Title	Genetic, enzymatic and metabolite profiling of the Lactobacillus casei group reveals strain biodiversity and potential applications for flavour diversification
Authors	Stefanovic, Ewelina;Kilcawley, Kieran N.;Rea, Mary C.;Fitzgerald, Gerald F.;McAuliffe, Olivia
Publication date	2017-04-04
Original Citation	Stefanovic, E., Kilcawley, K. N., Rea, M. C., Fitzgerald, G. F. and McAuliffe, O. (2017) 'Genetic, enzymatic and metabolite profiling of the Lactobacillus casei group reveals strain biodiversity and potential applications for flavour diversification', Journal of Applied Microbiology, 122(5), pp. 1245–1261. doi:10.1111/jam.13420
Type of publication	Article (peer-reviewed)
Link to publisher's version	10.1111/jam.13420
Rights	© 2017, the Society for Applied Microbiology. This is the peer reviewed version of the following article: Stefanovic, E., Kilcawley, K. N., Rea, M. C., Fitzgerald, G. F. and McAuliffe, O. (2017) 'Genetic, enzymatic and metabolite profiling of the Lactobacillus casei group reveals strain biodiversity and potential applications for flavour diversification', Journal of Applied Microbiology, 122(5), pp. 1245–1261, which has been published in final form at http://dx.doi.org/10.1111/jam.13420. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.
Download date	2025-02-08 10:28:02
Item downloaded from	https://hdl.handle.net/10468/4469

University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Received Date : 29-Nov-2016

Revised Date : 06-Feb-2017

Accepted Date : 08-Feb-2017

Article type : Original Article

Genetic, enzymatic and metabolite profiling of the *Lactobacillus casei* group reveals strain biodiversity and potential applications for flavour diversification

Ewelina Stefanovic[1,2*], Kieran N. Kilcawley[1], Mary C. Rea[1], Gerald F. Fitzgerald[2] and Olivia McAuliffe[1]

[1]Department of Food Bioscience, Teagasc Food Research Centre, Moorepark, Fermoy, Ireland.

[2]School of Microbiology, University College Cork, Cork, Ireland

*Correspondence to: Ewelina Stefanovic, Teagasc Food Research Centre Moorepark, Fermoy, Ireland. Tel: +3532542607, Fax: +3532542222, Email: ewelina.stefanovic@teagasc.ie

Running head: *Lactobacillus casei* biodiversity

Abstract

Aims: The *Lactobacillus casei* group represents a widely explored group of lactic acid bacteria, characterised by a high level of biodiversity. In this study, the genetic and phenotypic diversity of a collection of more than 300 isolates of the *L. casei* group and their potential to produce volatile metabolites important for flavour development in dairy products was examined.

Methods and Results: Following confirmation of species by 16S rRNA PCR, the diversity of the isolates was determined by pulsed field gel electrophoresis. The activities of enzymes involved in the proteolytic cascade were assessed and significant differences in the enzyme activities among the strains were observed. Ten strains were chosen based on the results of their enzymes activities and they were analysed for their ability to produce volatiles in media with increased concentrations of a

representative aromatic, branched-chain and sulfur amino acid. Volatiles were assessed using gas chromatography coupled to mass spectrometry. Strain-dependent differences in the range and type of volatiles produced were evident.

Conclusions: Strains of the *L. casei* group are characterised by genetic and metabolic diversity which supports variability in volatile production.

Significance and Impact of the Study: This study provides a screening approach for the knowledge-based selection of strains potentially enabling flavour diversification in fermented dairy products.

Keywords: *Lactobacillus*, dairy, proteinase, diversity, PFGE

Introduction

Lactobacillus is the largest and most diverse genus of the lactic acid bacteria (LAB), and to date (Nov 2016), comprises more than 170 species (http://www.ncbi.nlm.nih.gov/Taxonomy/Browser). The species of this genus show remarkable niche adaptation, and have been isolated from dairy products and other fermented foods, the human and animal gastrointestinal tract and from plant material (Claesson et al. 2007). The species *Lactobacillus casei* and *Lactobacillus paracasei*, along with *Lactobacillus rhamnosus*, are referred to as the *Lactobacillus casei* group, and are regarded as closely related, both phylogenetically and phenotypically. *Lactobacillus casei* and *Lactobacillus paracasei* represent some of the best explored species within the *Lactobacillus* genus with 89 genome sequences available (Nov 2016) for these species (https://www.ncbi.nlm.nih.gov/genome/genomes/). However, the taxonomic classification of strains is far from straightforward, as often new isolates are named as *L. casei* when they should be named *L. paracasei* since they are more closely related to *L. paracasei* type strain ATCC 334 than to ATCC393, the type strain of *L. casei,* according to the Judicial Commission of the International Committee of Systematics of Bacteria (Tindall 2008). This affects overall nomenclature and continues the on-going debate regarding these two species. Strains of these two species have been isolated from all the usual niches for lactobacilli (fermented products, gastrointestinal tract, environment) (Cai et al. 2009). Their broad ecological distribution reflects their metabolic flexibility and widespread application. The niche adaptability of these two species has been explained through genomic studies where the presence and absence of certain genes important for survival in different niches (specific

carbohydrate metabolism, bile salt resistance genes etc.) and comparative analysis has demonstrated the plasticity of their genomes and their liability to evolutionary changes (Cai et al. 2009; Broadbent et al. 2012; Smokvina et al. 2013; Yu et al. 2015). Acquisition of foreign genes, mainly through horizontal gene transfer, has enabled changes in the metabolic and nutritional capacities of these species and has led to adaptation to more dynamic habitats, such as the gastrointestinal tract and plant materials. Conversely, gene decay, evident in dairy isolates, has narrowed the potential habitats and accommodated niche specialisation (Makarova et al. 2006; Cai et al. 2009; Broadbent et al. 2012). The intra-species heterogeneity and the associated metabolic diversity have provided an opportunity to harness the metabolic potential of strains of the *Lactobacillus casei* group for application in a broad spectrum of fields, from health improvement to food production.

While certain strains of these species are perhaps best known for their characteristic health benefits (Sgouras et al. 2004; Herias et al. 2005; Ivory et al. 2008; Chen et al. 2014), other strains of *L. casei* and *L. paracasei* are commonly found as the dominant species of nonstarter lactic acid bacteria (NSLAB) in ripening cheese (Gobbetti et al. 2015) and are likely to play a role in the development of flavour in these products (Swearingen et al. 2001; Thage et al. 2005; Van Hoorde et al. 2010). The development of flavour results from a complex network of metabolic reactions, which include three main processes: sugar metabolism (glycolysis), lipid degradation (lipolysis), and protein catabolism (proteolysis). Although sugars, mainly lactose, and lipids can be metabolised to flavour compounds, the proteolytic cascade is seen as particularly important for flavour development (Smit et al. 2005). In LAB, this cascade begins with the activity of a surface proteinase, often called a cell-wall, or cell-envelope proteinase (CEP). The peptides produced by the activity of CEP are transported into the cell and degraded by the coordinated action of peptidases with different, but often partially overlapping, specificities. This joint activity of peptidases is crucial for achieving the desired level of proteolysis in cheese (Stressler et al. 2013). As a result of peptidase activity, free amino acids are released. Free amino acids can directly contribute to flavour (McSweeney and Sousa 2000), but it is their further metabolism that is seen as a key process in flavour formation (McSweeney and Sousa 2000; Yvon and Rijnen 2001; Rijnen et al. 2003). There are several pathways of amino acid metabolism in cheese, initiated by the activity of aminotransferases, lyases or decarboxylases (Ardo 2006). However, the majority of the most important flavour compounds originate in transamination pathway. Aminotransferases (AT) transfer the amino group to α-keto acid (most often α-ketoglutarate) (Jensen

and Ardo 2010). Nevertheless, transamination depends on the presence of an amino group receptor, usually α-ketoglutarate, which is produced by glutamate dehydrogenase (GDH) (Kieronczyk et al. 2004), although low level of α-ketoglutarate can be produced in cheese through glutamate catabolism (Christensen et al. 1999). GDH activity has been shown to be a limiting factor for transamination (Tanous et al. 2002) and as such, indirectly represents one of key enzymes for high flavour potential of LAB (Kieronczyk et al. 2003; Kieronczyk et al. 2004; Thage et al. 2005). The assessment of activities of enzymes of the proteolytic cascade could provide information regarding the flavour development capacity of strains.

The aim of this study was to investigate the genotypic and phenotypic diversity of strains of the *L. casei* group and to examine their potential to contribute to flavour development and diversification in dairy products. We focused on two species, *L. casei* and *L. paracasei*, as these species are most commonly associated with the non-starter flora in dairy products, and we designated the strains as belonging to *L. casei* group. The strategy employed included defining the genomic diversity of a selected bank of strains, subsequent assessment of activities of enzymes involved in proteolysis, and finally the determination of volatile flavour production in a single amino acid-enhanced media. The diversity observed at the genetic level was borne out at the phenotypic level. These differences facilitated variations in the metabolic activity resulting in the development of diverse volatile profiles among strains.

Materials and methods

Bacterial strains and culture conditions

A total of 252 strains of the *L. casei* group from the Teagasc Food Research Centre DPC Culture Collection were used in this study. The strains were isolated as a part non-starter microbiota of dairy products including Cheddar, Provola, Comte, and Gouda cheeses, and from other fermented products such as yoghurt and sourdough. The strains were cultivated in MRS media (Oxoid) at 30°C in aerobic conditions. Strain *Lactococcus lactis* subsp. *cremoris* Wg2 was cultivated in LM17 (Merck) at 30°C in aerobic conditions.

