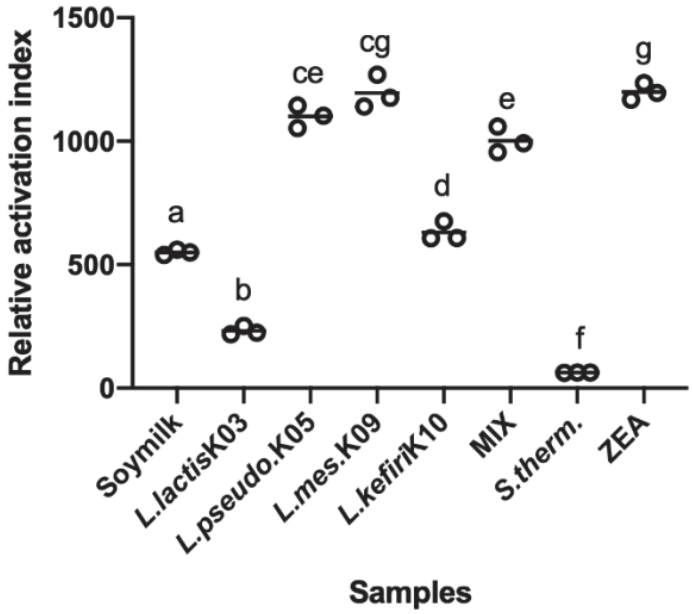


Figure 4. Dot-plot of the relative activation index for the fermented samples and controls. The experiment has been carried out in triplicate. The different letters a-g indicate the significant differences (p<0.05) found by applying the two-way unpaired Student's t-test. ZEA, zearalenone. MIX, fermented sample obtained by the combination of the four strains.

4.2.5. Discussion

In this work, we described the ability of selected food-associated lactic acid bacteria to modulate the phytoestrogenic activity of soy isoflavones.

First, we decided to inoculate kefir grains into soy drink and, after several steps, we observed a creamy and homogeneous product. Then, we isolated from the fermented soy drink four lactic acid bacterial strains belonging to EFSA QPS species list.

All isolated strains belonged to species shown to be common members of the kefir grain microbiota (Korsak et al., 2015; Kotova et al., 2016; Leite et al., 2010).

The absence of a strain capable to produce the kefiran exopolysaccharide, such as *Lactobacillus kefiranofaciens* among our isolates could be an explanation for the absence of granules in the fermented soy drink (Wang et al., 2008).

After isolating the strains, we were able to assess their growth in the soy drink. The results of the viable count, even after 48 h of fermentation, confirm a high presence of vital strains in all the final product (10^7/10^9 CFU/ml). According to literature, these concentrations are potentially adequate for the bacteria to survive the passage of the gastrointestinal tract and reach the intestine. (Derrien and van Hylckama Vlieg, 2015).

The results of the antibiotic MIC test showed reduced susceptibility of all strains to some of the tested molecules. The two *Leuconostoc* strains showed lower susceptibility to kanamycin, chloramphenicol and clindamycin. There is evidence in literature of the presence of resistance to these antibiotics for *Leuconostoc* strains dairy origin (Alegría et al., 2013; Adimpong et al., 2012; Flórez et al., 2016). *L. lactis* K03 strain showed reduced susceptibility to streptomycin, Toomey et al. (2010), in their work reported a strains of *L. lactis* resistant to the same antibiotic, which did not show any gene involved in this resistance.

L. kefiri K10 shows a reduced susceptibility to tetracycline that exceeds the maximum concentration tested. A previous work (Feichtinger et al., 2016) showed that *Lentilactobacillus buchneri*, a species phylogenetically close to *L. kefiri*, carries an intrinsic resistance to the same antibiotic.

The whole genome analysis of the four strains was performed in order to check the presence of acquired antibiotic resistance and virulence genes, according to the Pariza decision three for determining the safety of food microbial cultures (Pariza et al., 2015).

Antibiotic database search did not identify any potential antibiotic resistance gene that could have been acquired by the four tested bacteria. These results suggest that the observed antibiotic resistances above EFSA breakpoints could be associated to intrinsic non-transmissible feature, such as the presence of an abundant exopolysaccharide capsule, which is very probably produced by the bacteria here investigated.

In addition, no significant virulence factor or toxin genes were detected, suggesting that the four lactic acid bacterial strains here studied are safe for human consumption.

Concerning the naturally sugars in soy, the quantities of sucrose, stachyose, raffinose and verbascose were analyzed (Middelbos and Fahey Jr., 2008) before and after fermentation by UPLC-MS. This analysis showed that the samples containing *L. pseudomesenteroides* K05, *L. mesenteroides* K09 and *L. kefiri* K10 did not show the presence of sugars at the end of fermentation. This suggested that these strains are able to use the sugars naturally present in soybeans. The absence of these typical soy sugars can be interesting as concerns the elimination of flatulence, linked to the consumption of soy-based products. This unpleasant inconvenience is due to the lack of digestion of the oligosaccharides we tested and their fermentation in the intestine (Rackis, 1981).

Finally, we focused on the analysis of estrogenic activity. First, the concentrations of isoflavones, both in glycosylated and deglycosylated form, were evaluated by UPLC-MS soy products. The results showed an increase in deglycosylated forms of isoflavones in samples fermented with *L. pseudomesenteroides* K05*, L. mesenteroides* K09 and the mix of the four strains. In the products fermented with *L. lactis* K03, *L. kefiri* K10 and the commercial strain *S. thermophilus* the quantities of deglycosylated isoflavones are very similar to those found in unfermented soymilk. The formation of deglycosylated forms of isoflavones derives from the β-glucosidasic enzymatic activity (Ismail and Hayes, 2005), which is widely

spread in lactic acid bacteria (Michlmayr and Kneifel, 2013; Yuksekdag et al., 2017). All samples showed a total amount of isoflavones higher than the unfermented soy drink. This could suggest that acetylglucosides and malonylglucosides of isoflavones, not evaluated in this work, could also be reduced to the deglycosylated and glycosylated form (Ismail and Hayes, 2005).

We evaluated the estrogenic activity of fermented products using a yeast reporter system. This system allowed us to evaluate the response of our products to the estrogenic receptor ERα. This system was used by Välimaa et al. (2009) to detect zearalanone, an estrogenic mycotoxin produced by fungi of the genus *Fusarium*, in dairy products. Zearalanone is structurally similar to 17β-estradiol whose hormonal action is reported to be more powerful than the estrogenic activity of isoflavones (Bennett and Klich, 2003). Based on this, we adopted this system containing the receptor ERα, despite the higher affinity of isoflavones for the ERβ receptor (Morito et al, 2001). The peculiarity of this reporter system to connect ERα receptor activation to luciferase production allowed us to link estrogenic activity to the amount of light produced.

The results showed an estrogenic activity in samples fermented with the two strains of *Leuconostoc* similar to the activity of zearalanone. It was observed that the product fermented with *L. kefiri* K10 had a profile similar to the unfermented milk control, and the samples fermented with *L. lactis* K03 and *S. thermophilus* showed the lowest activity.

These data are in accordance with the results obtained analyzing the amount of glycosylated and deglycosylated isoflavones through UPLC-MS analysis. It can be noted that samples containing a higher concentration of genistein are also those with the highest estrogen response (*L. pseudomesenteroides* K05, *L. mesenteroides* K09 and the combination of four). This is also in agreement with the data related to the biosensor validation, where we noted a reporter response only for genistein and

equol. The lowest responses of *L. lactis* K03 and *S. thermophilus* samples compared to unfermented soy drink may have several explanations. First, the system has not been tested for its response to glycosylated isoflavones, not allowing us to understand how these forms can interact with it. Second, particular, the receptor does not respond for daidzein, as shown in the experiments carried out on individual compounds, presumably due to the higher specificity of phytoestrogens for the receptor ERβ.

4.2.6. References

Adimpong, D. B., Nielsen, D. S., Sørensen, K. I., Derkx, P. M., & Jespersen, L. (2012). Genotypic characterization and safety assessment of lactic acid bacteria from indigenous African fermented food products. *BMC microbiology*, *12*, 75.

Adlercreutz, C. H., Goldin, B. R., Gorbach, S. L., Höckerstedt, K. A., Watanabe, S., Hämäläinen, E. K., Markkanen, M. H., Mäkelä, T. H., Wähälä, K. T., & Adlercreutz, T. (1995). Soybean phytoestrogen intake and cancer risk. *The Journal of nutrition*, *125*(3 Suppl), 757S–770S.

Alegría, A., Delgado, S., Flórez, A. B., & Mayo, B. (2013) Identification, typing, and functional characterization of *Leuconostoc* spp. Strains from traditional, starter-free cheeses. *Dairy Science & Technology*. 93;657-673(2013).

Aloulou, A., Hamden, K., Elloumi, D., Ali, M. B., Hargafi, K., Jaouadi, B., Ayadi, F., Elfeki, A., & Ammar, E. (2012). Hypoglycemic and antilipidemic properties of kombucha tea in alloxan-induced diabetic rats. *BMC complementary and alternative medicine*, *12*, 63.

Atkinson, C., Frankenfeld, C. L., & Lampe, J. W. (2005). Gut bacterial metabolism of the soy isoflavone daidzein: exploring the relevance to human health. *Experimental biology and medicine (Maywood, N.J.)*, *230*(3), 155–170.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., & Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology: a journal of computational molecular cell biology*, *19*(5), 455–477.

Baum, J. A., Teng, H., Erdman, J. W., Jr, Weigel, R. M., Klein, B. P., Persky, V. W., Freels, S., Surya, P., Bakhit, R. M., Ramos, E., Shay, N. F., & Potter, S. M. (1998). Long-term intake of soy protein improves blood lipid profiles and increases mononuclear cell low-density-lipoprotein receptor messenger RNA in hypercholesterolemic, postmenopausal women. *The American journal of clinical nutrition*, *68*(3), 545–551.

Bennett, J. W., & Klich, M. (2003). Mycotoxins. *Clinical microbiology reviews*, *16*(3), 497–516.

Bergmann, S., Wild, D., Diekmann, O., Frank, R., Bracht, D., Chhatwal, G. S., & Hammerschmidt, S. (2003). Identification of a novel plasmin(ogen)-binding motif in surface displayed alpha-enolase of *Streptococcus pneumoniae*. *Molecular microbiology*, *49*(2), 411–423.

Bergmann, S., Rohde, M., & Hammerschmidt, S. (2004). Glyceraldehyde-3-phosphate dehydrogenase of *Streptococcus pneumoniae* is a surface-

displayed plasminogen-binding protein. *Infection and immunity*, *72*(4), 2416–2419.

Blaser, M. J., & Falkow, S. (2009). What are the consequences of the disappearing human microbiota?. *Nature reviews. Microbiology*, *7*(12), 887–894.

Borgia, P., & Sypherd, P. S. (1977). Control of beta-glucosidase synthesis in Mucor racemosus. *Journal of bacteriology*, *130*(2), 812–817.

Cassidy, A., Brown, J. E., Hawdon, A., Faughnan, M. S., King, L. J., Millward, J., Zimmer-Nechemias, L., Wolfe, B., & Setchell, K. D. (2006). Factors affecting the bioavailability of soy isoflavones in humans after ingestion of physiologically relevant levels from different soy foods. *The Journal of nutrition*, *136*(1), 45–51.

