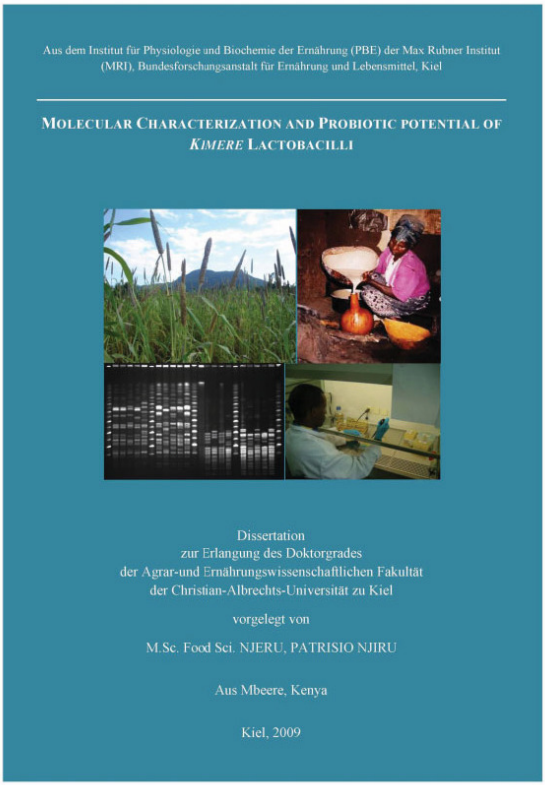
Patrisio Njeru (Autor)

Molecular Characterization and Probiotic Potential of Kimere Lactobacilli



https://cuvillier.de/de/shop/publications/1039

CHAPTER I

MOLECULAR METHODS IN PROBIOTIC RESEARCH

ABSTRACT

In searching for new novel probiotic strains, human and animal gut as well as fermented foods, which are the niches for these organisms are considered. Available evidence indicates that probiotic effects are strain specific. Hence results obtained from efficacy trials from one strain cannot be extrapolated to other strains. Therefore a strain to be applied as a probiotic has to be characterized accurately and independently to strain level, hence the need for accurate methods of characterization. Molecular tools have served this role before and continue to be applied in characterization and identification of probiotics to strain and even genomic level. Molecular methods applied in strain identification and characterization includes: - Amplified ribosomal DNA restriction analysis (ARDRA), ribotyping, pulsed field gel electrophoresis (PFGE), DNA-DNA hybridization, species specific PCR, sequencing etc. Although one method alone is inadequate for identification, it is not feasible to apply all available methods in a single study. Thus, this study employed a polyphasic approach of combining selected molecular methods reviewed below in identification and characterization of *Lactobacillus* strains from *Kimere*.

1 INTRODUCTION

Probiotics are dietary supplements that are beneficial to health and have been defined as "living micro-organisms, which upon ingestion in sufficient numbers exert health benefits beyond inherent basic nutrition" (Schrezenmeir & de Vrese 2001). Lactic acid bacteria (LAB) and especially strains from the *Lactobacillus* and *Bifidobacterium* genus form the largest group utilized as probiotics today. Their usage by man over long time without adverse health effects has acquired them a "generally regarded as safe" (GRAS) status, furthermore research has demonstrated that they possess health promoting properties which are strain specific. They are ubiquitous and found in many places including human gut and fermenting foods (Hammes & Hertel 2006). Their application as probiotics has called for accurate identification and

characterization schemes so as to ensure their safety, identity and correct labeling where they are used in food products.

Earlier characterization of LAB relied on simple microbiological methods based on physical and phenotypic characteristics. These characteristics includes:- morphology; cellular (shape, endospore, flagella, inclusion bodies, Gram staining) and colonial (color, dimensions, form) characteristics, growth requirements and characteristics (temperatures, pH, nutrients, salt concentration), fermentation profiles, configuration of the lactic acid formed and arginine hydrolysis among others (Kandler and Weiss, 1986; Vandamme et al. 1996). Their major drawback is reliance on phenotypic expression which is bound to be influenced by factors such as genotype and environment. They also suffer from limited accuracy, discriminating power, reproducibility and are tedious since several tests are required for characterization even at genus level. Characterization based on these methods to species level could only be considered as tentative pending confirmation with other powerful methods. A good example to weaknesses of classical methods is where bacteria strains with very similar nutritional requirements and growing under similar environmental conditions have erroneously been grouped as same species (Vandamme et al., 1996). This clearly indicates that similar phenotype cannot be equated to similarity. They are also not able to differentiate closely related LAB within a group such as *L. casei* group (Holzapfel et al. 1997). However these methods still remain important for preliminary classification and understanding the strain properties because per se, individual tests are easy and affordable.