DNA isolation, PCR and sequence analysis of the 16S rRNA gene

For DNA isolation, the GeneElute® Bacterial Genomic DNA Kit (Sigma Aldrich) was used according to the manufacturer's instruction. Subsequently, PCR amplification of the 16S rRNA gene was performed with 16S universal primers: UNI16_F: 5'- AGAGTTTGATCCTGGCTCAGG -3', UNI16_R: 5'-ACGGCAACCTTGTTACGAGTT-3' (Alander et al. 1999), which amplify nearly the entire length of 16S rRNA gene (Frank et al. 2008). PCR was performed using the following amplification conditions: initialization at 94°C for 5 min, 40 cycles of denaturation at 94°C for 45 s, annealing at 55°C for 35 s and elongation at 72°C for 1 min, and final elongation at 72°C for 10 min (Eppendorf Mastercycler Pro). Amplicons of size of about 1500bp were purified (Isolate II PCR and Gel Kit, Bioline) and sequenced using the Sanger method (GATC Biotech AG, Koln, Germany). The sequence data generated was compared to the NCBI nucleotide database (http://www.ncbi.nlm.nih.gov/) using the BLAST algorithm (Altschul et al. 1990). The top BLAST hit was taken as confirmation of species.

Pulsed Field Gel Electrophoresis (PFGE)

PFGE was performed as described by Simpson et al. (2002) with slight modifications. Bacterial strains were grown overnight at 30°C in MRS broth containing 20 mmol l^{-1} threonine. For each strain, 1 ml of cell suspension was centrifuged (15000 ***g***, 5 min), washed once in 500 µl of 10 mmol l^{-1} Tris-HCl, 1 mol l^{-1} NaCl, pH 7.6 and resuspended in 200 µl of the same solution, mixed with 200 µl of 2 % low melting point agarose in 0.125 mol l^{-1} EDTA pH 7.6 and left to solidify in moulds at room temperature. Plugs were subjected to cell lysis with a mix of 10 mg ml^{-1} lysozyme and 20 units $µl^{-1}$ mutanolysin in EC buffer (1 mol l^{-1} NaCl, 6 mmol l^{-1} Tris-HCl, 100 mmol l^{-1} EDTA, 1% (w/v) sarkosyl, pH 7.6) for 24 h at 37°C. Subsequently, plugs were subjected to proteolysis with proteinase K (0.5 mg ml^{-1} in 0.5 mol l^{-1} EDTA, 1% (w/v) sarkosyl, pH 8.0) and incubated for 24 h at 55°C. Proteolysis was performed twice, and plugs were washed in 1 mmol l^{-1} phenylmethanesulfonyl fluoride (PMSF) prepared in TE 10/1 buffer (10 mmol l^{-1} Tris-HCl, 1mmol l^{-1} EDTA, pH 8.0) for 1 h at 37°C. Slices (1-2 mm) were cut from the agarose plugs and washed 3 times for 30 min at room temperature with gentle shaking in TE 10/0.1 buffer (10 mmol l^{-1} Tris-HCl, 0.1 mmol l^{-1} EDTA, pH 8.0). Slices were then incubated with 100 µl of the restriction buffer Cut Smart® (New England Biolabs) at 4°C for at least 30 min. The buffer was removed and plugs were incubated for 24 h at 37°C with *Asc*I restriction enzyme (New England Biolabs) in the same buffer. The reaction was stopped by the addition of 0.5 ml of 0.5 mol l^{-1} EDTA pH 8.0. Following digestion, slices were loaded into the wells of a 200 ml 1% PFGE

grade agarose gel (prepared in 0.5 × dilution of TRIS borate-EDTA buffer concentrate, SigmaAldrich). The gels were run in the same 0.5 × TRIS-borate buffer using a CHEF-DR® II PFGE apparatus (Biorad) at 1V (6 V cm^{-1}) for 16 h at 14°C with the pulse ramped from 1 to 20 s. Gels were stained with ethidium bromide (0.5 µg ml^{-1}) for 1 h, and then destained in water for 1 h. Gels were photographed using Alpha Imager® 3400 (Alpha Innotech Corp).

Phylogenetic analysis

PFGE images were processed using BioNumerics® 7.5 software (Applied Maths). Dendrograms were made using the Unweighted Pair Group Method Using Average Linkage (UPGMA) distance matrix method (Sokal and Michener 1958) and curve based Pearson correlation.

Determination of CEP activity

Cell envelope proteinase (CEP) activity was determined using a modification of the method previously described by (Weimer et al. 1997; Gaudreau et al. 2005), which is based on the EnzCheck® kit Green Fluorescence E-6638 (Molecular Probes, Invitrogen). Strains were grown in 35 ml 10% (w/v) reconstituted skim milk (RSM) for 18 h at 30°C. Cells were centrifuged (4000 ***g***, 10 min, 4°C), and washed 3 times with 50 mmol l^{-1} Tris-HCl buffer pH 7.8 with 2 mmol l^{-1} $CaCl_2$ added. After washing, the optical density (OD_{600nm}) of cells was adjusted to approximately OD_{600}=10 in the same buffer. Components of the kit were prepared according to manufacturer's instructions. In a 96-well microplate, 100 µl of cell suspension and 100 µl of prepared BODIPY®FL casein solution were mixed and incubated for 24 h at 30°C. Fluorescence (Ex/Em 505/513 nm) was measured on a Synergy 2 reader (Bio-Tek Multi Detection Plate Reader), using optimal filters: 485/20 nm for extinction and 528/20 nm for emission. A proteinase K solution (2 µg ml^{-1}) was used as a positive control. Enzyme activities for each strain were expressed as direct fluorescence readings. All strains were evaluated in triplicate. *Lactococcus lactis* subsp. *cremoris* Wg2 was used as a control strain. A set of trypsin standards from 0.2 ng ml^{-1} to 70 µg ml^{-1} was made and their activity was measured as for the samples.

Determination of aminopeptidase activities

After incubation of strains in 10 ml MRS broth at 30°C for 18 h, cells were centrifuged (4000 ***g***, 10 min, 4°C) and washed twice with 50 mmol l^{-1} sodium-phosphate buffer pH 7.5, and resuspended in the same buffer to a final volume of 2 ml. To obtain cell free extracts (CFE), cells were disrupted by

sonication (Soniprep 150, MSE LTD) in 5 cycles of 15 s sonication on maximum amplitude (20 amplitude microns) and 45 s of cooling on ice. Sonicated samples were centrifuged (12000 ***g***, 10 min, 4°C) to remove cell debris.

Aminopeptidase assays were performed using a modified method for aminopeptidase activities defined by Jensen and Ardo (2010). Chromogenic substrates (L-Lys-*para*nitroanilide (*p*NA) (Sigma-Aldrich), H-Gly-Pro-*p*NA and H-Arg-*p*NA (Bachem) for PepN, PepX and PepC, respectively) were prepared as 1 mmol l^{-1} solutions in 50 mmol l^{-1} sodium-phosphate buffer pH 7.5. The assay mixture contained 50 µl of substrate solution and 50 µl of CFE. Absorbance was measured at 405 nm (Synergy HT, Bio-Tek Multi Detection Plate Reader) after 30 min of incubation at 30°C. The amount of *p*-nitroaniline released was determined by including a standard curve previously obtained for standard samples of *p*-nitroaniline ranging between 0 to 50 nmol. Aminopeptidase activities were expressed as nmol of *p*-nitroanilne released per min and mg of protein. No positive control was included in this assay, as no commercial enzyme was available. Blanks contained water instead of CFE. Development of yellow colour in the samples, originating from *p*-nitroaniline, and no colour development in the blank after incubation were considered as a sign of enzyme activity of CFE. Protein content was determined by using bicinchoninic acid assay (Pierce®BCA Protein Assay Kit, Thermoscientific). All strains we200-re analysed for all selected aminopeptidase activities in triplicate.

Determination of aromatic aminotransferase activity

The assay to determine aromatic aminotransferase activity was performed by following the conversion of phenylalanine to phenylpyruvate. The assay was based on a method described in Brandsma et al. (2008) with modifications. The assay mixture contained 20 mmol l^{-1} L-phenylalanine, 10 mmol l^{-1} α-ketoglutarate, 0.5 mmol l^{-1}sodium EDTA, 0.05 mmol l^{-1} pyridoxal-5'-phsophate all dissolved in 25 mmol l^{-1} borate buffer pH 8.5. In each well, 150 µl of mixture and 100 µl of CFE (prepared as described above) were mixed, and absorbance was measured after 12 hours incubation at 30°C at 290 nm (Synergy HT, Bio-Tek Multi Detection Plate Reader). The amount of phenylpyruvate released was determined from a standard curve obtained for a set of standards ranging from 5 to 450 nmol of sodium-phenylpyruvate. The aminotransferase activity was expressed as µmol of phenylpyruvate released per mg of protein. No positive control was included in this assay, as no commercial enzyme was available. Negative controls included CFE without added phenylalanine as a substrate and blanks contained water instead of CFE. Change of absorbance in samples containing CFE during

incubation time and no change of absorbance in negative controls and blanks were considered as an evidence of enzyme activity of CFE. Protein content was determined as previously described. All strains were analysed in triplicate.

Determination of glutamate dehydrogenase activity

The glutamate dehydrogenase assay was performed based on the principle described by (Kieronczyk et al. 2003) using a modification of the Megazyme L-Glutamic Acid Kit assay (K-GLUT®, Megazyme International Ireland Ltd). The modification involved supplementing the GDH of the kit with bacterial CFE, which allowed for quantitative determination of GDH activity of the bacterial strains. The original assay conditions and volumes were modified in order to quantify GDH activity in CFEs as follows. The final reaction mixture contained 10 µl of diaphorase, 40 µl of TEA buffer, 20 µl of glutamic acid solution (0.1 mg ml^{-1}), 20 µl of INT-NAD^+ solution (all of these supplied in K-GLUT® Kit) and 100 µl of CFE. Absorbance was measured at 492 nm after 1 h of incubation at 37°C (Synergy HT, Bio-Tek Multi Detection Plate Reader). One unit of GDH activity corresponded to the amount of enzyme that resulted in an increase of absorbance of 0.01 per 1 min. No positive control was included in this assay. Blank contained water instead of CFE. The development of red colour product (INT-formazan) in samples and no colour development in the blank after incubation were considered as an evidence of enzyme activity of CFE. Protein content was determined as previously described. Specific enzyme activity was expressed as the number of units (U) per mg of protein. All strains were analysed in duplicate.