Chang, Y., & Choue, R. (2013). Plasma pharmacokinetics and urinary excretion of isoflavones after ingestion of soy products with different aglycone/glucoside ratios in South Korean women. *Nutrition research and practice*, *7*(5), 393–399.

Chen, L., Yang., Yu, J., Yao, Z., Sun, L., Shen, Y., & Jin, Q. (2005). VFDB: a reference database for bacterial virulence factors. *Nucleic acids research*, *33*(Database issue), D325–D328. , J

Cohen, L. A., Zhao, Z., Pittman, B., & Scimeca, J. A. (2000). Effect of intact and isoflavone-depleted soy protein on NMU-induced rat mammary tumorigenesis. *Carcinogenesis*, *21*(5), 929–935.

Corroler, D., Desmasures, N., & Gueguen, M. (1999). Correlation between polymerase chain reaction analysis of the histidine biosynthesis operon, randomly amplified polymorphic DNA analysis and phenotypic characterization of dairy Lactococcus isolates. *Applied microbiology and biotechnology*, *51*(1), 91–99.

Cotter, P. D., Emerson, N., Gahan, C. G., & Hill, C. (1999). Identification and disruption of lisRK, a genetic locus encoding a two-component signal transduction system involved in stress tolerance and virulence in *Listeria monocytogenes*. *Journal of bacteriology*, *181*(21), 6840–6843.

Cotter, P. D., Guinane, C. M., & Hill, C. (2002). The LisRK signal transduction system determines the sensitivity of *Listeria monocytogenes* to nisin and cephalosporins. *Antimicrobial agents and chemotherapy*, *46*(9), 2784–2790.

de Oliveira Leite, A. M., Miguel, M. A., Peixoto, R. S., Rosado, A. S., Silva, J. T., & Paschoalin, V. M. (2013). Microbiological, technological and therapeutic properties of kefir: a natural probiotic beverage. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*, *44*(2), 341–349.

Derrien, M., & van Hylckama Vlieg, J. E. (2015). Fate, activity, and impact of ingested bacteria within the human gut microbiota. *Trends in microbiology*, *23*(6), 354–366.

EFSA. (2012) Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance. *EFSA journal, 10*(6):2740.

Faughnan, M. S., Hawdon, A., Ah-Singh, E., Brown, J., Millward, D. J., & Cassidy, A. (2004). Urinary isoflavone kinetics: the effect of age, gender, food matrix and chemical composition. *The British journal of nutrition*, *91*(4), 567–574.

Feichtinger, M., Mayrhofer, S., Kneifel, W., & Domig, K. J. (2016). Tetracycline Resistance Patterns of Lactobacillus buchneri Group Strains. *Journal of food protection*, *79*(10), 1741–1747.

Fiorda, F. A., de Melo Pereira, G. V., Thomaz-Soccol, V., Rakshit, S. K., Pagnoncelli, M., Vandenberghe, L., & Soccol, C. R. (2017). Microbiological, biochemical, and functional aspects of sugary kefir fermentation - A review. *Food microbiology*, *66*, 86–95.

Flórez, A. B., Campedelli, I., Delgado, S., Alegría, Á., Salvetti, E., Felis, G. E., Mayo, B., & Torriani, S. (2016). Antibiotic Susceptibility Profiles of Dairy Leuconostoc, Analysis of the Genetic Basis of Atypical Resistances and Transfer of Genes In Vitro and in a Food Matrix. *PloS one*, *11*(1), e0145203.

Garduño, R. A., Chong, A., Nasrallah, G. K., & Allan, D. S. (2011). The *Legionella pneumophila* Chaperonin - An Unusual Multifunctional Protein in Unusual Locations. *Frontiers in microbiology*, *2*, 122.

Garofalo, C., Osimani, A., Milanović, V., Aquilanti, L., De Filippis, F., Stellato, G., Di Mauro, S., Turchetti, B., Buzzini, P., Ercolini, D., & Clementi, F. (2015). Bacteria and yeast microbiota in milk kefir grains from different Italian regions. *Food microbiology*, *49*, 123–133.

Guglielmetti, S., Taverniti, V., Minuzzo, M., Arioli, S., Stuknyte, M., Karp, M., & Mora, D. (2010). Oral bacteria as potential probiotics for the pharyngeal mucosa. *Applied and environmental microbiology*, *76*(12), 3948–3958.

Gulitz, A., Stadie, J., Wenning, M., Ehrmann, M. A., & Vogel, R. F. (2011). The microbial diversity of water kefir. *International journal of food microbiology*, *151*(3), 284–288.

Haahtela T. (2019). A biodiversity hypothesis. *Allergy*, *74*(8), 1445–1456.

Harvey, K. L., Jarocki, V. M., Charles, I. G., & Djordjevic, S. P. (2019). The Diverse Functional Roles of Elongation Factor Tu (EF-Tu) in Microbial Pathogenesis. *Frontiers in microbiology*, *10*, 2351.

Hertzler, S. R., & Clancy, S. M. (2003). Kefir improves lactose digestion and tolerance in adults with lactose maldigestion. *Journal of the American Dietetic Association*, *103*(5), 582–587.

Ismail, B., & Hayes, K. (2005). Beta-glycosidase activity toward different glycosidic forms of isoflavones. *Journal of agricultural and food chemistry*, *53*(12), 4918–4924.

Jayachandran, M., & Xu, B. (2019). An insight into the health benefits of fermented soy products. *Food chemistry*, *271*, 362–371. https://doi.org/10.1016/j.foodchem.2018.0

Jeong, D., Kim, D. H., Kang, I. B., Kim, H., Song, K. Y., Kim, H. S., & Seo, K. H. (2017). Modulation of gut microbiota and increase in fecal water

content in mice induced by administration of Lactobacillus kefiranofaciens DN1. *Food & function*, *8*(2), 680–686.

Korsak, N., Taminiau, B., Leclercq, M., Nezer, C., Crevecoeur, S., Ferauche, C., Detry, E., Delcenserie, V., & Daube, G. (2015). Short communication: Evaluation of the microbiota of kefir samples using metagenetic analysis targeting the 16S and 26S ribosomal DNA fragments. *Journal of dairy science*, *98*(6), 3684–3689.

Kotova, I. B., Cherdyntseva, T. A., & Netrusov, A. I. (2016). Russian Kefir Grains Microbial Composition and Its Changes during Production Process. *Advances in experimental medicine and biology*, *932*, 93–121.

Kurzer M. S. (2000). Hormonal effects of soy isoflavones: studies in premenopausal and postmenopausal women. *The Journal of nutrition*, *130*(3), 660S–1S.

Lampe J. W. (2009). Is equol the key to the efficacy of soy foods?. *The American journal of clinical nutrition*, *89*(5), 1664S–1667S.

Křížová, L., Dadáková, K., Kašparovská, J., & Kašparovský, T. (2019). Isoflavones. *Molecules (Basel, Switzerland)*, *24*(6), 1076.

Landete, J. M., Arqués, J., Medina, M., Gaya, P., de Las Rivas, B., & Muñoz, R. (2016). Bioactivation of Phytoestrogens: Intestinal Bacteria and Health. *Critical reviews in food science and nutrition*, *56*(11), 1826–1843.

Leite, A. M., Mayo, B., Rachid, C. T., Peixoto, R. S., Silva, J. T., Paschoalin, V. M., & Delgado, S. (2012). Assessment of the microbial diversity of Brazilian kefir grains by PCR-DGGE and pyrosequencing analysis. *Food microbiology*, *31*(2), 215–221.

Leskinen, P., Michelini, E., Picard, D., Karp, M., & Virta, M. (2005). Bioluminescent yeast assays for detecting estrogenic and androgenic activity in different matrices. *Chemosphere*, *61*(2), 259–266.

Marco, M. L., Heeney, D., Binda, S., Cifelli, C. J., Cotter, P. D., Foligné, B., Gänzle, M., Kort, R., Pasin, G., Pihlanto, A., Smid, E. J., & Hutkins, R. (2017). Health benefits of fermented foods: microbiota and beyond. *Current opinion in biotechnology*, *44*, 94–102.

Martin, P. M., Horwitz, K. B., Ryan, D. S., & McGuire, W. L. (1978). Phytoestrogen interaction with estrogen receptors in human breast cancer cells. *Endocrinology*, *103*(5), 1860–1867.

Matsuda., S., Norimoto., F., Matsumoto., Y., Ohba., R., Teramoto., Y., Ohta., N., & Ueda., S. (1994). Solubilization of a novel isoflavone glucoside-hydrolyzing β-glucosidase from *Lactobacillus casei* subsp. *rhamnosu*s. *Journal of fermentation and bioengineering*, *77*, 439–441.

McDonald, L. C., McFeeters, R. F., Daeschel, M. A., & Fleming, H. P. (1987). A differential medium for the enumeration of homofermentative and heterofermentative lactic Acid bacteria. *Applied and environmental microbiology*, *53*(6), 1382–1384.

Michlmayr, H., & Kneifel, W. (2014). β-Glucosidase activities of lactic acid bacteria: mechanisms, impact on fermented food and human health. *FEMS microbiology letters*, *352*(1), 1–10.

Middelbos, I. S., Fahey Jr., G. C. (2008) Soybean carbohydrates. *AOCS Press, 2008*, 269-296.

Merenstein, D. J., Foster, J., & D'Amico, F. (2009). A randomized clinical trial measuring the influence of kefir on antibiotic-associated diarrhea: the measuring the influence of Kefir (MILK) Study. *Archives of pediatrics & adolescent medicine*, *163*(8), 750–754.

Morito, K., Hirose, T., Kinjo, J., Hirakawa, T., Okawa, M., Nohara, T., Ogawa, S., Inoue, S., Muramatsu, M., & Masamune, Y. (2001). Interaction of phytoestrogens with estrogen receptors alpha and beta. *Biological & pharmaceutical bulletin*, *24*(4), 351–356.

Nanra, J. S., Buitrago, S. M., Crawford, S., Ng, J., Fink, P. S., Hawkins, J., Scully, I. L., McNeil, L. K., Aste-Amézaga, J. M., Cooper, D., Jansen, K. U., & Anderson, A. S. (2013). Capsular polysaccharides are an important immune evasion mechanism for *Staphylococcus aureus*. *Human vaccines & immunotherapeutics*, *9*(3), 480–487.

Pariza, M. W., Gillies, K. O., Kraak-Ripple, S. F., Leyer, G., & Smith, A. B. (2015). Determining the safety of microbial cultures for consumption by humans and animals. *Regulatory toxicology and pharmacology : RTP*, *73*(1), 164–171.

Plé, C., Breton, J., Richoux, R., Nurdin, M., Deutsch, S. M., Falentin, H., Hervé, C., Chuat, V., Lemée, R., Maguin, E., Jan, G., Van de Guchte, M., & Foligné, B. (2016). Combining selected immunomodulatory Propionibacterium freudenreichii and Lactobacillus delbrueckii strains: Reverse engineering development of an anti-inflammatory cheese. *Molecular nutrition & food research*, *60*(4), 935–948.

Potter, S. M., Baum, J. A., Teng, H., Stillman, R. J., Shay, N. F., & Erdman, J. W., Jr (1998). Soy protein and isoflavones: their effects on blood lipids and bone density in postmenopausal women. *The American journal of clinical nutrition*, *68*(6 Suppl), 1375S–1379S.