The drawbacks of above classical methods have led to emergence of DNA based molecular methods for characterization of LAB (Holzapfel et al. 2001). Molecular methods have greatly increased the quality, reproducibility, accuracy and efficiency of identification of microorganisms and discrimination between closely related species (Singh et al. 2009). Speed and throughput has also been increased with some in-situ methods giving results of microbial composition within hours of sampling without going through cultivation stage. Their application in bacteria characterization has also led to reorganization and reclassification of earlier genus and even species (Dellaglio et al. 2005; Vancanneyt et al. 2006), identification of novel strains (Morita et al. 2007; Mañes-Lázaro et al. 2008) and even understanding the molecular phenotypic mechanisms such as virulence, sugar utilization, immunomodulation and resistance or

adaptability to various environments. Common molecular methods applied in LAB characterization have been reviewed before (Olive & Bean 1999; Holzapfel et al. 2001; McCartney 2002; Ben-Amor et al. 2007; Singh et al. 2009). Singh and co-workers (2009) grouped the molecular methods as PCR-independent and PCR dependent, while McCartney (2002) grouped them as those based on genetic finger prints, gene sequencing, specific primers and oligonucleotide probes. All these classifications overlap at some point and a clear-cut boundary cannot thus be stated. A good example is ARDRA which is a fingerprinting technique that combines PCR amplification and restriction digestion to produce genetic fingerprints which are used as basis of identification and the DNA sequencing method which relies on PCR amplification of the desired region before sequencing. This review compares closely applied techniques without strictly following groupings recommended by earlier reviewers. It is important to note that it is not tenable to apply all existing methods in one study, but selecting a few in a polyphasic approach has led to reliable results in identification and characterization of LAB (Sanchez et al. 2004; Rodas et al. 2005; Weiss et al. 2005).

2 GENETIC FINGER PRINTING METHODS

Finger printing methods are based on restriction fragment length polymorphism (RFLP). DNA is subjected to digestion by restriction endonucleases followed by electrophoresis which generate patterns that act as basis of differentiation (McCartney 2002). These methods differ in their discriminatory power, reproducibility and ease of interpretation (Holzapfel et al. 2001; Amor et al. 2007). The choice of the restriction enzyme is very critical since it determines the number of fragments and the method of their separation. Most applied finger printing methods include: - Amplified rDNA restriction analysis (ARDRA), ribotyping and pulsed field gel electrophoresis (PFGE) (Roy et al. 1996; O'Riordan & Fitzgerald 1999; Ventura & Zink 2002; Ismail 2007; Coudeyras et al. 2008).

Ribotyping combines restriction enzyme analysis (REA) and southern blot techniques, where restriction digestion of chromosomal DNA and agarose electrophoresis is followed by southern blotting with a 23S and 16S rRNA gene probe (Holzapfel et al. 2001). The method is highly reproducible, employs universal probes for all species due to conserved nature of the ribosomal gene and produces easy to interpret patterns. It has been applied before in characterization of various lactobacillus species. Giraffa et al (2000) applied ribotyping in studying diversity of 80

L. helveticus strains, and demonstrated its ability to differentiate to strain level. Miteva and co-workers (2001) were able to identify and differentiate *L. delbrueckii* strains using ribotyping method. Its major drawback is the involvement of southern blot step which is labour intensive (Tynkkynen et al. 1999; Holzapfel et al. 2001; Rodas et al. 2005; Coudeyras et al. 2008).

In ARDRA, the 16S - 23S rDNA gene including the intergenic spacer region is amplified by PCR using a set of universal primers and then subjected to REA. ARDRA has been applied in LAB characterization and has proved an easy, rapid and reproducible method that could be ideal for characterization of unidentified isolates. Using ARDRA, Rodas et al (2003) were able to correctly identify 342 lactobacillus isolates from wine with ease and accuracy. Collado & Hernández (2007) applied the method to identify various strains of *Lactobacillus, Streptococcus* and *Bifidobacterium* from fermented milk samples. It is however limited by inability to characterize below species level and its inability to differentiate very close relatives with 16S rRNA gene sequence relatedness >99%. A few studies have reported inability of ARDRA to differentiate very closely related strains such as *L. plantarum* and *L. pentosus* and also between *L. fermentum* and *L. cellobiosus* (Rodas et al. 2005). An attempt to differentiate *L. delbrueckii* strains by Miteva and co-workers only differentiated *L. delbrueckii* ssp *bulgaricus* from *L. delbrueckii* ssp *delbrueckii* and *L. delbrueckii* ssp *lactis*, which has the same EcoRI restriction map of the 16S rDNA (Miteva et al.2001).