Production of volatile compounds from single amino acid metabolism

In order to evaluate the metabolic activity of selected strains in the presence of a predominance of a single amino acid, a set of specific media was prepared. The media contained 50 g l^{-1} of Bactotryptone® (BD), 12 g l^{-1} NaCl and the specific selected amino acid added to final concentration of 50 mmol l^{-1}. The amino acids chosen for this analysis were phenylalanine, leucine and methionine to demonstrate the metabolic activity of strains towards aromatic, branched chain and sulfur amino acids, respectively. The corresponding media were designated as PEM (phenylalanine-enhanced media), LEM (leucine-enhanced media) and MEM (methionine-enhanced media), and for this analysis, 10 strains were chosen based on the results of all the enzymatic assays described above. Strain DPC1116 had the highest activity of PepN, one of the highest PepX activities, high PepN activity and medium CEP, AT, and GDH activities. Strain DPC2068 had medium CEP activity, high

PepN, PepC and PepX activities, medium ArAT activity and low GDH activity. Strain DPC2071, had the highest CEP activity, high PepN, PepC activities, medium PepX and ArAT activities and low GDH activity. Strain DPC3990 had high CEP activity, high PepN, PepC activities, medium PepX activity and low ArAT and GDH activities. Strain DPC4026 had low CEP activity, medium PepN and PepC activities, low PepX, ArAT and GDH activities. Strain DPC4206 had high activities in all enzyme assays. Strain DPC4536 had low CEP activity, medium activities for PepN, PepC and PepX and ArAT assays, but the highest activity of GDH. Strain DPC5408 had high CEP, medium PepN, PepC and PepX activities, and low activities of ArAT and GDH. Strain DPC6753 had high activity of CEP, low activities of PepN, PepC and PepX, and high activities of ArAT and GDH. Strain DPC6800 had low CEP activity, high PepN and PepC activities, and the highest PepX activity, high ArAT and medium GDH activity. Cells were prepared according to the protocol described by Van de Bunt et al. (2014) with some modifications. Briefly, strains were incubated overnight (30°C) in 10 ml of MRS, and were re-inoculated (1 % v/v) into 500 ml of MRS and incubated for 24 h at 30°C until they reached stationary phase. Cells were centrifuged (4000 ***g***, 10 min, 4°C), washed twice with 0.1 mol l^{-1} sodium phosphate buffer pH 6, and resuspended in 5 ml of the same buffer containing 15 % glycerol and kept at -80°C until required. Thawed cell suspensions (1 ml) were inoculated into 9 ml of prepared each of three media described above (PEM, LEM, MEM). Three replicate vials were made for each strain. For one replicate of each strain, cell counts (CFU ml^{-1}) were performed. A 100 µl aliquot was taken for plate counting at t=0 h and the strains were incubated for 48 h, after which another 100 µl aliquot was taken for plate counting (t=48 h). Plate counting was performed on MRS agar plates, which were incubated aerobically for 72 h at 30°C. Samples were kept in -20°C until required for volatile analysis. The control consisted of non-inoculated media. The samples and the control were tested in triplicate.

Head Space Solid Phase Microextraction Gas Chromatography Mass Spectrometry Analysis (HS SPME GC-MS)

For each sample of each of three media, 2 g of the sample was placed in an amber 20 ml screw capped HS-SPME vial with a silicone/PTFE septum (Apex Scientific). The vials were equilibrated to 40°C for 10 min with pulsed agitation of 5 seconds at 500 rpm using a heated stirrer module on a Shimadzu AOC 5000 plus autosampler. A single 50/30 µm Carboxen™/ divinylbenzene/ polydimethylsiloxane (CAR/DVB/PDMS, Agilent Technologies) fibre was used to perform solid phase microextraction (SPME). The SPME fibre was exposed to the headspace above the samples for 20

min at 40°C. After extraction, the fibre was injected into the GC inlet via a merlin microseal and desorbed for 2 min at 250 °C into a SPL injector with a SPME liner. Injections were made on Shimadzu 2010 Plus GC with a DB-5 (60 m x 0.25 mm x 0.25 µm, Agilent Technologies) column in splitless mode using a split/splitless injector. Helium was used as a carrier gas, which was maintained at 23 psi. The temperature of the column oven was set at 35°C, held for 5 min, increased at 6.5°C/min to 230°C then increased at 15°C/min to 320°C, yielding at total GC run time of 41.5 min. The mass spectrometer detector Shimadzu TQ8030 was run in single quad mode. The ion source temperature was 230°C, the interface temperature were set at 280°C and the MS mode was electronic ionization (-70 ev) with the mass range m/z scanned between 35 and 250 amu.

All samples were analysed in the same GC run. A set of external standards (dimethyl-sulfide, benzaldehyde, cyclohexanone, butyl acetate, acetone, and ethanol at concentrations of 10 ppm) was also run at the start of the sample set to ensure that both the HS-SPME extraction and MS detection was within specification. Blanks (empty vials) were injected regularly to monitor possible carry over. The SPME fibre was cleaned between samples using a bake-out station on the AOC 5000 at 270°C for 3 min to ensure no carry over between samples.

Data processing and compound identification

Chromatograms obtained by GC analysis were converted to .cdf format and processed by TargetView® (Markes International Ltd). Compounds of interest were chosen according to previously published reviews of flavour contributing compounds (Curioni and Bosset 2002; Singh et al. 2003; Smit et al. 2005). Identification of compounds was based on the results of a comparison with the NIST 2011 Mass Spectral Library (Scientific Instrument Services, Ringoes, NJ, USA) and an in-house library for flavour compounds and confirmed by calculating linear retention indices as described in (Van den Dool and Kratz 1963).

Statistical analysis

All enzymatic assay results were statistically analysed using one-way Analysis of Variance (ANOVA) followed by the least significant difference (LSD) post hoc test. ANOVA and LSD were also used for testing the significance of differences in cell counts in three media (PEM, LEM and MEM) at t=0 h and t=48 h. Principal Component Analysis (PCA) was used for analysis of GC-MS data. All listed statistical tests were performed in R statistical software (R Core Team 2015, R Foundation for Statistical Computing, Austria, https://R-project.org).

Results

Origins of *L. casei* group strain bank and confirmation of species

Initially, 310 isolates were selected for analysis in this study. The isolates had previously been assigned as *Lactobacillus casei, L. paracasei* or simply as *Lactobacillus* isolates in the Teagasc DPC Culture Collection and they originated from dairy products or sourdough. Each of the 310 isolates was subject to 16S rRNA sequencing, and based on the BLAST analysis of the amplified sequences, 252 out of 310 isolates were confirmed as either *L. casei* or *L. paracasei*.

Comparative phylogenetic analysis reveals extensive genomic diversity in *L. casei* group strain bank

To assess the diversity of the 252 confirmed isolates belonging to the *L. casei* group, PFGE was used to generate genomic fingerprints. Grouping of the distinct strains was performed by comparing PFGE fingerprints with BioNumerics®7.5 software, but also by a simple visual comparison, as some of fingerprints were distant in the dendrogram but very similar when checked manually. Analysis of the generated PFGE patterns revealed 98 distinct profiles among the 252 isolates, representing 98 distinct strains. Figure 1a represents a dendrogram of the PFGE profiles of the 98 diverse strains. Additionally, strain DPC4536, that has indistinguishable fingerprint to strain DPC4206 was included in further enzyme activity evaluation, to observe the potential phenotypic differences between these two strains (Figure 1b). This means that in total, 99 strains were analysed in enzymatic assays.

Key enzymatic assays show diverse activities among selected strains

To assess the cell envelope proteinase activity (CEP) of the strains, a kit based on the proteolysis of BODIPY®FL-labelled casein derivatives which release highly fluorescent peptides, was used. This assay is based on the principle that the measured increase in fluorescence is proportional to the proteinase activity. All 99 strains showed CEP activity, but the levels varied significantly from strain to strain. The CEP activity was expressed as measured fluorescence and it ranged from 80.3 arbitrary fluorescence units for strain DPC4764 to 229.7 arbitrary fluorescence units for strain DPC2071 (Figure 2), which corresponded to fluorescence measured when standard solutions of trypsin in the range of 0.4 to 4.0 μg ml^{-1} were used (data not shown). *Lactococcus lactis* subsp. *cremoris* Wg2 was used as a CEP-positive control strain, having been confirmed as such in a previous study (Kok et al. 1988; Laan and Konings 1989; Nikolic et al. 2009), and its CEP activity was 205.33 arbitrary fluorescence units. The strains DPC2071, DPC4206, DPC3990, (Figure 2) had similar or even higher

activity than the control strain *Lactococcus lactis* subsp. *cremoris* Wg2 and in total, 25 strains displayed activity that was not statistically different to the activity of *Lactococcus lactis* subsp. *cremoris* Wg2 (Supporting Information Table 1).

The strains showed significantly different activities towards the various aminopeptidase substrates tested (L-Lys-*para*nitroanilide (*p*NA), H-Gly-Pro-*p*NA and H-Arg-*p*NA for PepN, PepX and PepC, respectively). PepN activities ranged from 0 to 54.2 nmol *para*-nitroaniline (min*mg protein)$^{-1}$ for DPC5336 and DPC1116 respectively; PepC activities ranged from 0 for strains DPC4139, DPC4140 and DPC5410 to 50.3 nmol (min*mg protein)$^{-1}$ for DPC4680, and PepX activities from 0 (31 strains) to 39.2 nmol (min*mg protein)$^{-1}$ for DPC6800 (Figure 3, Supporting Information Table 1). Statistical analysis confirmed that aminopeptidase activities differed significantly between the strains. For example, for PepX, 39 strains showed activity below 5 nmol (min*mg protein)$^{-1}$, and therefore only eight of these strains were carried forward for further analysis of enzyme activities.