Prado, M. R., Blandón, L. M., Vandenberghe, L. P., Rodrigues, C., Castro, G. R., Thomaz-Soccol, V., & Soccol, C. R. (2015). Milk kefir: composition, microbial cultures, biological activities, and related products. *Frontiers in microbiology*, *6*, 1177.

Rackis, J. J. (1981) Flatulence caused by soya and its control through processing. *Journal of the American oil chemists' society, 58,* 503-509.

Rezac, S., Kok, C. R., Heermann, M., & Hutkins, R. (2018). Fermented Foods as a Dietary Source of Live Organisms. *Frontiers in microbiology*, *9*, 1785.

Rossiter, R. C., & Beck, A. B. (1966). Physiological and ecological studies on the oestrogenic isoflavones in subterranean clover (T. subterraneum L.). II. Effects of phosphate supply. *Australian journal of agricultural research, 17*(4), 447-456.

Sanlier, N., Gökcen, B. B., & Sezgin, A. C. (2019). Health benefits of fermented foods. *Critical reviews in food science and nutrition*, *59*(3), 506–527.

Setchell K. D. (1998). Phytoestrogens: the biochemistry, physiology, and implications for human health of soy isoflavones. *The American journal of clinical nutrition*, *68*(6 Suppl), 1333S–1346S.

Silva, K. R., Rodrigues, S. A., Filho, L. X., & Lima, A. S. (2009). Antimicrobial activity of broth fermented with kefir grains. *Applied biochemistry and biotechnology*, *152*(2), 316–325.

Suzuki, M. T., & Giovannoni, S. J. (1996). Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. *Applied and environmental microbiology*, *62*(2), 625–630.

Toomey, N., Bolton, D., & Fanning, S. (2010). Characterisation and transferability of antibiotic resistance genes from lactic acid bacteria isolated from Irish pork and beef abattoirs. *Research in microbiology*, *161*(2), 127–135.

Turan, İ., Dedeli, Ö., Bor, S., & İlter, T. (2014). Effects of a kefir supplement on symptoms, colonic transit, and bowel satisfaction score in patients with chronic constipation: a pilot study. *The Turkish journal of gastroenterology : the official journal of Turkish Society of Gastroenterology*, *25*(6), 650–656.

Välimaa, A. L., Kivistö, A. T., Leskinen, P. I., & Karp, M. T. (2010). A novel biosensor for the detection of zearalenone family mycotoxins in milk. *Journal of microbiological methods*, *80*(1), 44–48.

Wang, Y., Ahmed, Z., Feng, W., Li, C., & Song, S. (2008). Physicochemical properties of exopolysaccharide produced by Lactobacillus kefiranofaciens ZW3 isolated from Tibet kefir. *International journal of biological macromolecules*, *43*(3), 283–288.

Wang, H. J. & Murphy, P. A., (1996). Mass balance study of isoflavones during soybean processing. *Journal of agricultural and food chemistry, 44*(8), 2377-2383.

Wang, H. J. & Murphy, P. A., (1998). Isoflavone content in commercial soybean foods. *Journal of agricultural and food chemistry, 42*(8), 1666-1673.

Wen, Z. T., Suntharaligham, P., Cvitkovitch, D. G., & Burne, R. A. (2005). Trigger factor in Streptococcus mutans is involved in stress tolerance, competence development, and biofilm formation. *Infection and immunity*, *73*(1), 219–225.

Wick, R. R., Schultz, M. B., Zobel, J., & Holt, K. E. (2015). Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics (Oxford, England)*, *31*(20), 3350–3352.

Yang, J., Yoshida, Y., & Cisar, J. O. (2014). Genetic basis of coaggregation receptor polysaccharide biosynthesis in Streptococcus sanguinis and related species. *Molecular oral microbiology*, *29*(1), 24–31.

Yılmaz, İ., Dolar, M. E., & Özpınar, H. (2019). Effect of administering kefir on the changes in fecal microbiota and symptoms of inflammatory bowel disease: A randomized controlled trial. *The Turkish journal of gastroenterology: the official journal of Turkish Society of Gastroenterology*, *30*(3), 242–253.

Yoshida, Y., Palmer, R. J., Yang, J., Kolenbrander, P. E., & Cisar, J. O. (2006). Streptococcal receptor polysaccharides: recognition molecules for oral biofilm formation. *BMC oral health*, *6 Suppl 1*(Suppl 1), S12.

Yu, Z., Zhang, X., Li, S., Li, C., Li, D., & Yang, Z. (2013). Evaluation of probiotic properties of Lactobacillus plantarum strains isolated from Chinese sauerkraut. *World journal of microbiology & biotechnology*, *29*(3), 489–498.

Yuksekdag, Z., Cinar, B., Aslim, B., & Tukenmez, U. (2017). β-Glucosidase activity and bioconversion of isoflavone glycosides to aglycones by potential probiotic bacteria. *International journal of food properties, 20*(S3), S2878-S2886.

Zajšek, K., Kolar, M., & Goršek, A. (2011). Characterisation of the exopolysaccharide kefiran produced by lactic acid bacteria entrapped within natural kefir grains. *International journal of dairy technology, 64*, 544-548.

4.2.7 Authors contribution

Dalla Via, A. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Mantegazza, G. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Licata, A. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Gardana, C. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Gargari, G. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Alamprese, C. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.

Arioli, S. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Taverniti, V. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.
Guglielmetti, S. – Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy.

5. Conclusion and future perspective

In this PhD thesis we presented two works that investigates different activities that bacteria can perform on dietary compounds.

The first study covered the negative effects that the enzymatic activities of intestinal bacteria can have on certain food compounds. In particular, we focused on choline TMA-lyase microbiota enzymatic activity, that lead to production of trimethylamine oxide, a pro-atherosclerotic agent.
We have thereby developed a preliminary method capable of suggesting the existence of a relationship between the levels of TMAO excreted in urine, some taxonomic groups belonging to the most active bacterial families of the colonic microbiota, and a subgroup of the *cutC* gene ascribable to the choline-TMA conversion enzymes of Enterobacteriaceae, named cut-Kp, whose relative abundance can be determined with the qPCR protocol developed in this study. Nonetheless, considering the limitations listed above, particularly concerning dietary intake, it is plausible to hypothesize that the results of this study may vary in other populations.

In the second study we focused on another type of activity, linked to the bacterial modification of food compounds directly on food. During this work, four strains of lactic acid bacteria were isolated from a dairy product after adaption to soy drink. These strains were demonstrated to efficiently ferment the sugars in soy resulting in a creamy texture. In addition, the two strains belonging to the genus *Leuconostoc* were shown to remove completely the soy sugars stachyose and raffinose and increase the estrogenic activity of the product after fermentation. Considering that these bacteria belong to QPS species and have a safe antibiotic resistance profile, our study demonstrates that strains K05 and K09 are suitable for

the preparation of a new functional soy-based product, which can represent a source of isoflavones with enhanced estrogenic activity.

In conclusion, this PhD work has created the conditions for two lines of study based on an old and a new microbiological approach to modulate the action of the intestinal microbiota on our life.
The first study allowed the development of a protocol to profile a type of intestinal microbiota with harmful action. After subsequent trials that will take into account the limitations illustrated, this work can be included among the techniques able to profile all the inhabitants of the microbiota and to practice the so-called microbiota editing, a new frontier in this field of research (Madhusoodanan, 2020).
The second study has identified the starting points for the creation of a new fermented product. This product will be able to provide added value to the starting matrix, increasing its beneficial powers on human health. (Plé et al., 2016).

5.1. References

Madhusoodanan J. (2020). News Feature: Editing the microbiome. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(7), 3345–3348.

Plé, C., Breton, J., Richoux, R., Nurdin, M., Deutsch, S. M., Falentin, H., Hervé, C., Chuat, V., Lemée, R., Maguin, E., Jan, G., Van de Guchte, M., & Foligné, B. (2016). Combining selected immunomodulatory Propionibacterium freudenreichii and Lactobacillus delbrueckii strains: Reverse engineering development of an anti-inflammatory cheese. *Molecular nutrition & food research*, *60*(4), 935–948.

6. Appendices

6.1. Productivity

6.1.1. List of Pubblications

Taverniti, V., **Dalla Via, A.**, Minuzzo, M., Del Bo', C., Riso, P., Frøkiær, H., & Guglielmetti, S. (2017). In vitro assessment of the ability of probiotics, blueberry and food carbohydrates to prevent S. pyogenes adhesion on pharyngeal epithelium and modulate immune responses. *Food & function*, *8*(10), 3601–3609.

Taverniti, V., Koirala, R., **Dalla Via, A.**, Gargari, G., Leonardis, E., Arioli, S., & Guglielmetti, S. (2019). Effect of Cell Concentration on the Persistence in the Human Intestine of Four Probiotic Strains Administered through a Multispecies Formulation. *Nutrients*, *11*(2), 285.

Dalla Via, A., Gargari, G., Taverniti, V., Rondini, G., Velardi, I., Gambaro, V., Visconti, G. L., De Vitis, V., Gardana, C., Ragg, E., Pinto, A., Riso, P., & Guglielmetti, S. (2019). Urinary TMAO Levels Are Associated with the Taxonomic Composition of the Gut Microbiota and with the Choline TMA-Lyase Gene (*cutC*) Harbored by Enterobacteriaceae. *Nutrients*, *12*(1), 62.

6.1.2. Posters

Dalla Via, A., Taverniti, V., Gargari, G., Gambaro, V., Visconti, G. L., Pinto, A., Guglielmetti, S. (2018). The fecal abundance of bacterial choline utilization gene (cutC) is associated to specific fecal bacterial taxa and may predict urinary TMAO. Poster communication 6th World Congress on Targeting Microbiota, Porto, (P)

Dalla Via, A., Licata, A., Mantegazza, G., Gardana, C., Karp, M., Guglielmetti, S. (2019). Fermented soy drink with enhanced estrogenic activity. Microbiome and Probiotic series: Europe, Rotterdam, Rotterdam (NL)

nutrients

MDPI

Article

Urinary TMAO Levels Are Associated with the Taxonomic Composition of the Gut Microbiota and with the Choline TMA-Lyase Gene (*cutC*) Harbored by Enterobacteriaceae

Alessandro Dalla Via [1], Giorgio Gargari [1], Valentina Taverniti [1], Greta Rondini [1], Ilaria Velardi [1], Veniero Gambaro [2], Giacomo Luca Visconti [2], Valerio De Vitis [1], Claudio Gardana [1], Enzio Ragg [1], Andrea Pinto [1], Patrizia Riso [1] and Simone Guglielmetti [2,*]

[1] Department of Food, Environmental and Nutritional Sciences (DeFENS), University of Milan, 20122 Milan, Italy; alessandro.dallavia@unimi.it (A.D.V.); gargari.g@gmail.com (G.G.); valentina.taverniti@unimi.it (V.T.); gremary8687@gmail.com (G.R.); i.velardi@hotmail.com (I.V.); valeriodevitis88@gmail.com (V.D.V.); claudio.gardana@unimi.it (C.G.); enzio.ragg@unimi.it (E.R.); andrea.pinto@unimi.it (A.P.); patrizia.riso@unimi.it (P.R.)

[2] Department of Pharmaceutical Sciences, University of Milan, 20122 Milan, Italy; veniero.gambaro@unimi.it (V.G.); giacomo.visconti@unimi.it (G.L.V.)