PFGE involves embedding whole cells into molten agarose, subjecting them to in-situ detergent enzyme lysis followed by rare-cutting restriction enzyme digestion and resolving the fragments in agarose gel using alternating electropholetic fields. Its high discriminatory power to strain level has earned it a “gold standard” status for strain differentiation (Olive & Bean 1999), in that it contains whole chromosomal DNA and has a high intra-species discriminatory power. Sawadogo-Lingani and co-workers (2007) demonstrated diversity using PFGE based on sampling sites for *L. fermentum* strains isolated from sorghum beer. Other studies too have applied PFGE in studying diversity of *Lactobacillus* strains (Rodas et al. 2005; Coudeyras et al. 2008; Giraffa & Neviani, 2000) and have found PFGE method to have high intra-species discrimination and able to differentiate at strain level, which outdoes its main disadvantage of being labour intensive and days taken before obtaining results. Faster modes of PFGE are being sought, which could maintain same discriminatory power and give results within shorter time (Briczinski & Roberts 2006). PFGE is still the number one choice for strain differentiation.

3 POLYMERASE CHAIN REACTION (PCR)-BASED METHODS

PCR dependent methods are based on amplification of DNA fragments or specific genes using a pair or a set of oligonucleotide primers. The products obtained can be sequenced to obtain identity or digested using restriction endonucleases and resolving the PCR products in agarose gel electrophoresis and patterns compared with a known type strain. When specific primers are applied, presence or absence (+/-) of amplified target fragment is used as means of identification. Most common PCR methods target the conserved and variable regions within the 16S-23S rRNA gene and other housekeeping genes. They include species-specific PCR (Tilsala-Timisjärvi & Alatossava 1998; Kwon et al. 2004; Dickson et al. 2005; Coton et al. 2008), random amplification of polymorphic DNA-polymerase chain reaction (RAPD-PCR) (Tilsala-Timisjärvi & Alatossava 1998; Sanchez et al. 2004; Rodas et al. 2005; Ruiz et al. 2008; Franciosi et al. 2009), analysis of repetitive genomic elements (rep-PCR) (Gevers et al., 2001; Kostinek et al. 2005; Georgieva et al. 2008 ; Mohammed et al. 2009), Multiplex PCR -restriction enzyme analysis (Lee et al. 2004; Sul et al. 2007) and primers targeting internal *rrn* spacer region and specific genes such as *recA* and *pheS* (Nour 1998; Dubernet et al., 2002; Petrovic et al. 2006; de Angelis et al. 2007; Naser et al. 2007; Rodríguez et al. 2007; Claesson et al, 2008).

Species-specific PCR employs oligonucleotide primers designed from the highly variable regions of the 16S rRNA gene (Drake et al. 1996) to amplify a specific target within a population. Various PCR primers targeting certain species have been designed. Coton and co-workers demonstrated its ability of species-specific PCR to handle large number of isolates with ease and differentiated over 150 heterofermentative *L. brevis, L. fermentum*, and *L. parabuchneri* strains. Several earlier studies have applied species-specific PCR for LAB identification and characterization, and a general consensus is that it is a fast, easy and reliable tool to offer correct identification. This is because the method tests for presence or absence of a certain target species, strain or even gene and mostly does not requires cultivation of the organism (Ward & Timmins 1999; Spano et al. 2002; Dickson et al. 2005). Hence it could be applied when screening large number of samples or as a quality control method for probiotics to ascertain presence/absence of claimed strains and differentiation of strains (Schillinger et al. 2003).

Another commonly used PCR based method is RAPD which is based on short arbitrary primers and low stringency annealing conditions that randomly amplify DNA fragments, which are separated to give a fingerprint. The method has found wide application is characterization of

probiotics and general LAB strains because it is easy and fast (Torriani et al. 2001; Weiss et al. 2005) though it suffers poor reproducibility (Tilsala-Timisjärvi &Alatossava, 1998; Roy et al. 1999; McCartney 2002) and inability to differentiate closely related relatives with high 16S rDNA sequence similarity such as *L. fermentum* and *L. cellobiosus* (Rodas et al. 2005).