When aromatic aminotransferase activity was measured, strain DPC5411 showed the highest activity with 3.28 µmoles of phenylpyruvate released per mg protein while DPC4805 had the lowest activity of 0.28 µmol mg protein^{-1}. Statistical analysis revealed significant differences among the strains for aromatic aminotransferase activity (Figure 4). The GDH activity ranged from 0 for strains DPC6084, DPC4802 and DPC4026 and 17.5 U mg^{-1} of protein for strain DPC4536 and differences in GDH activities were shown to be significantly different (Figure 5).

HS SPME GC-MS volatile analysis confirms metabolic diversity of selected strains

Strains DPC1116, DPC2068, DPC2071, DPC3990, DPC4026, DPC5408, DPC6753, DPC6800, and two strains which have indistinguishable genetic fingerprints (DPC4206 and DPC4536) were selected as candidates for volatile analysis on the basis of their spectrum of key proteolytic enzyme activities. The strains were assessed for their capacity to metabolise amino acids in three distinct media containing elevated levels of a single amino acid (phenylalanine (PEM), leucine (LEM) or methionine (MEM)). The enumeration of cells in these media, pre- and post-incubation, is presented in Supporting Information Table 2. In all three amino acid-enhanced media, six of the ten strains (DPC1116, DPC2071, DPC4206, DPC4536, DPC5408, DPC6800) showed significantly lower cell numbers after 48 hours incubation. In the case of strain DPC4026, significantly lower cell number after 48 hours of incubation was observed in only MEM.

Compounds selected as flavour-contributing in samples in all three media are listed in Table 3 of Supporting Information. The compounds of interest were selected according to the previously published reviews of compounds considered as main flavour contributors in cheese (Curioni and Bosset 2002; Singh et al. 2003; Smit et al. 2005). The highest number of the relevant volatile compounds (47) were present in PEM media, with metabolites containing aromatic ring structures, such as benzenacetaldehyde, 3-ethyl-benzaldehyde, 1,3-xylene, tetramethyl-benzene, hexyl-benzene and methyl-naphtalene exclusively present in this medium. Acetic and butanoic acids and long-chain ketones (C7-C13) were also only detected in PEM. In addition, PEM samples had the highest number of alcohols, aldehydes, and ketones. Of the 47 compounds selected, significant differences in the relative abundance of 27 compounds were observed. In LEM samples, 25 volatiles were detected, and 24 were present in significantly different amounts for various strains. In MEM samples, 22 volatiles were detected and significant differences in abundances of all 22 compounds were observed between the strains tested. No specific metabolites, present in LEM or MEM exclusively, were identified.

The relative abundances of compounds for which significant differences among the strains, including the control, were observed, are presented in Figure 6. In all three media, 1-butanol was present at the highest abundances compared to all other volatiles (no significant difference between the strains including the control, in PEM), and in PEM and MEM, the abundance of this compound was the highest in the control. Several strains showed unique abilities to produce certain volatiles. Strain DPC2068 produced significantly higher amounts of butyl-3-methylbutanoate, a metabolite of leucine, in PEM compared to all other strains, while in MEM and LEM, this strain was the only producer of this compound. The same observation was made in the case of butyl-2-methyl propanoate (metabolite of leucine), butyl butanoate, butyl propanoate (secondary metabolites of amino acids) in all three media. In MEM and LEM, this strain produced a significantly higher abundance of dimethyl-trisulfide (DMTS), originating from methionine, compared to all other strains. An interesting observation was made for diacetyl and acetoin, important flavour compounds originating from sugar metabolism. Strain DPC4026 was the only producer of diacetyl in both LEM and MEM, and in PEM the production by DPC4026 was significantly higher compared to all other strains. DPC4026 also produced the highest abundance of acetoin in PEM and LEM, and significantly higher abundance of acetoin compared to all other strains in MEM. Strain DPC4206 produced the highest abundance of dimtehyl-disulfide (DMDS)

and methanethiol, methionine metabolites, in both MEM and LEM, and the abundance of methanethiol in LEM was significantly higher compared to all other strains. In LEM, this strain produced the highest amount of 3-methyl-butanal (leucine metabolite), while in PEM it produced the highest levels of butanoic acid. Strains DPC4206 and DPC1116 produced the highest abundances of 3-methyl-butanol (leucine metabolite) in all three media, and DPC1116 produced significantly higher abundance of this compound in MEM compared to all other strains. Strain DPC6800 produced the highest abundance of the two main compounds arising from phenylalanine metabolism, benzaldehyde and benzyl-alcohol, and the abundance of benzyl-alcohol was significantly higher compared to all other strains in PEM.

PCA plots obtained after analysis of total ion chromatograms for 10 chosen strains in three media (PEM, LEM and MEM) are presented in Figure 7. PCA plots were generated by using only those compounds (variables) for which significant differences in abundance among the strains, including the control, were observed. In the case of PEM samples, dimension (PC1) described 26.1 % variation and dimension 2 (PC2) described 22.0 % total variation between the strains. In the PCA plot for LEM samples, PC1 described 32.2 % of variation, while PC2 described 26.6 % variation between the strains. In the PCA plot for strains inoculated in MEM, PC1 described 33.8 % variation, and PC2 described 28.1 % variation. The control clearly separated from all strains tested on plots for PEM and MEM, while on LEM plot, the control was positioned in the central part of the plot, along with the majority of other strains. In all three plots, two strains, DPC2068 and DPC 4206, were positioned more separately from other strains and the control. The position of DPC2068 was determined by the relative abundance of butyl-2-methyl-propanoate, butyl-3-methyl-butanoate, butyl propanoate and butyl butanoate, in all plots, and additionally by 3-methyl-butyl-acetate and butyl butanoate in PEM and MEM, and DMTS in MEM. The position of strain DPC4206 was associated with DMDS, butanoic acid and ethyl benzene in PEM, 3-methyl-butanal, 3-methyl-butanoic acid, 3-methyl-butanol, DMDS and methanethiol in LEM and DMDS, 3-methyl-butanol, acetone, butanone and methanethiol in MEM.

Two strains with the same genomic fingerprint show different phenotypic characteristics

In addition to the 98 strains with diverse genomic fingerprints, strain DPC4536, that has an indistinguishable PFGE pattern from strain DPC4206 (Figure 1b) was included in the enzymatic assays. Different activities were determined for all enzymes analysed and significant differences in activities were observed for CEP and GDH (Figures 2, 3, 4 and 5, Supporting Information Table 1).

These two strains were also compared for their abilities to produce volatile compounds in the three amino acid-enriched specific media. The volatile profiles differed in types and abundances of compounds detected (Figure 6), with specificities of DPC4206 metabolic characteristics described in the above section. The differences in the metabolite production is also visible in PCA plots (Figure 7), where DPC4536 is located closer to other strains, while DPC4206 is one of two strains that were the most separated from the other strains.

Discussion

Lactobacillus is by far the largest and most diverse genus of LAB, with *L. casei* and *L. paracasei* being some of the best explored species of this genus (Broadbent et al. 2012). The genome evolution was elucidated in comparative genome analysis and it explained the adaptation of this species to numerous dynamic, nutritionally variable environments, such as gut, plant and milk The wide ranging habitats of these species make them relevant subjects for research on genetic diversity and niche expansion.

In the literature, interchangeable use of the names *L. casei* and *L. paracasei* occurs. There is an on-going debate about type strains for *L. casei* and *L. paracasei*, and whether they constitute one or two species (Dellaglio et al. 1991; Dicks et al. 1996; Dellaglio et al. 2002). Any newly isolated strains are often designated as *L. casei*, but they should be named *L. paracasei,* according to current taxonomy and defined type strains (Tindall 2008). Attempts have been made to introduce novel approaches for differentiation of *L. casei, L. paracasei*, *L. zeae* and *L. rhamnosus* such as the use of species specific 16S primers (Ward and Timmins 1999), and Temporal temperature gradient gel electrophoresis (TTGE) of obtained 16S rRNA PCR amplicons (Vasquez et al. 2001). Additionally, combining results of restriction endonuclease analysis of total DNA, TTGE of 16S rRNA PCR and ribotyping (Vasquez et al. 2005) showed that numerous heterogeneities found in 16S rRNA genes in *L. casei/paracasei* and related species *L. zeae* and *L. rhamnosus* make definitive separation of these species complex and difficult. Recently, a new method based on high resolution melting analysis of PCR amplicons obtained with sets of species discriminating primers (Iacumin et al. 2015) has been proposed. Although for the majority of the 194 strains tested, this method gave satisfactory results in species identification. However, six strains showed inconsistencies in identification. In the case of strains from the DPC culture collection, analysis with strain-specific primers did not enable satisfactory

differentiation of the two species (data not shown) and to that end, a more general approach that involved conventional 16S rRNA sequencing was used to confirm the isolates as part of the *L. casei* group. In addition, the strain genomic profiling was assessed and 98 distinct PFGE profiles were detected, confirming the variability of the isolates. Surprisingly, although the strains originated mainly from cheese, the observed level of diversity was high. The reason behind this could be that the isolation source does not necessarily match the original habitat, as they could be part of the NSLAB flora or starter mixture used during cheese manufacture, or come from other sources during cheese handling and ripening, e.g. contamination from personnel, surfaces etc. Besides that, isolates were collected over a period of more than 20 years, and this time-span contributed most likely to the variety of isolates available in the DPC culture collection.