* Correspondence: simone.guglielmetti@unimi.it

Received: 7 December 2019; Accepted: 19 December 2019; Published: 25 December 2019

Abstract: Gut microbiota metabolization of dietary choline may promote atherosclerosis through trimethylamine (TMA), which is rapidly absorbed and converted in the liver to proatherogenic trimethylamine-N-oxide (TMAO). The aim of this study was to verify whether TMAO urinary levels may be associated with the fecal relative abundance of specific bacterial taxa and the bacterial choline TMA-lyase gene *cutC*. The analysis of sequences available in GenBank grouped the *cutC* gene into two main clusters, cut-Dd and cut-Kp. A quantitative real-time polymerase chain reaction (qPCR) protocol was developed to quantify *cutC* and was used with DNA isolated from three fecal samples collected weekly over the course of three consecutive weeks from 16 healthy adults. The same DNA was used for 16S rRNA gene profiling. Concomitantly, urine was used to quantify TMAO by ultra-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS). All samples were positive for *cutC* and TMAO. Correlation analysis showed that the cut-Kp gene cluster was significantly associated with *Enterobacteriaceae*. Linear mixed models revealed that urinary TMAO levels may be predicted by fecal cut-Kp and by 23 operational taxonomic units (OTUs). Most of the OTUs significantly associated with TMAO were also significantly associated with cut-Kp, confirming the possible relationship between these two factors. In conclusion, this preliminary method-development study suggests the existence of a relationship between TMAO excreted in urine, specific fecal bacterial OTUs, and a *cutC* subgroup ascribable to the choline-TMA conversion enzymes of *Enterobacteriaceae*.

Keywords: choline; trimethylamine; trimethylamine n-oxide; 16S rRNA gene profiling; qPCR; linear mixed models

1. Introduction

From infancy, the microorganisms colonizing the human gastrointestinal tract (GIT), collectively known as GIT microbiota, act as a "hidden" metabolic organ that exerts indispensable functions for the development and physiology of the human organism, such as the production of vitamins,

modulation of the immune system, competitive exclusion toward exogenous pathogenic bacteria, xenobiotic detoxification, and production of short-chain fatty acids [1]. Nonetheless, detrimental activities have also been associated with gut commensal microorganisms, such as the production of carcinogens by the bacterial nitroreductases and azoreductases [2], or the conversion of primary bile acids to toxic compounds by the microbiota-associated enzyme cholesterol dehydrogenase and 7-α-dehydroxylase [3]. In addition, it was proposed that the intestinal bacterial enzymatic activities that produce trimethylamine (TMA) may promote atherosclerosis. TMA, in fact, is readily absorbed from the intestinal tract and, once in the liver, is converted into trimethylamine-N-oxide (TMAO) [4], whose plasma level has been identified as a metabolite strongly associated with atherosclerosis in a large case-control cohort for cardiovascular disease [5]. In particular, TMAO was proposed to promote atherogenesis by increasing cholesterol in macrophages and enhancing the accumulation of foam cells in artery walls [4,5]. Nonetheless, the literature has contradicted the role of TMAO, and recent studies have questioned its deleterious role in the cardiovascular system [6], suggesting, on the contrary, that TMAO could have protective functions [7,8].

Reportedly, a dominant contribution to the production of TMA in the gut comes from the microbial metabolism of diet-derived substrates such as carnitine- and choline-containing molecules [4,5,9]. Choline is an essential nutrient that is used by cells to synthesize membrane phospholipids. Furthermore, choline is the precursor of the neurotransmitter acetylcholine and a major source for methyl groups via its metabolite, trimethylglycine (betaine) [10]. The main dietary sources of the choline moiety, which is mostly present in food as lecithin (i.e., phosphatidylcholine), were reported to be eggs, liver, soybeans, and pork [11]. Although they are also present in numerous other foods [12], recent surveys in the USA indicated that choline may be underconsumed in specific populations (e.g., pregnant women and vegans) [13]. Based on the average observed choline intake in healthy European populations, a panel of the European Food Safety Authority set the adequate intake of choline at 400 mg/day [14].

Recent literature has suggested that the enhanced abundance of choline utilization genes in the intestinal microbiome is associated with increased TMA levels in the gut and, subsequently, with a higher hepatic production of TMAO. Proof of the importance of choline-derived TMA in the context of TMAO toxicity was recently provided by the study of Craciun and Balskus, in which the specific inhibition in mouse intestine of the microbial choline TMA-lyase (the primary enzymatic activity involved in the production of TMA from choline [15]) resulted in a significant reduction in plasma TMAO levels and recovery from dietary-induced platelet aggregation and thrombus formation [16].

Choline TMA-lyase is discontinuously distributed in bacterial taxa. Consequently, it was speculated that the phylogenetic composition of the microbiota is plausibly a poor predictor of the intestinal potential to convert choline into TMA [15,17,18]. However, in another study, the taxonomic structure of the gut microbiota was used to predict genes involved in choline metabolism [19] by means of PICRUSt, a bioinformatic tool used to infer the functional profiles of the microbial communities from 16S rRNA gene profiling data [20]. Although the toxicity of TMAO has been extensively investigated in the last 10 years, the association potentially existing among host TMAO levels, gut microbiota composition, and the intestinal microbial metabolization of choline has been only marginally considered. In this context, we developed a molecular protocol for the targeted quantification in the fecal microbiome of the bacterial gene *cutC* coding for the glycyl radical enzyme homolog choline TMA-lyase [15,21]. This protocol was applied to quantify the *cutC* gene abundance in the fecal samples collected at different time points from a group of healthy adults. Then, the obtained results were analyzed in comparison with the bacterial taxonomic composition and the urinary levels of TMAO concomitantly determined in the same population to deduce the potential association of excreted TMAO with gut microbial taxa and/or specific choline TMA-lyase enzymes.

2. Materials and Methods

2.1. Design and Use of Primers Targeting the cutC *Gene*

The primers used in polymerase chain reaction (PCR) for the amplification of the *cutC* gene were designed as follows. The GenBank database and Conserved Domain Database (CDD) at the National Center for Biotechnology Information (NCBI) were queried to select 52 nonredundant representative bacterial proteins of the choline trimethylamine-lyase protein family TIGR04394 (choline_CutC; EC Number 4.3.99.4), including the CutC enzymes of *Desulfovibrio desulfuricans* [4], and *Klebsiella pneumoniae* [22]. Then, the corresponding CDS nucleotide sequences of selected proteins were used to build a UPMGA tree upon ClustalW multiple alignments. According to the obtained dendrogram, sequences were clustered in two groups: One including the *cutC* sequence of *K. pneumoniae*, named cut-Kp, and one including the *cutC* sequence of *D. desulfuricans*, named cut-Dd (Supplementary Figure S1). Finally, a pair of primers was designed in the most conserved regions of each group of sequences: cut-Dd-F, 5′-CGTGTTGACCAGTACATGTA-3′ and cut-Dd-R 5′-GCTGGTAACCTGCGAAGAA-3′ (expected amplicon of 185 bp); cut-Kp-F, 5′-GATCTGACCTATCTGATTATGG-3′, and cut-Kp-R, 5′-TTGTGGAGCATCATCTTGAT-3′ (expected amplicon of 190 bp).

2.2. PCR Detection of cutC *Gene in Single Strains*

The two primer pairs designed as described above were used in endpoint PCR with the genomic DNA extracted from 64 bacterial strains (Table S1). Reaction mix was prepared in 25 µL, including 0.5 units of DreamTaq Polymerase (ThermoFisher, Fermentas, Waltham, MA, USA), 1× concentration of DreamTaq Polymerase Buffer (ThermoFisher, Fermentas,), 0.25 µM of each primer, 200 µM of deoxyribonucleotide triphosphate (dNTPs), and 0.5 mM of MgCl2. The PCR cycle program used was the following: Initial denaturation at 95 °C for 2 min, followed by 35 cycles of denaturation at 94 °C for 45 s, annealing at 58 °C for 45 s for the cut-Dd couple and 56 °C for 45 s for the cut-Kp couple, and extension at 72 °C for 20 s. A final extension of 7 min at 72 °C was then applied.

2.3. Detection of Choline-Utilization Activity in Single Strains

Bacterial strains were grown in the respective culture medium (reported in Table S1) for 48 h. Afterward, the biomasses were collected by centrifugation at 9500 g for 10 min. The cell pellets were then washed with sterile PBS and resuspended in fresh medium with the addition of 0.2% filter-sterilized choline. Bacteria were incubated at 37 °C for 48 h in glass tubes with screw cap. Afterward, supernatants were collected and used for mass spectrometry (MS) and nuclear magnetic resonance (NMR) analyses. The MS analyzes were performed by directly injecting 5 µL of diluted broth cultures after the removal of the bacterial cells by centrifugation and subsequent filtration with a 0.45-µm syringe filter. In detail, the broth cultures were analyzed in full scan in the range from 50 u to 400 u on an HR-MS Orbitrap model Exactive with a HESI-II probe for electrospray ionization (Thermo Scientific, San Jose, CA, USA). The resolution, gain control, mass tolerance, and maximum ion injection time was set to 50 K, 1E6, 2 ppm, and 100 ms, respectively. The MS data were processed using Xcalibur software (Thermo Scientific). Choline and TMA were used as reference standard. Choline and TMA were also directly detected in broth cultures by ^{1}H-NMR with a 60 MHz benchtop NMR spectrometer Spinsolve 60 Carbon Ultra, Magritek GmbH (Aachen, Germany).

2.4. Study Population

Study participants were recruited within the University campus. In total, four females and 12 males aged 21–45 (mean: 29.8 years) were enrolled (Table S2). The inclusion criteria were as follows: Healthy adult volunteers of both sexes who provided signed informed consent of their participation in the study. The exclusion criteria were as follows: Antibiotic consumption in the month preceding the start of the study, consumption of antacids or prokinetic gastrointestinal drugs, episodes of viral or bacterial enteritis in the two months prior to the study, episodes of gastric or duodenal ulcers in the

previous five years, pregnancy or breastfeeding, recent history of alcohol abuse or suspected drug use, and any severe disease that may interfere with treatment. Ethical permission was granted by the University of Milan Ethics Committee (ref: opinion no. 37/16, 15 December 2016).

2.5. Collection of Fecal and Urine Samples

Three fecal sample were collected weekly over the course of three consecutive weeks from each volunteer. All the participants were asked to follow their regular diet during the three weeks. Concomitantly to the fecal sample, the volunteers provided 24-h urine collection.

Urine samples were collected over 24 h in sterile tanks and on the same days that fecal samples were been collected. The volume of collected urine was recorded in order to calculate the daily excretion of trimethylamine oxide (TMAO). Immediately after delivery, part of the urine samples was transferred in 10-mL sterile tubes and stored at −80 °C until analysis.

