International Journal of Food Microbiology 145 (2011) 113–120

Contents lists available at ScienceDirect

International Journal of Food Microbiology

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

Multilocus sequence typing of *Propionibacterium freudenreichii*

Marion Dalmasso [a,b,1], Pierre Nicolas [c,1], Hélène Falentin [a,b], Florence Valence [b,d], Jarna Tanskanen [e], Hanna Jatila [e], Tuomas Salusjärvi [e], Anne Thierry [a,b,*]

[a] INRA, UMR1253 Science et Technologie du Lait et de l'Œuf, F-35042 Rennes, France
[b] AGROCAMPUS OUEST, UMR1253 Science et Technologie du Lait et de l'Œuf, F-35042 Rennes, France
[c] INRA, UR1077 Mathématique Informatique et Génome, F-78350 Jouy-en-Josas, France
[d] CIRM-BIA, INRA, UMR1253 Science et Technologie du Lait et de l'Œuf, F-35042 Rennes, France
[e] Valio Ltd, PO Box 30, FI-00039 VALIO, Helsinki, Finland

ARTICLE INFO

Article history:
Received 31 August 2010
Received in revised form 24 November 2010
Accepted 25 November 2010
Available online 1 December 2010

Keywords:
Multilocus sequence typing
Propionibacterium freudenreichii
Recombination
Mutation
Subspecies

ABSTRACT

Propionibacterium freudenreichii is used as a ripening culture in Swiss cheese manufacture. This study investigates the molecular diversity and the population structure of this bacterium via multilocus sequence typing (MLST). Internal fragments of seven genes sequenced for 113 strains of different subspecies and origins allowed the resolution of 46 sequence types (STs) with occurrence frequencies ranging from 1 to 11. The core genome of the species harbours a low level of nucleotide polymorphism. In our data, single nucleotide polymorphisms account for only 2.28% of the concatenated sequences, and the average polymorphism rate in pairwise comparisons is 0.46%. The analyses reveal quantitatively comparable contributions of recombination and mutation in nucleotide changes at core genome loci along cell lineages. Remarkably, the STs exhibit little if any dairy biotope specialization. Phenotypic characterisation of the strains, based on their aptitude to use lactose and nitrate, shows that the two previously identified subspecies (*freudenreichii* and *shermani*) do not reflect the ancestral relationships in the *P. freudenreichii* population. The considerable phenotypic heterogeneity, found even at the ST level, suggests instead a history of recurrent switches between phenotypes.

1. Introduction

Propionibacterium freudenreichii belongs to the dairy group of propionic acid bacteria included in the *Actinobacteria* class of high G+C content Gram-positive bacteria. The species has mostly been isolated from dairy products and especially from Swiss cheese in which its population density generally reaches 10^9 colony-forming units (cfu)/g. By converting lactic acid into propionic and acetic acids and CO_2, it is the main bacterium involved in the holes and flavor formations in Swiss-type cheeses, like Emmental and Leerdammer© cheeses (Langsrud and Reinbold, 1973). It also contributes to the formation of other aroma compounds in cheese thanks to the conversion of branched-chain amino acids into short branched-chain fatty acids, and to the hydrolysis of milk fat to free fatty acids (Dherbécourt et al., 2010; Thierry et al., 2002; Thierry et al., 2005). A growing attention is paid to the probiotic potential of *P. freudenreichii* such as modulation of the microbiota, enzymatic activities, immune system, inflammation and carcinogenesis, at the gut level (Cousin et al., 2010; Jan et al., 2007). *P. freudenreichii* is divided into two subspecies, *P. freudenreichii* subsp. *freudenreichii* and *P. freudenreichii* subsp. *shermanii*, on the basis of lactose fermentation and nitrate reductase activity (Cummins and Johnson, 1986). Many *P. freudenreichii* activities are strain-dependant, like its survival and metabolic activity in the digestive tract (Lan et al., 2007), its autolytic activity (Lemée et al., 1995) and its production of aroma compounds (Dherbécourt et al., 2008; Thierry et al., 2005). Despite this diversity of reported phenotypes, there is a lack of information concerning the molecular diversity in the species and the relationships between strains that could define its population structure. The few data concerning the molecular diversity of propionibacteria strains are based on fingerprinting characterizations such as Randomly Amplified Polymorphic DNA (RAPD)-PCR and Pulsed-Field Gel Electrophoresis (PFGE) (Gautier et al., 1996; Rehberger, 1993; Rossi et al., 1998). These methods have proven to be useful for the identification of *Propionibacterium* strains in some cases. Nevertheless, RAPD-PCR and PFGE remain clustering techniques that hardly provide information about *P. freudenreichii* nucleotide diversity and mode of evolution.