It is envisaged that differences in genomic structure present a natural basis for genetic variation among the strains, which would further facilitate strain-to-strain variation in their phenotypic characteristics, including their flavour-forming ability. *L. casei* and *L. paracasei* are often part of NSLAB flora of cheese during ripening, and are seen as an important contributors to flavour development, due to the metabolic capacity of these strains that balances the degradation of substrates present and leads to formation of volatiles affecting organoleptic characteristics of cheese (Banks and Williams 2004; Sgarbi et al. 2013).

The metabolic activity of microorganisms present in cheese during ripening results in development of flavour compounds (Marilley and Casey 2004), and the metabolic products of three biochemical pathways: glycolysis, lipolysis and proteolysis are seen as cheese flavour contributors (McSweeney and Sousa 2000). Lactate is the main metabolite of the primary milk sugar lactose, and it can be metabolised to flavour compounds such as diacetyl, acetoin, acetaldehyde, or acetic acid, via a pyruvate intermediate (Smit et al. 2005). Lipolytic reactions during cheese ripening result in free fatty acid production, and short and intermediate chain fatty acids either contribute to flavour themselves, or represent the starting molecules for the production of other flavour compounds (Collins et al. 2003). Although products of both glycolysis and lipolysis can contribute to flavour, in bacterial ripened cheeses catabolic products of proteolytic reactions, mainly metabolites of free amino acids, represents the major flavour contributing metabolic pathway (Smit et al. 2005), and because of this, activities of the enzymes of the proteolytic cascade were assessed.

The activity of the cell-envelope proteinase (CEP), which starts the cascade by cleaving casein molecules to shorter peptides, was shown to vary from strain to strain. This enzyme has an important role in flavour development, as casein hydrolysis is a cascade process and greater proteolytic activity in the earlier steps will result in the exponential generation of more flavour compounds in later metabolic steps. Various approaches have been developed to determine CEP activity, such as ones based on absorbance measurement after cleavage of chromogenic substrate (Fernández de Palencia et al. 1997; Hebert et al. 2008), or fluorescent measurement after degradation of fluorescently labelled caseins such as used in this study (Weimer et al. 1997; Wakai et al. 2013). Our results show that CEP activity of strains of the *L. casei* group is quite variable. These results correspond to the results obtained by Weimer et al. (1997), which demonstrated inter- and intraspecies differences in CEP activities for *Lactococcus lactis, L. casei* and *L. helveticus.*

The next step of proteolysis, the degradation of peptides to free amino acids is catabolised by aminopeptidases, some of which cleave only one type of amino acid and have a very narrow activity (glutamyl-aminopeptidase, PepA), while others, such as general aminopeptidases break the bond between various amino acids at the N-terminal end of the peptide (general aminopeptidases PepN, PepC). Besides these, important enzymes in efficient peptidolysis are dipeptidyl aminopeptidases, such as PepX (Magboul and McSweeney 1999), as they remove proline residues thus enabling further degradation of proteins. Gonzalez et al. (2010) demonstrated that cell-free extracts of *L. paracasei* isolated from cheese exhibited the highest activity towards Ala-, Lys-, Pro-, and Leu-*p*NA substrates compared to other LAB isolates which included leuconostocs, lactococci and enterococci. Similar findings were reported by Ayad et al. (2004) and Macedo et al. (2000), where strains of *L. paracasei* were shown to have the highest aminopeptidase activities compared to leuconostoc, lactococcal or enterococcal strains. In this study, strain dependent activities of *L. casei* group were confirmed.

A study conducted on a series of lactococcal strains revealed that the activity of aromatic aminotransferases resulted in a more diverse volatile profile than the activity of branched chain amino acid aminotransferases (Rijnen et al. 2003). On the other hand, the specificity of aminotransferases towards a certain type of amino acid is not absolute as shown in lactococcal strains where aromatic aminotransferase was able to degrade aromatic amino acids, but also methionine, a sulfur-containing amino acid (Rijnen et al. 2003) and leucine, a branched chain amino acid (Christensen et al. 1999).

With this in mind, determination of aromatic aminotransferase activity was seen as a suitable test for general aminotransferase activity determination. The results obtained in this study for strains of the *L. casei* group illustrate a diverse range of aromatic aminotransferase activities, which is one of the crucial steps for the diversity of flavour compounds produced and confirms strain-specificity of the enzyme activity previously reported (Thage et al. 2004). Similar results to the ones from this study were reported for *L. helveticus* and *L. danicus* grown in MRS (Jensen and Ardo 2010; Pedersen et al. 2013).

Transamination, the transfer of an amino group from an amino acid to a keto acid, is enhanced when α-ketoglutaric acid is present as the amino group acceptor. Strains possessing glutamate dehydrogenase (GDH), that converts glutamate to α-ketoglutarate, are more likely to have an impact on flavour (Williams et al. 2006). Activity of GDH can depend on NAD or NADP as cofactors. Previously, Kieronczyk et al. (2003) reported that *L. paracasei* strains (INF15D, 2756) and *L. casei* 1244 did not possess NAD-dependent GDH activity, while *Lactococcus lactis* subsp. *cremoris* NCDO763 expressed low activity. Similarly, no NAD-GDH activity was observed for any of the NSLAB lactobacilli (*L. plantarum, L rhamnosus, L. parabucknerii, L. curvatus*) (De Angelis et al. 2010) or for *L. plantarum* or *L. paracasei* strains in the study of Tanous et al. (2002), but it was detected in the case of *Lactococcus lactis* subsp. *lactis* NCDO 1867. Conversely, Williams et al. (2006) showed that several *Lactobacillus* species possessed both NAD and NADP dependent GDH activity. In contrast, in the present study, NAD-GDH activity was detected in all strains and the activities were higher than those obtained for NADP-GDH activity by De Angelis et al. (2010), yet they correspond to those obtained by Kieronczyk et al. (2003) for NADP-activity of GDH in *L. paracasei* INF15D.

All of the enzyme assays performed showed that strains in our culture collection bank expressed a range of activities of proteolytic enzymes supporting the diversity observed on the genetic level. Statistical analysis revealed groups that differed in activities in each enzymatic assay, confirming the significance of the observed activity variations.

Metabolism of amino acids is considered as the most important process contributing to the development of flavour compounds (Yvon and Rijnen 2001). These compounds include short chain acids, alcohols and aldehydes, often with an aliphatic branch (volatiles originating from leucine, valine, isoleucine metabolism), as well compounds with aromatic ring structures originating from phenylalanine or tyrosine, and sulfur-containing compounds, which are the product of methionine

metabolism (Yvon and Rijnen 2001; Singh et al. 2003). Some of the most important flavour-forming compounds that arise from phenylalanine metabolism are benzaldehyde (almond flavour), phenylethanol (rose flower) (Curioni and Bosset 2002) and benzenacetaldehyde (bitter almond aromatic flavour) (Jung et al. 2013). The leucine metabolites, such as 3-methyl-butanol, 3-methyl-butanal and 3-methyl-butanoic acid, have malty, fresh cheese and rancid-sweet odours, respectively (Smit et al. 2005). Volatile sulfur compounds are important in overall cheese flavour. They include DMDS, DMTS and methanethiol which are described as having an onion, garlic and cabbage odour, respectively (Singh et al. 2003; Yvon 2006). Secondary products, such as esters, also contribute to flavour, mainly with sweet fruity notes. For this reason, diversity in activity of amino acid converting enzymes was assessed. The base of the media used was pancreatic digest of casein, which contained all of the amino acids ranging between 0.4 % for aspartic acid up to 5.5 % for lysine, expressed as % of free amino acids. This medium was modified by addition of a single predominant amino acid, to explore metabolic preferences of the strains. The strains for these assays were selected on the basis of their varying enzyme activities, but most importantly, their different aminotransferase activities (Supporting Information Table 1). Since amino acids dominated as the metabolic precursors in these media, as expected, the volatiles identified included the breakdown products typical of amino acid catabolism, confirming the general amino acid metabolic activity of the strains.

Figure 6 presents the relative abundances of compounds identified in samples of all three media in significantly different abundances among the strains, including the control. The overall trend in these experiments was that while more metabolites were present in samples of PEM, many of these compounds were not detected in LEM and MEM. In addition, the relative abundances of compounds were higher, often significantly higher, in PEM, compared to MEM and LEM (data not shown); however, the reason for this observation is not clear. In samples of all three media, 1-butanol was present in the highest abundances. This compound most probably originated from components of media generated during sterilisation. Butanol present in such a high abundances was the substrate for butyl-esters formation in samples of all three media. Long chain ketones were detected particularly in PEM. This could mean that in PEM a greater ratio of cell division/lysis occurred, leading to more long-chain ketones, normally of fat origin, and in this fat free medium which contained no added fats probably came from the metabolism of lipids released from the cell membrane after cell lysis.

Interestingly, 2,3-butanedione (diacetyl) was detected in LEM and MEM, and 3-hydroxy-butanone (acetoin), was detected in all three media (no significant difference in PEM). These two compounds are important flavour contributors (buttery flavour) and they most probably originated from low level of sugar present in Bactotryptone® itself, as no sugar source was added during media preparation. In LEM and MEM, only DPC4026 produced 2,3-butanedione, suggesting a potential energy source for this strain. Among the ten strains that were analysed in PEM, LEM and MEM, strain DPC6800 showed the highest activity of aromatic aminotransferase as determined by the *in vitro* assay, and this strain was confirmed as the highest producer of the most important molecules arising from phenylalanine, such as benzaldehyde and benzyl-alcohol. However, ethyl-benzene, 1,3-xylene, tetramethyl-benzene, etc. were present in higher concentrations in samples of other strains, and they probably emerged in further degradation of phenylalanine and its metabolites.