2.6. Analysis of cutC *Gene by Quantitative Real-Time PCR*

The cutC gene was quantified in fecal DNA with quantitative real-time PCR (qPCR) with both primer pairs, cut-Dd and cut-Kp. To this aim, DNA was extracted from feces using the kit PowerLyzer® PowerFecal® DNA Isolation Kit (MO BIO Laboratories, Inc.), starting from 0.25 ± 0.02 mg of sample according to the manufacturer's instructions. Primer pairs were tested with a gradient qPCR in a range of eight temperatures in order to find the most efficient annealing temperature using DNA of Streptococcus dysgalactiae 485 and Klebsiella sp. A1.2 as reference DNA. In addition, the amplification efficiency of the two pairs of primers was tested in qPCR experiments with six serial 1:3 dilutions of genomic DNA isolated from *Streptococcus dysgalactiae* 485, *Klebsiella* sp. A1.2, and human fecal metagenomic DNA. All DNA (bacterial and metagenomic) serial dilutions were tested with primer concentrations of 0.5 μM, 0.4 μM, and 0.3 μM. Efficiency curves were obtained with Bio-Rad software by setting samples as "standard" and obtaining a curve with efficiency (E) parameter and R2 value. Based on the results of these setup experiments, primers were then used at a final concentration of 0.5 μM, as with this concentration, we obtained an R2 value of 0.98. In addition, two randomly selected fecal DNA samples were tested at the different concentration by adding 70 ng, 50 ng, 25 ng, and 10 ng in qPCR reactions. Based on Ct value comparison between the different DNA concentrations, the *cutC* gene quantification was subsequently performed using 50 ng of total DNA. The reaction mix contained the SsoFast TM Eva-SuperGreen Supermix 2× (Bio-Rad Laboratories), deionized Milli-Q water (Millipore), and primers. All DNA samples (5 μL in each well) were tested in technical duplicate. The qPCR cycles employed were the following: Initial denaturation at 95 °C for 3 min, followed by 44 cycles of denaturation at 95 °C for 30 s, annealing at 58 °C (for cut-Dd primers) or 58.5 °C (for cut-Kp primers) for 30 s, and elongation at 72 °C for 5 s. A final denaturation ramp between 65 °C and 95 °C for 5 s was performed for the melting curve analysis. Moreover, specificity of qPCR reaction was confirmed by checking the presence of only one amplification and of the expected size in electrophoresis on a 2% agarose gel. A total of 48 fecal samples were analyzed. Each sample was analyzed with each primer set in duplicate. The $2^{-\Delta\Delta Ct}$ method was used for the relative quantification of *cutC* gene, using the EUB panbacterial primers [Muyzer] targeting the 16S rRNA gene as reference. Data were reported as relative increase of *cutC* copy number compared to the level of the sample that showed the highest significant Ct in qPCR set as 1.

2.7. Analysis of the Bacterial Taxonomic Composition of Fecal Samples

The bacterial community structure of the fecal microbiota was analyzed as described elsewhere [23,24], with DNA extracted from feces as described in Section 2.2. In brief, extracted DNA was analyzed through 16S rRNA gene profiling. Sequencing reads were generated at the Institute for Genome Sciences (University of Maryland, School of Medicine, Baltimore, MD, USA) with Illumina HiSeq 2500 rapid run sequencing of the V3–V4 variable region. Sequencing reads were equally distributed among the samples. Sequences were filtered and trimmed based on their quality.

We obtained a sequence length of 301 bp for both R1 and R2 sequences with an average quality score (Phred score) higher than 35. Sequencing reads were rarefied at 5000 per sample. Subsequently, sequence reads were analyzed through the bioinformatic pipeline Quantitative Insights into Microbial Ecology (QIIME) version 1.9.1 [25] with the GreenGenes database updated to version 13.5. The relative abundance of bacteria in each fecal sample was reported at the taxonomic levels of phylum, class, order, family, genus, and operational taxonomic units (OTUs). Sequence were deposited in the European Nucleotide Archive (ENA) of the European Bioinformatics Institute under accession code PRJEB34169.

2.8. TMAO Quantification in Urine Samples

TMAO levels in urine samples were determined by ultra-performance liquid chromatography coupled to tandem mass spectrometry (UPLC-MS/MS) (Waters Acquity UPLC system). The analysis method involved the use of a totally porous column with stationary C8 stable bond (Agilent Poroshell C8-SB) and a mobile phase consisting of a gradient acetonitrile and formate buffer (3 mM of ammonium formate and 0.1% formic acid). The UPLC system was equipped with a triple quadrupole detector, which allowed the development of a "multiple reaction monitoring" (MRM) method for the analysis of TMAO. In detail, once thawed at room temperature and after centrifugation at 6000 rpm for 5 min, 25 µL of urine sample were diluted in 950 µL of UPLC mobile phase (1/1 (*v*/*v*) acetonitrile/ultra-pure sterile water + 0.025% of formic acid), and 25 µL of deuterated internal standard solution (1 ppm, TMAO-d9, Spectra 2000) were used for the normalization of results [26]. The UPLC samples were prepared mixing 950 µL of mobile phase [1/1 (*v*/*v*) acetonitrile/ultra-pure sterile water + 0.025% formic acid), 25 µL of urine sample, and 25 µL of deuterium-labeled methyl d9-TMAO solution (1 ppm; Spectra 2000 S.r.l., Roma, Italy). Mobile phase: 1/1 (*v*/*v*) acetonitrile/ultra-pure sterile water + 0.025% of formic acid. The run time per sample was 8 min. Sample freezing and thawing or their prolonged storage at room temperature did not have an impact on the TMAO quantification. A triple set of working standards of TMAO (trimethylamine N-oxide dihydrate, Fluka) at concentrations of 5 ppm, 50 ppm, 100 ppm was prepared according to the method described above, replacing the 25 µL of urine sample with 25 µL of standard solution. The average response factor was used for calculation.

2.9. Statistical Analysis

Statistical analyses of data were carried out using R statistic software (version 3.4.2). Concerning *cutC* gene and TMAO data, intrasubject variability was defined "high" when variance among the three replicates results were higher than twice the median of all variances. Correlation analyses were performed using the Kendall and Spearman formula with the items specified in the text as predictors and dependent variables. Significance was set at $p \leq 0.05$, and mean differences in the range $0.05 < p < 0.10$ were accepted as trends. To find associations among TMAO levels, bacterial taxa relative abundance, and *cutC* gene abundance, the machine learning supervised linear mixed model (LMM) algorithm was used. In brief, the LMM was performed using "lmer" function in the "lme4" library [27]. All samples were used in the LMM analysis (n = 48), considering that three measurements were available for each subject. The Akaike's Information Criterion (AIC) was used to test the goodness of fit of the LMM. The AIC index/value depends on the ANOVA test results between two models: The model that considered the effect of the predictors and the null model.

3. Results

3.1. Distribution of the cutC *Gene among Bacterial Taxa*

According to the literature, the ability of intestinal bacteria to convert the choline moiety to TMA is primarily associated with a recently discovered choline utilization (cut) genetic region harboring the *cutC* gene, which encodes a glycyl radical enzyme catalyzing C–N bond cleavage [15,18]. For this reason, we designed primers specifically targeting the *cutC* gene. These primers were intended for quantitative PCR (qPCR) experiments, and we avoided the use of degenerations in their sequence.

In contrast, to target all putative *cutC* sequences identified in GenBank, we clustered the putative *cutC* genes into two groups (named Dd and Kp) according to sequence similarity (Supplementary Figure S1) and designed a pair of primers for each group in the most conserved sequence regions. Group Dd included putative *cutC* genes from *Firmicutes* (*Anaerococcus*, *Clostridium*, *Enterococcus*, *Streptococcus*), *Proteobacteria* (*Desulfotalea*, *Desulfovibrio*, *Enterobacter*), and *Actinobacteria* (*Olsenella*). Group Kp comprised putative *cutC* gene sequences from *Proteobacteria* (*Aeromonas*, *Enterobacter*, *Erwinia*, *Escherichia*, *Klebsiella*, *Pectobacterium*, *Pelobacter*, *Proteus*, *Providencia*, *Raoultella*, *Serratia*) and *Firmicutes* (*Desulfosporosinus*, *Enterococcus*).

Subsequently, the two primer sets were used in endpoint PCR reactions to test the presence of putative *cutC* genes within the genomic DNA isolated from the pure cultures of 64 bacterial strains. We obtained an amplicon of the expected size from seven strains. Specifically, strains *Streptococcus dysgalactiae* 485, *S. dysgalactiae* 486, and *S. dysgalactiae* A1.3 gave a band of the expected size with primers cut-Dd. In addition, strains *Enterococcus gilvus* MD179, *Enterococcus hirae* MD160, *Klebsiella oxytoca* MIMgr, and *Klebsiella* sp. MIMgr were positive with primers cut-Kp (Figure 1A,B). MS and NMR analyses revealed the ability to metabolize choline and produce TMA only for the same seven strains that resulted in positive PCR experiments (Figure 1C and Supplementary Figure S2).

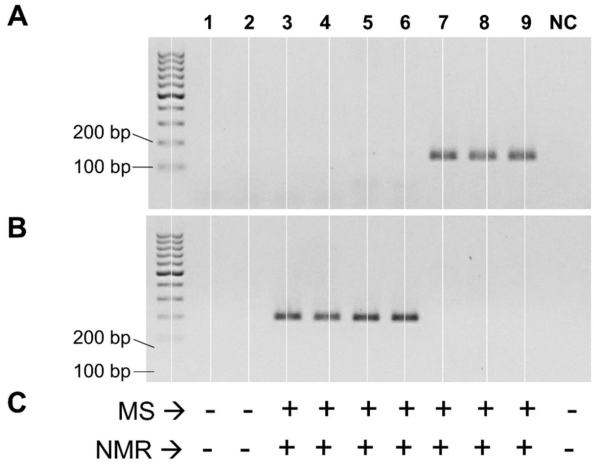


Figure 1. Detection of the choline-utilization activity in pure bacterial cultures. Panels (**A**,**B**) represent agarose gel resulting from end-point PCR with primers cut-Dd (**A**) and cut-Kp (**B**). Panel (**C**) summarizes the detection of TMA in cell-free broth by mass spectrometry (MS) and nuclear magnetic resonance (NMR); +, TMA detected; -, TMA not detected. Lanes: **1**, *Escherichia coli* 3.1; **2**, *Lactococcus garvieae* FMBgr; **3**, *Enterococcus gilvus* MD160; **4**, *Enterococcus hirae* MD179; **5**, *Klebsiella oxytoca* MIMgr; **6**, *Klebsiella* sp. A1.2; **7**, *Streptococcus dysgalactiae* 485; **8**, *Streptococcus dysgalactiae* 486; **9**, *Streptococcus dysgalactiae* A 1.2; NC, negative control (i.e., M17 broth incubated without bacteria).

3.2. Bacterial Taxonomic Structure of the Fecal Microbiota

The metagenomic DNA isolated from the feces collected at three time points from 16 healthy adults (n = 48) was used in 16S rRNA gene profiling experiments. A total of 12,588,795 filtered high-quality sequence reads were generated with an average of 13,340 ± 8677 (mean ± standard deviation; max-min 11,594–4570) per sample.

We failed to stratify samples according to the 16S rRNA gene profiling data, indicating that fecal bacterial community structure was homogeneous among samples and among subjects (Supplementary Figure S3). In addition, we also observed that the overall composition of the fecal microbiota in each subject remained mostly stable over the three collection time points (Supplementary Figure S3). Globally, 182 bacterial genera were estimated, with a minimum of 36 and a maximum of 98 genera per fecal sample. *Bacteroides* was the most prevalent genus, followed by four genera of the order *Clostridiales* (undefined *Ruminococcaceae*, undefined *Lachnospiraceae*, *Ruminococcus*, and *Faecalibacterium*)

(Supplementary Figure S4A). At the family level, most of the reads were ascribed to only three families, i.e., *Ruminococcaceae*, *Bacteroidaceae*, and *Lachnospiraceae* (Supplementary Figure S4B).