Multilocus sequence typing (MLST), a method based on partial nucleotide sequences of multiple genes, has emerged as an alternative typing tool to study bacterial populations. It is recognized for its high degree of resolution and its capacity to rapidly characterize large numbers of isolates (Maiden et al., 1998). As MLST is based on DNA sequence analysis, it allows a phylogenetic approach of bacterial population that other typing tools do not permit. Therefore, it has

* Corresponding author. Mailing address: INRA, AGROCAMPUS OUEST, UMR1253 Science et Technologie du Lait et de l'Œuf, 65 rue de Saint-Brieuc, F-35042 Rennes cedex, France. Tel.: +33 2 23 48 53 37; fax: +33 2 23 48 53 50.
E-mail address: anne.thierry@rennes.inra.fr (A. Thierry).
[1] Marion Dalmasso and Pierre Nicolas contributed equally to the manuscript.

0168-1605/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2010.11.037

been largely used for epidemiologic and phylogenetic studies of pathogenic bacteria such as *Neisseria* (Hanage et al., 2005) or *Staphylococcus aureus* (Sakwinska et al., 2009). MLST has more recently been applied to bacteria used in the food industry like *Oenococcus oeni* (de Las Rivas et al., 2004), *Lactobacillus casei* (Cai et al., 2007), *Lactobacillus plantarum* (de las Rivas et al., 2006), and *Streptococcus thermophilus* (Delorme et al., 2010).

The present study was undertaken to investigate the population structure of *P. freudenreichii* by means of the first MLST scheme developed for a member of the *Propionibacterium* genus. The results of our analysis indicate that the core genome of *P. freudenreichii* displays a low level of nucleotide diversity and undergoes frequent recombination events, both characteristics contributing to the species cohesiveness. Our results also show that the lineages of recent common ancestry as defined by the sequence types exhibit broad dairy biotope ranges and phenotypic diversity.

2. Materials and methods

2.1. Bacterial strains, growth conditions and DNA extractions

A total of 113 *P. freudenreichii* strains from CIRM-BIA (Centre International de Ressources Microbiennes — Bactéries d'Intérêt Alimentaire, INRA, Rennes, France), Valio Ltd (Helsinki, Finland) and Actilait (Institut technique du lait et des produits laitiers, Rennes, France) collections were used (Table 1). Their origin is representative of the main biotopes where *P. freudenreichii* strains have been isolated so far: Swiss cheeses (58 strains), other cheeses (25 strains), milk (17 strains), other biotopes (5 strains) and unknown origin (8 strains). Strains originated from ten countries, the best represented being Finland (57 strains) and France (36 strains).

All the strains were routinely cultured in YEL broth (Malik et al., 1968) at 30 °C under microaerophilic conditions for 48 h. Cells were harvested from 2 ml of each culture (8000 g, 10 min, ambient temperature), suspended in 180 µl lysis buffer (lysozyme 20 mg/ml, mutanolysin 50 U/ml in 20 mM Tris–HCl [pH 8], 2 mM EDTA and 1% (v/v) Triton X 100), and incubated for 1 h at 37 °C. DNA was then extracted by means of DNeasy tissues kit (QIAGEN, Courtaboeuf, France).

2.2. Subspecies characterization

All the strains were tested for the presence of a nitrate reductase activity and for lactose fermentation. Nitrate reductase activity was detected by means of the Griess reagent (Biomérieux, Marcy l'Etoile, France) after strain incubation on broth containing nitrate (potassium nitrate 1.5 g/l (VWR International, Fontenay-sous-Bois, France), tryptone 10 g/l, yeast extract 5 g/l, glucose 1 g/l) for 48 h at 30 °C under microaerophilic conditions. Lactose fermentation was tested on modified API 50CH media (tryptone 10 g/l, yeast extract 5 g/l, K_2HPO_4 0.25 g/l, $MnSO_4$ 0.05 g/l, and bromocresol purple 0.17 g/l). One milliliter of a lactose solution (200 g/l, Panreac, Lyon, France) was added to 9 ml of modified API 50CH medium. Strains were incubated at 30 °C for 48 h under microaerophilic conditions. All the tests were carried out in triplicate.