Principal component analysis (PCA) was used to emphasize variation and determine strong patterns in the datasets. PCA identified differences in total volatile production between the strains, and revealed, two strains, DPC2068 and DPC4206, that differed in metabolic potential compared to other strains in all three media (Figure 7). The position of strain DPC2068 in the PCA plots from all three media was predominately due to its association with metabolites originating from the metabolism of branched chain amino acids. Apparently, this strain has high activity towards branched chain amino acids even in the media with less availability of these and abundances of other amino acid, such as PEM and MEM. On the other hand, strain DPC4206 has the most diverse metabolic activity compared to other strains in all three media, and it was able to produce the broadest range of volatiles, often of the highest relative abundance (Figure 6). These two strains showed outstanding volatiles patterns and their metabolic activity could lead to diverse flavour development in fermented dairy products.

Interestingly, strains DPC4206 and DPC4536, which have the same PFGE fingerprints, were shown to have considerably different phenotypic characteristics, based both on the enzymatic assays and volatile analysis (Figures 6 and 7). The PFGE analysis for these two strains was performed with additional enzymes (*Apa*I, *Cla*I, and combination of both *Apa*I and *Cla*I), as well as RAPD PCR, and no differences in band patterns for these two strains was observed (data not shown). Similar results were obtained during analysis of *Listeria monocytogenes* strains in our laboratory (unpublished data). These findings highlight that strains sharing the same PFGE pattern do not necessarily have same genetic and, subsequently, phenotypic characteristics. This study showed that PFGE is not a

definitive tool to determine strains genetic diversity, but rather a robust method used for assessing differences in genomic structure and observing larger evolutional events, such as large insertions, deletions and rearrangements of DNA (Cai et al. 2007). Although used a golden standard for assessing strain diversity based on whole genome restriction analysis for long time, nowadays PFGE has slowly being replaced by the whole genome sequencing, which enables deeper insight in gene content differences among the strains and becomes a preferable method to record subtle genetic differences which would not be apparent in the PFGE profiles alone. For this reason, whole genome sequence analysis of the two strains, DPC4206 and DPC4536 is currently performed in our laboratory, in order to reveal the genetic basis of different phenotypic characteristics of these two strains observed in this study.

The study presented in this paper gives an insight into both the genetic and phenotypic diversity of strains of *L. casei* group. The observed level of genetic diversity can be considered as very broad, since the majority of isolates have the same origin of isolation. The analysed strains, including the two strains with the identical genetic fingerprint, showed variable phenotypic traits, as observed in assays determining the activities of proteolytic cascade enzymes. Additionally, the strains demonstrated different capacities for production of flavour compounds, and two strains, DPC2068 and DPC4206, were particularly diverse in their volatiles production. It can be inferred that strains of *L. casei* group have different abilities for volatile production, which makes them potentially useful for dairy product flavour diversification.

Acknowledgements

The authors wish to thank to Douwe van Sinderen for providing strain *Lactococcus lacts* subsp. *lactis* Wg2. Also, the authors are grateful to Anne Thierry for assistance with R software for statistical analysis. Ewelina Stefanovic is in a receipt of Teagasc Walsh Fellowship.

Conflict of interest

The authors declare no conflict of interest.

References:

Alander, M., Satokari, R., Korpela, R., Saxelin, M., Vilpponen-Salmela, T., Mattila-Sandholm, T. and von Wright, A. (1999) Persistence of colonization of human colonic mucosa by a probiotic strain, *Lactobacillus rhamnosus* GG, after oral consumption. *Appl Environ Microbiol* **65**, 351-354.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403-410.

Ardo, Y. (2006) Flavour formation by amino acid catabolism. *Biotechnol Adv* **24**, 238-242.

Ayad, E.H.E., Nashat, S., El-Sadek, N., Metwaly, H. and El-Soda, M. (2004) Selection of wild lactic acid bacteria isolated from traditional Egyptian dairy products according to production and technological criteria. *Food Microbiol* **21**, 715-725.

Banks, J.M. and Williams, A.G. (2004) The role of the nonstarter lactic acid bacteria in Cheddar cheese ripening. *Int J Dairy Technol* **57**, 145-152.

Brandsma, J.B., Floris, E., Dijkstra, A.R.D., Rijnen, L., Wouters, J.A. and Meijer, W.C. (2008) Natural diversity of aminotransferases and dehydrogenase activity in a large collection of *Lactococcus lactis* strains. *Int Dairy J* **18**, 1103-1108.

Broadbent, J.R., Neeno-Eckwall, E.C., Stahl, B., Tandee, K., Cai, H., Morovic, W., Horvath, P., Heidenreich, J., Perna, N.T., Barrangou, R. and Steele, J.L. (2012) Analysis of the *Lactobacillus casei* supragenome and its influence in species evolution and lifestyle adaptation. *BMC genomics* **13**, 533.

Cai, H., Rodriguez, B.T., Zhang, W., Broadbent, J.R. and Steele, J.L. (2007) Genotypic and phenotypic characterization of Lactobacillus casei strains isolated from different ecological niches suggests frequent recombination and niche specificity. *Microbiology* **153**, 2655-2665.

Cai, H., Thompson, R., Budinich, M.F., Broadbent, J.R. and Steele, J.L. (2009) Genome sequence and comparative genome analysis of *Lactobacillus casei*: insights into their niche-associated evolution. *Genome Biol Evol* **1**, 239-257.

Chen, P., Zhang, Q., Dang, H., Liu, X., Tian, F., Zhao, J., Chen, Y., Zhang, H. and Chen, W. (2014) Antidiabetic effect of *Lactobacillus casei* CCFM0412 on mice with type 2 diabetes induced by a high-fat diet and streptozotocin. *Nutrition* **30**, 1061-1068.

Christensen, J.E., Dudley, E.G., Pederson, J.A. and Steele, J.L. (1999) Peptidases and amino acid catabolism in lactic acid bacteria. *Ant Leeuwenhoek* **76**, 217-246.

Claesson, M.J., van Sinderen, D. and O'Toole, P.W. (2007) The genus *Lactobacillus*-a genomic basis for understanding its diversity. *FEMS Microbiol Letters* **269**, 22-28.

Collins, Y.F., McSweeney, P.L.H. and Wilkinson, M.G. (2003) Lipolysis and free fatty acid catabolism in cheese: a review of current knowledge. *Int Dairy J* **13**, 841-866.

Curioni, P.M.G. and Bosset, J.O. (2002) Key odorants in various cheese types as determined by gas chromatography-olfactometry. *Int Dairy J* **12**, 959-984.

De Angelis, M., Calasso, M., Di Cagno, R., Siragusa, S., Minervini, F. and Gobbetti, M. (2010) NADP-glutamate dehydrogenase activity in nonstarter lactic acid bacteria: effects of temperature, pH and NaCl on enzyme activity and expression. *J Appl Microbiol* **109**, 1763-1774.

Dellaglio, F., Dicks, L.M.T., Toit, M.D. and Torriani, S. (1991) Designation of ATCC 334 in Place of ATCC 393 (NCDO 161) as the Neotype Strain of *Lactobacillus casei* subsp. casei and Rejection of the Name *Lactobacillus paracasei* (Collins et al., 1989). *Int J Syst Evol Microbiol* **41**, 340-342.

Dellaglio, F., Felis, G.E. and Torriani, S. (2002) The status of the species *Lactobacillus casei* (Orla-Jensen 1916) Hansen and Lessel 1971 and *Lactobacillus paracasei* Collins et al. 1989. Request for an opinion. *Int J Syst Evol Microbiol* **52**, 285-287.

Dicks, L.M.T., Plessis, E.M.D., Dellaglio, F. and Lauer, E. (1996) Reclassification of *Lactobacillus casei* subsp. *casei* ATCC 393 and *Lactobacillus rhamnosus* ATCC 15820 as *Lactobacillus zeae* norn. rev., Designation of ATCC 334 as the Neotype of *L. casei* subsp. *casei,* and Rejection of the Name *Lactobacillus paracasei*. *Int J Syst Evol Microbiol* **46**, 337-340.

Fernández de Palencia, P., Pelaez, C., Romero, C. and Martin-Hernandez, M. (1997) Purification and characterization of the cell wall proteinase of *Lactobacillus casei* subsp. *casei* IFPL 731 isolated from raw goat's milk cheese. *J Agric Food Chem* **45**, 3401-3405.

Frank, J.A., Reich, C.I., Sharma, S., Weisbaum, J.S., Wilson, B.A. and Olsen, G.J. (2008) Critical evaluation of two primers commonly used for amplification of bacterial 16S rRNA genes. *Appl Environ Microbiol* **74**, 2461-2470.

Gaudreau, H., Champagne, C.P. and Jelen, P. (2005) The use of crude cellular extracts of *Lactobacillus delbrueckii* ssp *bulgaricus* 11842 to stimulate growth of a probiotic *Lactobacillus rhamnosus* culture in milk. *Enzyme Microb Technol* **36**, 83-90.

Gobbetti, M., De Angelis, M., Di Cagno, R., Mancini, L. and Fox, P.F. (2015) Pros and cons for using non-starter lactic acid bacteria (NSLAB) as secondary/adjunct starters for cheese ripening. *Trends Food Sci Tech* **45**, 167-178.

Gonzalez, L., Sacristan, N., Arenas, R., Fresno, J.M. and Eugenia Tornadijo, M. (2010) Enzymatic activity of lactic acid bacteria (with antimicrobial properties) isolated from a traditional Spanish cheese. *Food Microbiol* **27**, 592-597.

Hebert, E.M., Mamone, G., Picariello, G., Raya, R.R., Savoy, G., Ferranti, P. and Addeo, F. (2008) Characterization of the pattern of αs1-and β-casein breakdown and release of a bioactive peptide by a cell envelope proteinase from *Lactobacillus delbrueckii* subsp. *lactis* CRL 581. *Appl Environ Microbiol* **74**, 3682-3689.

Herias, M.V., Koninkx, J.F., Vos, J.G., Huis in't Veld, J.H. and van Dijk, J.E. (2005) Probiotic effects of *Lactobacillus casei* on DSS-induced ulcerative colitis in mice. *Int J Food Microbiol* **103**, 143-155.