3.3. *Putative* cutC *Genes in Human Fecal Metagenomic DNA*

In order to investigate the presence of *cutC* genes in the human gut microbiome, the cut-Dd and cut-Kp primer sets were used in qPCR experiments using the same fecal metagenomic DNA as a template from healthy adults used for microbiota profiling. All analyzed fecal samples gave a positive signal in qPCR with both primer pairs (Figure 2). In general, cut-Kp was detected at a higher relative concentration than cut-Dd (median ΔΔCt of 5.33 and 0.85 for cut-Kp and cut-Dd, respectively) (Figure 2A,B). In addition, with both cut-Kp and cut-Dd, six volunteers out of 16 showed a variance among the three replicates that was higher than twice the median of all variances, indicating a higher intrasubject variability (Figure 2A,B).

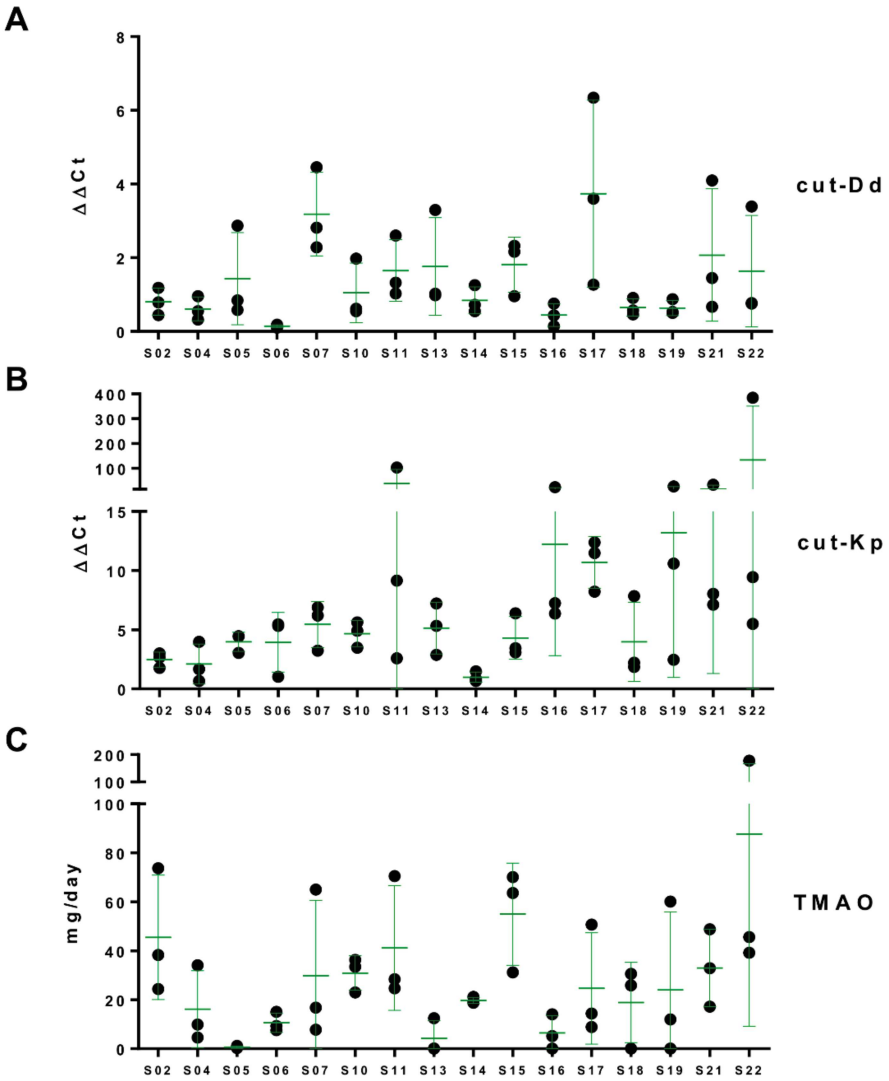


Figure 2. Fecal levels of the *cutC* gene and daily urinary excretion of trimethylamine-N-oxide (TMAO). The relative abundance of *cutC* was determined by quantitative real-time polymerase chain reaction (qPCR) with the primer pair cut-Dd-F/R (panel **A**) and cut-Kp-F/R (**B**). The TMAO concentration was determined by ultra-performance liquid chromatography coupled with tandem mass spectrometry (UPLC-MS/MS) in urine collected over 24 h (**C**). Green bars represent the mean ± standard deviation of three measurements per subject.

Subsequently, we performed correlation analyses between the *cutC* abundances determined with qPCR and the 16S rRNA gene profiling data to find potential relationships between the choline TMA-lyase genes and specific bacterial taxa of the fecal microbiota. To this end, we used the median relative abundance of bacterial taxa in fecal samples as predictors, whereas the dependent variables considered were the median abundances of cut-Dd and cut-Kp determined by qPCR per subject. We found that cut-Dd was positively correlated with taxa belonging to the phylum *Firmicutes*, including an undefined *Mogibacteriaceae* genus, *Oscillospira*, and the family *Christensenellaceae*. On the contrary, cut-Dd was negatively correlated with the *Firmicutes* order *Bacillales*, the *Firmicutes* genus *Streptococcus*, and the *Proteobacteria* genus *Haemophilus* (Supplementary Figure S5). Conversely, cut-Kp was positively associated with *Proteobacteria*. In particular, inside this phylum, a significant correlation was found with the family *Enterobacteriaceae* (Supplementary Figure S5).

3.4. Daily Urinary Excretion of TMAO

Subjects were asked to collect 24-h urine specimens the same days when the fecal samples were taken. Then, the levels of TMAO were quantified by UPLC-MS in all urine samples, revealing wide variability among the investigated healthy adults, with levels of urinary TMAO excretion ranging from less than 1 mg to more than 175 mg per day (Figure 2). We also observed an evident intrasubject variability in five volunteers whose TMAO excretion showed a variance among the three replicates that was higher than twice the median of all variances (Figure 2C). In particular, four out of the five volunteers with wide intrasubject variability (i.e., S07, S11, S19, and S22) were found to possess high intrasubject variability for *cutC* gene levels determined in qPCR experiments (Figure 2).

3.5. Associations among Urinary TMAO, Fecal cutC*, and Fecal Bacterial Taxa*

A linear mixed model was used to infer potential significant relationships among the datasets collected from volunteers at the three time points considered (Figure 3). TMAO was significantly associated with the cut-Kp/cut-Dd synergy ($p < 0.001$). Furthermore, studying the association of the single *cut* gene types, we observed that the relationship with TMAO was mainly determined by cut-Kp (Figure 3). In addition, we found a significant association between TMAO and 23 operational taxonomic units (OTUs). Conversely, cut-Kp and cut-Dd were significantly associated with 18 and eight OTUs, respectively. Notably, most of the OTUs that were significantly associated with cut-Kp (i.e., 15 out of 18) were also associated with TMAO, confirming the relationship between these two variables. Nine of the identified OTUs belonged to the phylum *Bacteroidetes*, while the remaining 21 were ascribed to *Firmicutes*. In addition, 80% of the OTUs (n = 24) belonged to only three families: *Bacteroidaceae*, *Lachnospiraceae*, and *Ruminococcaceae*. In particular, the most significant association (i.e., $p < 0.001$) referred to *Bacteroides caccae*, an undefined *Lachnospiraceae* genus, and several undefined *Ruminoccaceae* species (for TMAO and cut-Kp), *Bacteroides fragilis*, and an undefined *Clostridiales* species (for cut-Kp only) and an *Oscillospira* species (for cut-Dd) (Figure 3).

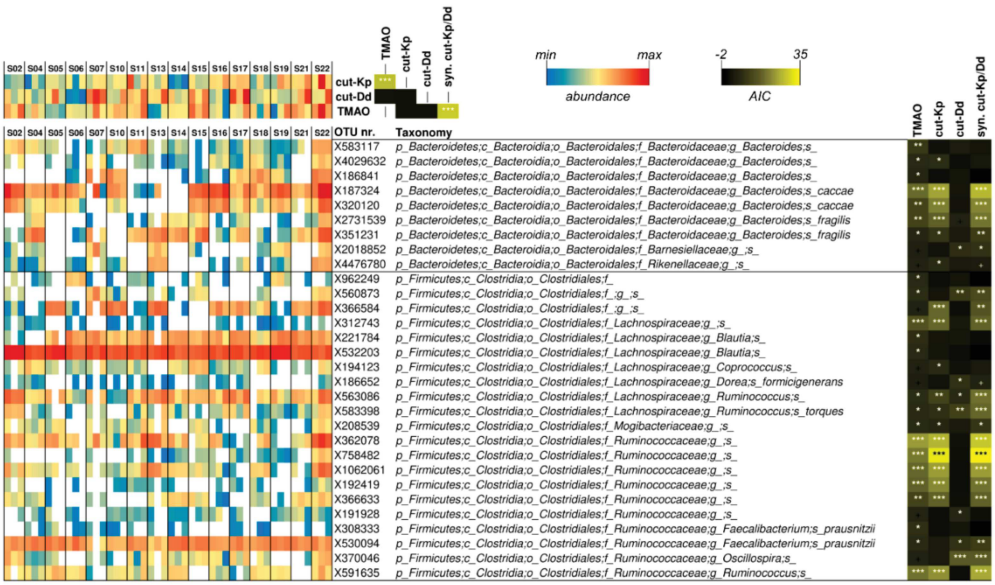


Figure 3. Analysis of the associations among fecal *cutC* gene abundances, fecal bacterial operational taxonomic units (OTUs), and urinary excreted TMAO carried out through a linear mixed model (LMM). Only OTUs that showed a significant association with *cutC* or TMAO are reported. The heatmap on the right represents TMAO levels, and *cutC* gene and OTU relative abundances. White boxes in the blue-yellow-red heatmap indicate that the OTU was not detected in that specific sample. The taxonomic lineage of each taxon is shown: p, phylum; c, class; o, order; f, family; g, genus; s, species. The black-yellow heatmap represents the Akaike's information criterion (AIC) values of the LMM analysis. Asterisks indicate significant associations: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; $^{+}$, $p < 0.1$. syn. cut-Kp/Dd = synergy between cut-Kp and cut-Dd in LMM analysis.

4. Discussion

A growing number of studies have linked host TMAO levels to different diseases or prepathological metabolic states [28,29]. Conversely, TMAO has also been proposed as a beneficial factor that may promote protein stabilization and protect cells from osmotic and hydrostatic stresses according to a compensatory response mechanism [30]. The biological role of TMAO is therefore still debated. Nonetheless, a growing number of scientific studies have suggested that this molecule may play an important role in health and diseases [6].

It has been suggested that an important contribution to the hepatic production of TMAO is given by the TMA produced in the gut by microbial degradation of TMA-containing dietary molecules [31]. In particular, TMAO levels and their physiological consequences were shown to be significantly affected by the TMA derived from choline [15]. In this context, we studied the levels of TMAO excreted daily with urine, the composition of the intestinal microbiota, and the abundance of the choline TMA-lyase gene *cutC* in a group of healthy adult subjects with an Italian dietary pattern. The aim of this observational study was to verify whether TMAO levels excreted with the urine might be associated with the relative abundance of specific bacterial taxa and the bacterial gene *cutC* in feces. Literature focusing on the relationship among these three elements, particularly in non-diseased populations, is limited and partially contradicting [15,17–19].