2.3. Loci, primers, PCR conditions and sequencing

The genome sequence of the type strain *P. freudenreichii* CIRM-BIA1^T (CIP103027^T) (Falentin et al., 2010) was used for choosing the genes encoding the following proteins: RNA polymerase β-subunit (*rpoB*), adenylate kinase (*adk*), carboxylic ester hydrolases (*pf169* and *pf1637*), DNA recombinase A (*recA*), fumarate hydratase (*fumC*), cell-wall polysaccharide synthase (*gtf*). Primers were designed with Primer3Plus software (Untergasser et al., 2007) and their sequences are specified in Table 2. PCR reactions were performed in 50 µl containing Taq DNA polymerase buffer (100 mM Tris–HCl [pH 8.8], 500 mM KCl, 0.8% Nonidet P40), 2 mM $MgCl_2$, 200 µM of dNTPs, 0.5 µM of each primer, 0.5 U Taq DNA polymerase (Fermentas, Saint-Rémy-Les-Chevreuses, France) and 2 µl of template DNA (equivalent to 20–50 ng). The PCR amplifications were carried out with an initial denaturation step at 95 °C for 3 min, followed by 40 cycles of denaturation at 95 °C for 30 s, annealing for 30 s at the appropriate primer set temperature (Table 2) and extension at 72 °C for 30 s; and ending with 72 °C for 10 min. For each PCR product, both forward and reverse DNA strands were sequenced using the same primers as for PCR by the sequencing service of AGOWA genomics (Berlin, Germany).

2.4. Description of the sequence polymorphism

Each distinct gene sequence was arbitrarily assigned an allele number or allele type (AT) and each unique combination of the seven gene ATs was assigned a sequence type (ST) (Maiden, 2006). The mean G + C content at the seven loci was calculated. The detection of intra- and inter-loci recombination events and the minimal number of homoplasies were respectively assessed by means of the statistics R_{min} and h using LDhat and DNApars (PHYLIP) (Nicolas et al., 2008). R_{min} is the Hudson and Kaplan's lower bound on the minimal number of recombination events needed to explain the sequence data in absence of homoplasy (Hudson and Kaplan, 1985). The number of apparent homoplasies, h, corresponds to the minimal number of homoplasies that would be needed to explain the data in absence of recombination. The associated statistical test evaluates whether or not h is compatible with the level of polymorphism if no recombination is involved. It is suitable for detecting recombination between fairly related strains with similar sequences, *i.e.* differing by less than 5% (Maynard Smith and Smith, 1998).

2.5. Phylogenetic analysis

Phylogenetic analysis was performed using ClonalFrame software version 1.1 (Didelot and Falush, 2007). ClonalFrame grounds the reconstruction of the genealogical relationship between sequences on a model that accounts for both vertical transmission of the genetic material in cell lineage (the clonal frame) and the occurrence of recombination tracts of variable length. Divergence is generated by mutations as well as and by recombination events that consists of the localized replacement of the sequence by a divergent sequence. In addition to the genealogy of the clonal frame, the parameters of this model are the population mutation rate (θ), the population recombination rate (R), the average length of the recombination tracts (δ) and the fraction of nucleotide divergence in sequences incorporated by recombination (ν). Population mutation and recombination rates correspond to the rates of accumulation of these events in two diverging lineages when time is measured in coalescent units; one coalescent unit of time being the average time to the last common ancestor for a pair of isolates sampled randomly within the population. Inference is carried out in a Bayesian framework using a MCMC (Markov Chain Monte Carlo) algorithm. The number of burn-in cycles of the MCMC algorithm and further iterations were both fixed at a number of 250,000 (five times the default settings). The thinning interval was set to 50, resulting in a posterior sample of size 1000. In this analysis, each ST was represented by a unique sequence and the algorithm was run simultaneously on the 7 loci without concatenating the sequences. Posterior samples were simulated using the default settings where all parameters are estimated or with given values for some parameters other than R. Robustness of the inferred topology for the genealogy with respect to the starting point of the algorithm was verified by comparing the results obtained with three available options for the initial tree: a fully unbalanced tree, a uniformly chosen coalescent tree and the scaled UPGMA tree. The estimated matrix of pairwise-distances was used to draw a phylogenetic tree using