Iacumin, L., Ginaldi, F., Manzano, M., Anastasi, V., Reale, A., Zotta, T., Rossi, F., Coppola, R. and Comi, G. (2015) High resolution melting analysis (HRM) as a new tool for the identification of species belonging to the *Lactobacillus casei* group and comparison with species-specific PCRs and multiplex PCR. *Food Microbiol* **46**, 357-367.

Ivory, K., Chambers, S.J., Pin, C., Prieto, E., Arques, J.L. and Nicoletti, C. (2008) Oral delivery of *Lactobacillus casei* Shirota modifies allergen-induced immune responses in allergic rhinitis. *Clin Exp Allergy* **38**, 1282-1289.

Jensen, M.P. and Ardo, Y. (2010) Variation in aminopeptidase and aminotransferase activities of six cheese related *Lactobacillus helveticus* strains. *Int Dairy J* **20**, 149-155.

Jung, H.J., Ganesan, P., Lee, S.J. and Kwak, H.S. (2013) Comparative study of flavor in cholesterol-removed Gouda cheese and Gouda cheese during ripening. *J Dairy Sci* **96**, 1972-1983.

Kieronczyk, A., Skeie, S., Langsrud, T., Le Bars, D. and Yvon, M. (2004) The nature of aroma compounds produced in a cheese model by glutamate dehydrogenase positive *Lactobacillus* INF15D depends on its relative aminotransferase activities towards the different amino acids. *Int Dairy J* **14**, 227-235.

Kieronczyk, A., Skeie, S., Langsrud, T. and Yvon, M. (2003) Cooperation between *Lactococcus lactis* and nonstarter lactobacilli in the formation of cheese aroma from amino acids. *Appl Environ Microbiol* **69**, 734-739.

Kok, J., Leenhouts, K.J., Haandrikman, A.J., Ledeboer, A.M. and Venema, G. (1988) Nucleotide sequence of the cell wall proteinase gene of *Streptococcus cremoris* Wg2. *Appl Environ Microbiol* **54**, 231-238.

Laan, H. and Konings, W.N. (1989) Mechanism of Proteinase Release from *Lactococcus lactis* subsp. *cremoris* Wg2. *Appl Environ Microbiol* **55**, 3101-3106.

Macedo, A.C., Vieira, M., Pocas, R. and Malcata, F.X. (2000) Peptide hydrolase system of lactic acid bacteria isolated from Serra da Estrela cheese. *Int Dairy J* **10**, 769-774.

Magboul, A.A.A. and McSweeney, P.L.H. (1999) Purification and characterization of an aminopeptidase from *Lactobacillus curvatus* DPC2024. *Int Dairy J* **9**, 107-116.

Makarova, K., Slesarev, A., Wolf, Y., Sorokin, A., Mirkin, B., Koonin, E., Pavlov, A., Pavlova, N., Karamychev, V., Polouchine, N., Shakhova, V., Grigoriev, I., Lou, Y., Rohksar, D., Lucas, S., Huang, K., Goodstein, D.M., Hawkins, T., Plengvidhya, V., Welker, D., Hughes, J., Goh, Y., Benson, A., Baldwin, K., Lee, J.H., Diaz-Muniz, I., Dosti, B., Smeianov, V., Wechter, W., Barabote, R., Lorca, G., Altermann, E., Barrangou, R., Ganesan, B., Xie, Y., Rawsthorne, H., Tamir, D., Parker, C., Breidt, F., Broadbent, J., Hutkins, R., O'Sullivan, D., Steele, J., Unlu, G., Saier, M., Klaenhammer, T., Richardson, P., Kozyavkin, S., Weimer, B. and Mills, D. (2006) Comparative genomics of the lactic acid bacteria. *Proc Natl Acad Sci U S A* **103**, 15611-15616.

Marilley, L. and Casey, M.G. (2004) Flavours of cheese products: metabolic pathways, analytical tools and identification of producing strains. *Int J Food Microbiol* **90**, 139-159.

McSweeney, P.L.H. and Sousa, M.J. (2000) Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review. *Lait* **80**, 293-324.

Nikolic, M., Tolinacki, M., Fira, D., Golic, N. and Topisirovic, L. (2009) Variation in specificity of the PrtP extracellular proteinases in *Lactococcus lactis* and *Lactobacillus paracasei* subsp. *paracasei*. *Folia Microbiol* **54**, 188-194.

Pedersen, T.B., Ristagno, D., McSweeney, P.L.H., Vogensen, F.K. and Ardo, Y. (2013) Potential impact on cheese flavour of heterofermentative bacteria from starter cultures. *Int Dairy J* **33**, 112-119.

Rijnen, L., Yvon, M., van Kranenburg, R., Courtin, P., Verheul, A., Chambellon, E. and Smit, G. (2003) Lactococcal aminotransferases AraT and BcaT are key enzymes for the formation of aroma compounds from amino acids in cheese. *Int Dairy J* **13**, 805-812.

Sgarbi, E., Lazzi, C., Tabanelli, G., Gatti, M., Neviani, E. and Gardini, F. (2013) Nonstarter lactic acid bacteria volatilomes produced using cheese components. *Journal of dairy science* **96**, 4223-4234.

Sgouras, D., Maragkoudakis, P., Petraki, K., Martinez-Gonzalez, B., Eriotou, E., Michopoulos, S., Kalantzopoulos, G., Tsakalidou, E. and Mentis, A. (2004) In Vitro and In Vivo Inhibition of *Helicobacter pylori* by *Lactobacillus casei* Strain Shirota. *Appl Environ Microbiol* **70**, 518-526.

Simpson, P.J., Stanton, C., Fitzgerald, G.F. and Ross, R.P. (2002) Genomic Diversity within the Genus *Pediococcus* as Revealed by Randomly Amplified Polymorphic DNA PCR and Pulsed-Field Gel Electrophoresis. *Appl Environ Microbiol* **68**, 765-771.

Singh, T., Drake, M. and Cadwallader, K. (2003) Flavor of Cheddar cheese: a chemical and sensory perspective. *Compr Rev Food Sci Food Saf* **2**, 166-189.

Smit, G., Smit, B.A. and Engels, W.J. (2005) Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. *FEMS Microbiol Rev* **29**, 591-610.

Smokvina, T., Wels, M., Polka, J., Chervaux, C., Brisse, S., Boekhorst, J., van Hylckama Vlieg, J.E. and Siezen, R.J. (2013) *Lactobacillus paracasei* comparative genomics: towards species pan-genome definition and exploitation of diversity. *PloS one* **8**, e68731.

Sokal, R.R. and Michener, C.D. (1958) A statistical method for evaluating systematic relationships. *Kans Univ Sci Bull* **28**, 1409-1438.

Stressler, T., Eisele, T., Schlayer, M., Lutz-Wahl, S. and Fischer, L. (2013) Characterization of the recombinant exopeptidases PepX and PepN from *Lactobacillus helveticus* ATCC 12046 important for food protein hydrolysis. *PloS one* **8**, e70055.

Swearingen, P.A., O'Sullivan, D.J. and Warthesen, J.J. (2001) Isolation, characterization, and influence of native, nonstarter lactic acid bacteria on Cheddar cheese quality. *Journal of dairy science* **84**, 50-59.

Tanous, C., Kieronczyk, A., Helinck, S., Chambellon, E. and Yvon, M. (2002) Glutamate dehydrogenase activity: a major criterion for the selection of flavour-producing lactic acid bacteria strains. *Ant Leeuwenhoek* **82**, 271-278.

Thage, B.V., Broe, M.L., Petersen, M.H., Petersen, M.A., Bennedsen, M. and Ardo, Y. (2005) Aroma development in semi-hard reduced-fat cheese inoculated with *Lactobacillus paracasei* strains with different aminotransferase profiles. *Int Dairy J* **15**, 795-805.

Thage, B.V., Houlberg, U. and Ardo, Y. (2004) Amino acid transamination in permeabilised cells of *Lactobacillus helveticus*, *Lb. paracasei* and *Lb. danicus*. *J Dairy Res* **71**, 461-470.

Tindall, B. (2008) The type strain of *Lactobacillus casei* is ATCC 393, ATCC 334 cannot serve as the type because it represents a different taxon, the name *Lactobacillus paracasei* and its subspecies names are not rejected and the revival of the name *'Lactobacillus zeae'*contravenes Rules 51b (1) and (2) of the International Code of Nomenclature of Bacteria. Opinion 82. *Int J Syst Evol Microbiol* **58**, 1764-1765.

Van de Bunt, B., Bron, P.A., Sijtsma, L., de Vos, W.M. and Hugenholtz, J. (2014) Use of non-growing *Lactococcus lactis* cell suspensions for production of volatile metabolites with direct relevance for flavour formation during dairy fermentations. *Microb Cell Fact* **13**, 1-9.

Van den Dool, H. and Kratz, P.D. (1963) A generalization of the retention index system including linear temperature programmed gas—liquid partition chromatography. *J Chromatogr A* **11**, 463-471.

Van Hoorde, K., Van Leuven, I., Dirinck, P., Heyndrickx, M., Coudijzer, K., Vandamme, P. and Huys, G. (2010) Selection, application and monitoring of *Lactobacillus paracasei* strains as adjunct cultures in the production of Gouda-type cheeses. *Int J Food Microbiol* **144**, 226-235.

Vasquez, A., Ahrne, S., Pettersson, B. and Molin, G. (2001) Temporal temperature gradient gel electrophoresis (TTGE) as a tool for identification of *Lactobacillus casei, Lactobacillus paracasei, Lactobacillus zeae* and *Lactobacillus rhamnosus*. *Lett Appl Microbiol* **32**, 215-219.

Vasquez, A., Molin, G., Pettersson, B., Antonsson, M. and Ahrne, S. (2005) DNA-based classification and sequence heterogeneities in the 16S rRNA genes of *Lactobacillus casei/paracasei* and related species. *Syst Appl Microbiol* **28**, 430-441.