The gene *cutC*, encoding the lyase enzyme essential for the conversion of choline into TMA [32], is not evenly distributed across bacterial taxa due to gene loss and horizontal gene transfer events that differently involve strains within the same species [15,18,33]. Therefore, predicting the choline degradation potential of a microbial ecosystem solely based on the taxonomic composition has many intrinsic limitations. The use of primers selectively targeting a specific enzymatic conserved domain may overcome this problem, permitting the selective quantification of the abundance of a gene coding for a specific enzymatic activity in the metagenomic DNA. A similar approach was used

by Martinez-Del Campo et al., who designed degenerate primers for the PCR amplification of the *cutC* gene from fecal metagenomic DNA and single strains [18]. The use of degenerate primers was necessitated by the fact that the CutC protein possesses sequence heterogeneity. In particular, Martinez-Del Campo et al. showed that the amino acid sequences deduced from the predicted bacterial *cutC* genes can be clustered into two groups (clades 1 and 2, [18]), which correspond to the CutC types I and II identified by Jameson et al. within a neighbor-joining phylogenetic tree constructed from amino acid sequences of glycyl radical enzymes [32]. The same result was found in our study by generating a distance tree based on the nucleotide sequences of putative *cutC* genes (Supplementary Figure S1). In particular, cluster cut-Dd corresponded to clade 1 and CutC type I, whereas cut-Kp included sequences coding for putative proteins found in clade 2 and CutC type II reported by the authors of [18] and [32], respectively.

For this reason, we developed two nondegenerate primer pairs located at the level of the catalytic site of the encoded enzyme that were useful for the amplification in (q)PCR experiments of the two clusters of the gene *cutC*.

When the two primer sets were used with the DNA of single strains, the only positive amplification signals were obtained with the bacteria that demonstrated the ability to metabolize choline in the biotransformation assay and produce TMA, confirming the suitability of these molecular probes to target choline-TMA-converting bacteria. Specifically, the bacterial strains identified here as able to degrade choline to TMA include species previously confirmed to exert this conversion, such as *Streptococcus dysgalactiae* [18]. In addition, we found *Klebsiella oxytoca*, which was reported to harbor a putative cut gene cluster [34], but has never been confirmed phenotypically. We also identified two positive *Enterococcus* strains. Reportedly, TMA production from choline has also been described for some enterococci, but not for the species *E. gilvus*, which is often isolated from food matrices, including meat, milk, and cheeses [35,36], and for the zoonotic pathogen *E. hirae* [37].

The qPCR experiments conducted showed that putative bacterial *cutC* genes were present in the fecal samples of all healthy adult subjects investigated. The high prevalence of this bacterial gene in the human gut microbiome was reported in a previous study, in which the presence of *cutC* homologs was observed in 96.6% of the assembled stool metagenomes of healthy individuals from the Human Microbiome Project (HMP) [18].

Reportedly, most of the TMA produced in the gut is absorbed into the portal circulation by passive diffusion [38]. Then, approximately 95% of the absorbed TMA is oxidized in the liver by flavin monooxygenases and excreted in the urine within 24 h [31,39]. Therefore, in this study, we performed a quantification of TMAO levels in urine samples obtained by 24-h collection.

The data presented here revealed a marked variability of both *cutC* and TMAO levels over the three time points considered in approximately 40% of volunteers. This instability was plausibly due to the variability of the daily food consumption of each subject. In this study, volunteers were free to follow their usual diet. Therefore, the analysis of multiple time points at approximately one-week intervals was useful to address the observed temporal instability of these parameters. To the best of our knowledge, this is the first work to report the stability of intestinal *cutC* and urinary TMAO levels over time.

This study has several limitations:

1. First, we quantified the abundance of a gene of the intestinal microbiome without considering if and how much this gene was expressed. This could therefore limit the possibility of associating the abundance of this gene with its product.
2. Furthermore, the production of TMA, in addition to the presence of the bacterial gene that allows its production (*cutC*), depends on the availability of the choline substrate, which mainly comes from the diet.
3. Nonetheless, the contribution to the TMA produced in the intestine and, consequently, to the TMAO generated in the liver, derives from different chemical moieties (mainly choline, betaine, and carnitine) and includes different microbial metabolic pathways, such as those involving the

carnitine monooxygenase CntAB and the glycine betaine reductase GrdH, in addition to the choline TMA-lyase CutC [40].

4. In addition, TMAO urinary levels may also depend on host factors that may largely vary from subject to subject, such as (i) the gut-to-blood barrier permeability to TMA [41], (ii) the oxidation of TMA in the liver by flavin monooxygenase [5], and the kidney function [42].
5. Finally, TMAO can also be ingested directly from foods such as fish and seafood, which are naturally rich in this molecule [43].

However, despite the limitations described above, this study showed that changes in urine TMAO levels are associated with changes in the fecal abundance of the *cutC* gene and variations in the relative abundance of several bacterial taxonomic units of the fecal microbiota. In particular, TMAO was significantly associated with the levels of a specific subcategory of the *cutC* gene, which we named cut-Kp here. This result could be explained by the relative abundance of cut-Kp, which, by qPCR results, was approximately six-times higher than that of cut-Dd. According to correlation analysis, the most important contribution to cut-Kp gene abundance is provided by *Proteobacteria*, particularly by *Enterobacteriaceae*. This result is supported by the fact that cut-Kp has been quantified with primers designed on a cluster of gene sequences having the *cutC* of the *Enterobacteriaceae* species *K. pneumoniae* as a reference. Reportedly, the analysis of human gut metagenomes revealed a high proportion of the genera *Klebsiella* and *Escherichia*, which harbor three potential TMA-producing pathways, suggesting the importance of these bacteria for TMA cycling in the human gut [44].

Most OTUs that were found to be significantly associated with TMAO also had cut-Kp, confirming the relationship between TMAO and cut-Kp levels. A few OTUs were also associated with cut-Dd. All the taxonomic units associated with TMAO and *cutC* belong to only two taxonomic orders, *Bacteroidales* and *Clostridiales*. In particular, almost all the OTUs are attributable to only three families: *Bacteroidaceae*, *Lachnospiraceae*, and *Ruminococcaceae*. Notably, these families have been identified as the most metabolically active bacteria of the human microbiota and play a dominant role in the colonic fermentation of dietary fibers [45,46]. Reportedly, many of these bacteria do not display choline-utilization activities (e.g., cut genes have never been identified in *Bacteroidetes* and *Faecalibacterium*). Nonetheless, we can hypothesize an indirect association of these bacteria with *cutC* and TMAO based on the speculation that the higher presence of these bacteria might determine a greater utilization of the available nutritional sources in the colon, reducing substrates for the remaining bacterial communities. The latter may then receive selective pressure for the expansion of the activities related to the metabolization of the residual energy and carbon sources such as choline, resulting in increased TMA production.

5. Conclusions

Here, we described the results of a preliminary method-development study, which suggests the existence of a relationship between the levels of TMAO excreted in urine, some intestinal taxonomic groups belonging to the most active bacterial families of the colonic microbiota, and a subgroup of the *cutC* gene ascribable to the choline-TMA conversion enzymes of *Enterobacteriaceae*, named cut-Kp, whose relative abundance can be determined with the qPCR protocol developed in this study. Nonetheless, considering the limitations listed above, particularly concerning dietary intake, it is plausible to hypothesize that the results of this study may vary in other populations.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-6643/12/1/62/s1, Figure S1: UPGMA hierarchical clustering based on ClustalW alignment of amino acid sequences of the choline trimethylamine lyase CutC. Figure S2. Verification of choline utilization and TMA production by single bacterial strains. Figure S3: Bacterial community structure of fecal samples. Figure S4: Tukey box and whisker plots representing the most abundant genera (A) and families (B) detected by 16S rRNA gene profiling in fecal samples collected from the adult volunteers participating in this study. Figure S5: Correlations among the fecal relative abundances of the choline TMA-lyase gene *cutC* and bacterial taxa. Table S1: Bacterial strains used for the screening of choline utilization activity. Table S2: Basic characteristics of the study participants.

Author Contributions: Conceptualization, S.G. and P.R.; methodology, S.G., A.D.V., V.T., I.V., V.G., G.R., E.R., C.G., V.D.V. and A.P.; formal analysis, S.G. and G.G; investigation, A.D.V., V.T., G.R., I.V., G.L.V. and A.P.; resources, S.G., V.G. and A.P.; data curation, S.G., A.D.V., V.T., G.G., and A.P.; writing—original draft preparation, S.G.; writing—review and editing, all authors; visualization, S.G., G.G. and A.D.V.; supervision, S.G, V.T. and P.R.; project administration, S.G.; funding acquisition, S.G., P.R. and V.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by the University of Milan Funding "Linea 2-2016", PRISM Project. We acknowledge the European Joint Programming Initiative "A Healthy Diet for a Healthy Life" (JPI-HDHL—http://www.healthydietforhealthylife.eu/) and Mipaaft (Italy; D.M. 8245/7303/2016) for contributing to the grant awarded to G.G.

Acknowledgments: We thank Renata Piccinini for kindly providing the *S. dysgalactiae* strains.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mohajeri, M.H.; Brummer, R.J.M.; Rastall, R.A.; Weersma, R.K.; Harmsen, H.J.M.; Faas, M.; Eggersdorfer, M. The role of the microbiome for human health: From basic science to clinical applications. *Eur. J. Nutr.* **2018**, *57*, 1–14. [CrossRef] [PubMed]
2. Claus, S.P.; Guillou, H.; Ellero-Simatos, S. The gut microbiota: A major player in the toxicity of environmental pollutants? *NPJ Biofilms Microbiomes* **2016**, *2*, 16003. [CrossRef] [PubMed]
3. Ridlon, J.M.; Harris, S.C.; Bhowmik, S.; Kang, D.J.; Hylemon, P.B. Consequences of bile salt biotransformations by intestinal bacteria. *Gut Microbes* **2016**, *7*, 22–39. [CrossRef] [PubMed]
4. Tang, W.H.; Wang, Z.; Levison, B.S.; Koeth, R.A.; Britt, E.B.; Fu, X.; Wu, Y.; Hazen, S.L. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N. Engl. J. Med.* **2013**, *368*, 1575–1584. [CrossRef] [PubMed]
5. Wang, Z.; Klipfell, E.; Bennett, B.J.; Koeth, R.; Levison, B.S.; Dugar, B.; Feldstein, A.E.; Britt, E.B.; Fu, X.; Chung, Y.M.; et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **2011**, *472*, 57–63. [CrossRef]
6. Nowinski, A.; Ufnal, M. Trimethylamine N-oxide: A harmful, protective or diagnostic marker in lifestyle diseases? *Nutrition* **2018**, *46*, 7–12. [CrossRef]
7. Huc, T.; Drapala, A.; Gawrys, M.; Konop, M.; Bielinska, K.; Zaorska, E.; Samborowska, E.; Wyczalkowska-Tomasik, A.; Paczek, L.; Dadlez, M.; et al. Chronic, low-dose TMAO treatment reduces diastolic dysfunction and heart fibrosis in hypertensive rats. *Am. J. Physiol. Heart Circ. Physiol.* **2018**, *315*, H1805–H1820. [CrossRef]
8. Collins, H.L.; Drazul-Schrader, D.; Sulpizio, A.C.; Koster, P.D.; Williamson, Y.; Adelman, S.J.; Owen, K.; Sanli, T.; Bellamine, A. L-Carnitine intake and high trimethylamine N-oxide plasma levels correlate with low aortic lesions in ApoE(-/-) transgenic mice expressing CETP. *Atherosclerosis* **2016**, *244*, 29–37. [CrossRef]
9. Koeth, R.A.; Wang, Z.; Levison, B.S.; Buffa, J.A.; Org, E.; Sheehy, B.T.; Britt, E.B.; Fu, X.; Wu, Y.; Li, L.; et al. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat. Med.* **2013**, *19*, 576–585. [CrossRef]
10. Zeisel, S.H.; Blusztajn, J.K. Choline and human nutrition. *Annu. Rev. Nutr.* **1994**, *14*, 269–296. [CrossRef]
11. Zeisel, S.H.; Mar, M.H.; Howe, J.C.; Holden, J.M. Concentrations of choline-containing compounds and betaine in common foods. *J. Nutr.* **2003**, *133*, 1302–1307. [CrossRef] [PubMed]
12. Wiedeman, A.M.; Barr, S.I.; Green, T.J.; Xu, Z.; Innis, S.M.; Kitts, D.D. Dietary choline intake: Current state of knowledge across the life cycle. *Nutrients* **2018**, *10*, 1513. [CrossRef] [PubMed]
13. Wallace, T.C.; Blusztajn, J.K.; Caudill, M.A.; Klatt, K.C.; Natker, E.; Zeisel, S.H.; Zelman, K.M. Choline: The underconsumed and underappreciated essential nutrient. *Nutr. Today* **2018**, *53*, 240–253. [CrossRef] [PubMed]
14. EFSA Panel on Dietetic Products, Nutrition and Allergies (NDA). Dietary reference values for choline. *EFSA J.* **2016**, *14*, 70.
15. Craciun, S.; Balskus, E.P. Microbial conversion of choline to trimethylamine requires a glycyl radical enzyme. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 21307–21312. [CrossRef] [PubMed]

16. Roberts, A.B.; Gu, X.; Buffa, J.A.; Hurd, A.G.; Wang, Z.; Zhu, W.; Gupta, N.; Skye, S.M.; Cody, D.B.; Levison, B.S.; et al. Development of a gut microbe-targeted nonlethal therapeutic to inhibit thrombosis potential. *Nat. Med.* **2018**, *24*, 1407–1417. [CrossRef]
17. Rath, S.; Heidrich, B.; Pieper, D.H.; Vital, M. Uncovering the trimethylamine-producing bacteria of the human gut microbiota. *Microbiome* **2017**, *5*, 54. [CrossRef]
18. Martinez-del Campo, A.; Bodea, S.; Hamer, H.A.; Marks, J.A.; Haiser, H.J.; Turnbaugh, P.J.; Balskus, E.P. Characterization and detection of a widely distributed gene cluster that predicts anaerobic choline utilization by human gut bacteria. *mBio* **2015**, *6*, e00042-15. [CrossRef]
19. Xu, K.Y.; Xia, G.H.; Lu, J.Q.; Chen, M.X.; Zhen, X.; Wang, S.; You, C.; Nie, J.; Zhou, H.W.; Yin, J. Impaired renal function and dysbiosis of gut microbiota contribute to increased trimethylamine-N-oxide in chronic kidney disease patients. *Sci. Rep.* **2017**, *7*, 1445. [CrossRef]
20. Langille, M.G.; Zaneveld, J.; Caporaso, J.G.; McDonald, D.; Knights, D.; Reyes, J.A.; Clemente, J.C.; Burkepile, D.E.; Vega Thurber, R.L.; Knight, R.; et al. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nat. Biotechnol.* **2013**, *31*, 814–821. [CrossRef]
21. Craciun, S.; Marks, J.A.; Balskus, E.P. Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. *ACS Chem. Biol.* **2014**, *9*, 1408–1413. [CrossRef] [PubMed]
22. Kalnins, G.; Kuka, J.; Grinberga, S.; Makrecka-Kuka, M.; Liepinsh, E.; Dambrova, M.; Tars, K. Structure and function of CutC choline lyase from human microbiota bacterium klebsiella pneumoniae. *J. Biol. Chem.* **2015**, *290*, 21732–21740. [CrossRef] [PubMed]
23. Cattaneo, C.; Gargari, G.; Koirala, R.; Laureati, M.; Riso, P.; Guglielmetti, S.; Pagliarini, E. New insights into the relationship between taste perception and oral microbiota composition. *Sci. Rep.* **2019**, *9*, 3549. [CrossRef] [PubMed]
24. Gargari, G.; Taverniti, V.; Gardana, C.; Cremon, C.; Canducci, F.; Pagano, I.; Barbaro, M.R.; Bellacosa, L.; Castellazzi, A.M.; Valsecchi, C.; et al. Fecal clostridiales distribution and short-chain fatty acids reflect bowel habits in irritable bowel syndrome. *Environ. Microbiol.* **2018**, *20*, 3201–3213. [CrossRef] [PubMed]
25. Caporaso, J.G.; Kuczynski, J.; Stombaugh, J.; Bittinger, K.; Bushman, F.D.; Costello, E.K.; Fierer, N.; Pena, A.G.; Goodrich, J.K.; Gordon, J.I.; et al. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* **2010**, *7*, 335–336. [CrossRef] [PubMed]
26. Johnson, D.W. A flow injection electrospray ionization tandem mass spectrometric method for the simultaneous measurement of trimethylamine and trimethylamine N-oxide in urine. *J. Mass Spectrom.* **2008**, *43*, 495–499. [CrossRef]
27. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw. Artic.* **2015**, *67*, 1–48. [CrossRef]
28. Chhibber-Goel, J.; Singhal, V.; Parakh, N.; Bhargava, B.; Sharma, A. The metabolite Trimethylamine-N-Oxide is an emergent biomarker of human health. *Curr. Med. Chem.* **2017**, *24*, 3942–3953. [CrossRef]
29. Janeiro, M.H.; Ramirez, M.J.; Milagro, F.I.; Martinez, J.A.; Solas, M. Implication of Trimethylamine N-Oxide (TMAO) in disease: Potential biomarker or new therapeutic target. *Nutrients* **2018**, *10*, 1398. [CrossRef]
30. Ufnal, M.; Nowinski, A. Is increased plasma TMAO a compensatory response to hydrostatic and osmotic stress in cardiovascular diseases? *Med. Hypotheses* **2019**, *130*, 109271. [CrossRef]
31. Zeisel, S.H.; Warrier, M. Trimethylamine N-Oxide, the microbiome, and heart and kidney disease. *Annu. Rev. Nutr.* **2017**, *37*, 157–181. [CrossRef] [PubMed]
32. Jameson, E.; Fu, T.; Brown, I.R.; Paszkiewicz, K.; Purdy, K.J.; Frank, S.; Chen, Y. Anaerobic choline metabolism in microcompartments promotes growth and swarming of Proteus mirabilis. *Environ. Microbiol.* **2016**, *18*, 2886–2898. [CrossRef] [PubMed]
33. Falony, G.; Vieira-Silva, S.; Raes, J. Microbiology meets big data: The case of gut microbiota-derived Trimethylamine. *Annu. Rev. Microbiol.* **2015**, *69*, 305–321. [CrossRef] [PubMed]
34. Zarzycki, J.; Erbilgin, O.; Kerfeld, C.A. Bioinformatic characterization of glycyl radical enzyme-associated bacterial microcompartments. *Appl. Environ. Microbiol.* **2015**, *81*, 8315–8329. [CrossRef] [PubMed]
35. Fracalanzza, S.A.; Scheidegger, E.M.; Santos, P.F.; Leite, P.C.; Teixeira, L.M. Antimicrobial resistance profiles of enterococci isolated from poultry meat and pasteurized milk in Rio de Janeiro, Brazil. *Mem. Inst. Oswaldo Cruz* **2007**, *102*, 853–859. [CrossRef] [PubMed]
36. Colombo, E.; Franzetti, L.; Frusca, M.; Scarpellini, M. Phenotypic and genotypic characterization of lactic acid bacteria isolated from Artisanal Italian goat cheese. *J. Food Prot.* **2010**, *73*, 657–662. [CrossRef] [PubMed]

37. De Jong, A.; Simjee, S.; Rose, M.; Moyaert, H.; El Garch, F.; Youala, M.; Group, E.S. Antimicrobial resistance monitoring in commensal enterococci from healthy cattle, pigs and chickens across Europe during 2004-14 (EASSA Study). *J. Antimicrob. Chemother.* **2019**, *74*, 921–930. [CrossRef]
38. Bennett, B.J.; de Aguiar Vallim, T.Q.; Wang, Z.; Shih, D.M.; Meng, Y.; Gregory, J.; Allayee, H.; Lee, R.; Graham, M.; Crooke, R.; et al. Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab.* **2013**, *17*, 49–60. [CrossRef]
39. Zhang, A.Q.; Mitchell, S.; Smith, R. Fish odour syndrome: Verification of carrier detection test. *J. Inherit. Metab. Dis.* **1995**, *18*, 669–674. [CrossRef]
40. Jameson, E.; Quareshy, M.; Chen, Y. Methodological considerations for the identification of choline and carnitine-degrading bacteria in the gut. *Methods* **2018**, *149*, 42–48. [CrossRef]
41. Ufnal, M.; Pham, K. The gut-blood barrier permeability—A new marker in cardiovascular and metabolic diseases? *Med. Hypotheses* **2017**, *98*, 35–37. [CrossRef] [PubMed]
42. Johnson, C.; Prokopienko, A.J.; West, R.E., 3rd; Nolin, T.D.; Stubbs, J.R. Decreased kidney function is associated with enhanced hepatic flavin monooxygenase activity and increased circulating Trimethylamine N-Oxide concentrations in mice. *Drug Metab. Dispos. Biol. Fate Chem.* **2018**, *46*, 1304–1309. [CrossRef] [PubMed]
43. Seibel, B.A.; Walsh, P.J. Trimethylamine oxide accumulation in marine animals: Relationship to acylglycerol storage. *J. Exp. Biol.* **2002**, *205*, 297–306. [PubMed]
44. Jameson, E.; Doxey, A.C.; Airs, R.; Purdy, K.J.; Murrell, J.C.; Chen, Y. Metagenomic data-mining reveals contrasting microbial populations responsible for trimethylamine formation in human gut and marine ecosystems. *Microb. Genom.* **2016**, *2*, e000080. [CrossRef]
45. Wilson, M. The indigenous microbiota of the gastrointestinal tract. In *The Human Microbiota in Health and Disease: An Ecological and Community-Based Approach*, 1st ed.; CRC Press, Taylor & Francis Group: Boca Raton, FL, USA, 2018; 472p.
46. Gosalbes, M.J.; Durban, A.; Pignatelli, M.; Abellan, J.J.; Jimenez-Hernandez, N.; Perez-Cobas, A.E.; Latorre, A.; Moya, A. Metatranscriptomic approach to analyze the functional human gut microbiota. *PLoS ONE* **2011**, *6*, e17447. [CrossRef]