Wakai, T., Shinoda, T., Uchida, N., Hattori, M., Nakamura, Y., Beresford, T., Ross, R.P. and Yamamoto, N. (2013) Comparative analysis of proteolytic enzymes need for processing of antihypertensive peptides between *Lactobacillus helveticus* CM4 and DPC4571. *J Biosci Bioeng* **115**, 246-252.

Ward, L.J. and Timmins, M.J. (1999) Differentiation of *Lactobacillus casei, Lactobacillus paracasei* and *Lactobacillus rhamnosus* by polymerase chain reaction. *Lett Appl Microbiol* **29**, 90-92.

Weimer, B., Dias, B., Ummadi, M., Broadbent, J., Brennand, C., Jaegi, J., Johnson, M., Milani, F., Steele, J. and Sisson, D. (1997) Influence of NaCl and pH on intracellular enzymes that influence Cheddar cheese ripening. *Lait* **77**, 383-398.

Williams, A.G., Withers, S.E., Brechany, E.Y. and Banks, J.M. (2006) Glutamate dehydrogenase activity in lactobacilli and the use of glutamate dehydrogenase-producing adjunct *Lactobacillus* spp. cultures in the manufacture of cheddar cheese. *J Appl Microbiol* **101**, 1062-1075.

Yu, S., Peng, Y., Zheng, Y. and Chen, W. (2015) Comparative Genome Analysis of *Lactobacillus casei*: Insights into Genomic Diversification for Niche Expansion. *Indian J Microbiol* **55**, 102-107.

Yvon, M. (2006) Key enzymes for flavour formation by lactic acid bacteria. *Aust J Dairy Technol* **61**, 88-96.

Yvon, M. and Rijnen, L. (2001) Cheese flavour formation by amino acid catabolism. *Int Dairy J* **11**, 185-201.

Figure legends:

Figure 1: (a) Dendrogram of PFGE fingerprints of 98 diverse strains of *L. casei* group generated by BioNumerics® 7.5 software, using UPGMA distance matrix method, and Pearson correlation, and (b) genetic fingerprints of two strains, DPC4206 and DPC4536, with indistinguishable PFGE profiles.

Figure 2: Cell envelope proteinase (CEP) activities of strains of *Lactobacillus casei* group as determined by EnzCheck® kit following incubation at 30°C for 24 h. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test, $P<0.05$. Strains were analysed in triplicate. Error-bars present standard deviation. The graph presents activities of ten representative strains, including the DPC4764 with the lowest activity observed, and *Lactococcus lactis* spp. *cremoris* Wg2, which was used as a positive control.

Figure 3: Heat map of aminopeptidase (PepN, PepC, PepX) activities of strains of *Lactobacillus casei* group as determined by measuring cleavage of corresponding chromogenic substrates (L-Lys-*p*NA, Arg-*p*NA and Gly-Pro-*p*NA) for PepN, PepC and PepX, respectively. Results are expressed as nmol of released *p*-nitroaniline $(min*mg\ of\ protein)^{-1}$. Strains were analysed in triplicate.

Figure 4: Aromatic aminotransferase activities of strains of *Lactobacillus casei* group determined by measuring the absorbance of phenylpyruvate, the final product of transamination between phenylalanine and α-ketoglutarate. Results are expressed as µmol of released phenylpyruvate $(min*mg\ of\ protein)^{-1}$. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test, $P<0.05$. Strains were analysed in triplicate. Error-bars present

standard deviation. The graph presents activities of ten representative strains, inlcuding the DPC5411 and DPC4805 with the highest and the lowest activity observed, respectively.

Figure 5: Glutamate dehydrogenase (GDH) activities of strains of *Lactobacillus casei* group by following change in absorbance during a reaction catalysed by GDH enzyme in which glutamic acid is converted to 2-oxoglutarate in the presence of NAD^+. Results are presented as Units of enzyme activity per mg of protein, where the unit represents the amount of enzyme giving an increase of absorbance of 0.01 per 1 min. Bars sharing the same letter show no significant difference according to least significant difference (LSD) post hoc test, $P<0.05$. Strains were analysed in duplicate. Error-bars present standard deviation. The graph presents activities of ten representative strains, with strain DPC4536 showing the highest, and strain DPC4026 showing the lowest GDH activity.

Figure 6: Abundance, in arbitrary units, of compounds for which significant differences according to least significance test (LSD), $P<0.05$ were observed among the tested strains, including the control, in samples of ten strains of *Lactobacillus casei* group incubated for 48 h at 30 °C in three different media: phenylalanine-enhanced medium (PEM) (a), leucine-enhanced medium (LEM) (b), methionine-enhanced medium (MEM). The control consisted of un- inoculated medium incubated under the same conditions. Strains and the control were tested in triplicate in all three media.

Figure 7: PCA plots of compounds for which significant differences according to least significance test (LSD), ($P<0.05$) produced by ten strains of *L. casei* group incubated for 48 h at 30 °C in three different media: phenylalanine-enhanced medium (PEM) (a), leucine-enhanced medium (LEM) (b), methionine-enhanced medium (MEM) (c) and detected as volatile compounds using HS SPME GC-MS system. The control consisted of un-inoculated medium. Strains and the control were tested in triplicate in all three media.

Supporting Information:

Supporting Information Table 1: Results of the least significant test (LSD) ($P<0.05$) performed after Analysis of Variance (ANOVA), for the results obtained for all analyses strains of *L. casei* group in enzymatic assays: cell envelope proteinase, aminopeptidases PepN, PepC and PepX, aromatic aminotransferase (ArAT) and glutamate dehydrogenase (GDH).

Sheet 1 (CEP): Cell envelope proteinase activities of strains of *Lactobacillus casei* group as determined by EnzCheck® kit following incubation at 30°C for 24 h. Results are expressed as direct flourescence readings. In total, 99 strains were analysed in triplicate. Additionally, strain *Lactococcus lactis* spp. *cremoris* Wg2, and sample of proteinase K were used as positive controls. Strains sharing the same letter show no significant differences in enzyme after performing the Least Significance Test for comparison of mean values ($P<0.05$).

Sheet 2 (PepN): Peptidase N activities of strains of *Lactobacillus casei* group determined by measuring cleavage of corresponding chromogenic substrate (L-Lys-*p*NA). Results are expressed as nmol of released p-nitroaniline per min and mg of protein. In total, 99 strains were analysed in triplicate. Strains sharing the same letter show no significant differences in enzyme after performing the Least Significance Test for comparison of mean values ($P<0.05$).

Sheet 3 (PepC): Peptidase C activities of strains of *Lactobacillus casei* group determined by measuring cleavage of corresponding chromogenic substrate (Arg-pNA). Results are expressed as nmol of released p-nitroaniline per min and mg of protein. In total, 99 strains were analysed in triplicate. Strains sharing the same letter show no significant differences in enzyme after performing the Least Significance Test for comparison of mean values ($P<0.05$).

Sheet 4 (PepX): Peptidase X activities of strains of *Lactobacillus casei* group determined by measuring cleavage of corresponding chromogenic substrate (Gly-Pro-pNA). Results are expressed as nmol of released p-nitroaniline per min and mg of protein.. In total, 99 strains were analysed in triplicate. Strains sharing the same letter show no significant differences in enzyme after performing the Least Significance Test for comparison of mean values ($P<0.05$). 31 out of 39 strains, that had PepX activity below 5 nmol $(min*mg\ protein)^{-1}$ were not carried forward for further analysis of the remaining enzyme activities (ArAT and GDH).

Sheet 5 (ArAT): Aromatic aminotransferase activities of strains of *Lactobacillus casei* group determined by measuring the absorbance of phenylpyruvate, the final product of transamination between phenylalanine and α-ketoglutarate. Results are expressed as µmol of released phenylpyruvate per min and mg of protein. In total, 68 strains were analysed in triplicate. Strains sharing the same letter show no significant differences in enzyme after performing the Least Significance Test for comparison of mean values ($P<0.05$).

Sheet 6 (GDH): Glutamate dehydrogenase activities of strains of *Lactobacillus casei* group determined by measuring the absorbance of phenylpyruvate, the final product of transamination between phenylalanine and α-ketoglutarate. Results are expressed as units (U) of enzyme activity per mg of protein. One unit of GDH activity corresponds to the amount of enzyme that resulted in an increase of absorbance of 0.01 per 1 min. In total, 68 strains were analysed in triplicate. Strains sharing the same letter show no significant differences in enzyme after performing the Least Significance Test for comparison of mean values ($P<0.05$).

Supporting Information Table 2: Cell numbers (log10) of strains of *Lactobacillus casei group* in media with a predominance of a single amino acid (phenylalanine, leucine and methionine, PEM, LEM, MEM, respectively) upon inoculation (t=0 h) and after incubation at 30°C (t=48 h). Strains where significant decreases in cell numbers were observed according to the least significant differences test (LSD) ($P<0.05$) are presented in bold.

Supporting Information Table 3: Compounds generated by strains of *L. casei* group in phenylalanine-enhanced media (PEM), leucine-enhanced media (LEM), and methionine-enhanced media (MEM) after 48 h incubation at 30°C. Compounds were detected by gas-chromatography-mass spectrometry and identified according to their linear retention indices (RI) and by comparison of mass-spectra with National Institute of Science and Technology (NIST) 2011 Mass Spectral Library. Compounds that were detected in significantly different relative abundances in samples, including the control, according to least significant differences (LSD) test ($P<0.05$), are marked with an asterix. Strains and the controls were analysed in triplicate. Controls consisted of un-inoculated media (PEM, LEM, MEM as appropriate). The information on aroma notes were obtained from "The LRI and Odour Database" at www.odour.org.uk, and publications (Curioni and Bosset 2002; Singh et al. 2003; Smit et al. 2005).