4 HYBRIDIZATION METHODS

Hybridization methods employ labeled oligonucleotide probes called "tracers" to hybridize against target sequences within bacterial DNA or DNA fragments. Universal probes are used for larger groupings targeting conserved regions whereas different variable regions are used for genus-specific probes to species-specific probes. It involves heating the double strand DNA at high temperatures to separate the strands into single strands (denaturation) and mixing the target and probe single strands, allowing them to re-associate on cooling (annealing). Perhaps the most widely used hybridization method is the DNA-DNA hybridization, which has been used in measuring the degree of genetic similarity between complete genomes by measuring the amount of heat required to melt the hydrogen bonds between hybrids in a thermal gradient. The melting temperature required to separate hydrogen bond in a hybrid is a function of the number of correctly base-paired nucleotides, thus it is a measure of the degree of genetic similarity between the two single strands forming the hybrid (Vandamme et al. 1999). All hybridization methods use this principle although with various modifications depending on the application. Other examples of hybridization methods used in characterization of LAB includes: - Dot blot hybridization (Giraffa & Neviani 1999), Southern blot (Southern 1975), representational difference analysis (RDA) (Lisitsyn et al, 1993; Hubank and Schatz, 1994; Konstantinov et al. 2005), subtractive hybridization (Sargent and Dawid, 1983), suppression subtractive hybridization (Diatchenko et al, 1996) and Microarrays (Schena et al. 1995).

Dot-blot hybridization technique involves probing DNA extracts from either bacterial isolates or subtracted libraries on a charged nylon membrane and hybridization with a labeled probe. It is a simplified form of the Southern blot (Southern 1975), that eliminates tedious steps of gel electrophoresis and capillary blotting (O'Sullivan, 1999). It is a fast method and has an advantage of testing multiple samples concurrently which is seldom the case in other blotting methods. The other advantage is that the membranes can easily be stripped off the probe and another probe used for hybridization, hence samples can be tested against many probes. It has

been used as an identification method in LAB (Ehrmann et al. 1994; Quere et al. 1997; Giraffa & Neviani 1999) and also as a quality control method for library screening.

Microarray technique involves the extraction of mRNA from the samples of interest, converting to cDNA, labelling, hybridization into microarray chips, and visualizing under laser light to detect the bound fragments. One chip contains thousands of gene sequences and the data is analyzed using available computer programs. Gene expression profiling using microarray is a very useful tool in quantification of large numbers of mRNA and reveals the cellular process at gene level. It has been applied in immunomodulatory effect of probiotic strain *Lactobacillus casei* Shirota (Baken et al, 2006), and transcriptional response of bile-salt resistance in *Lactobacillus plantarum* (Bron et al. 2006; Whitehead et al. 2008). Using microarray analysis Whitehead and co-workers were able to elucidate the mechanism through which a bile resistant *L. reuteri* strain ATCC 55730, copes in the intestines (Whitehead et al. 2008). Microarray methods have an advantage of high throughput in that many genes can be studied on one plate, but are limited in that they are expensive and prior knowledge of sequence is required (Lander 1999). They however do not measure DNA compositions but gene expressions hence are rarely used for characterization of organisms.

Subtractive hybridization involves hybridizing a target sequence called the “tester” against excess mRNA of another sequence called the “driver”, then separation of the unhybridized fractions (target) from hybridized common sequences. Common sequences between the two form cDNA/mRNA hybrids and different sequences from the tester remain as single and double strands and are separated using hydroxylapatite chromatography. The eluted single-stranded cDNA represent genes that are only expressed in the tester, which can then be cloned and sequenced (Sargent and Dawid, 1983). The major drawbacks of the original subtractive hybridization method are its inability to subtract rare transcripts and requirement of high amount of mRNA/cDNA. This original method does not seem to have gained a lot of application in LAB characterization.

Suppression subtractive hybridization (SSH) method by Diatchenko et al (1996) was an improvement of normal subtractive hybridization, combining normalization and suppression effect of PCR in a single cycle. SSH is based on PCR suppression by long inverted terminal repeats (LITRs) that allows selective amplification of molecules that are flanked by different

adapters at opposing termini (asymmetrically flanked molecules) (Gurskaya et al. 1996; Lukyanov et al, 1997). Tester cDNAs are ligated to Adaptors 1 and 2R hence asymmetrical, leaving driver cDNA without primer annealing sites (symmetrical). Furthermore, application of primers from the sequences of the terminal repeats favours the intramolecular annealing of the LITRs over the intermolecular annealing of the shorter PCR primer to these sequences, thus no exponential amplification occurs from these LITR containing sequences. This ensures selective amplification of differentially expressed genes and at the same time suppresses the amplification of abundant genes, hence no need to separate single and double-stranded cDNAs (Diatchenko et al 1996). The SSH procedure (Figure 1-1) is accomplished using 7 steps summarized below:-

1. The tester and driver ds cDNAs are digested with a frequent cutting restriction enzyme to yield blunt ends such as Hae III, RsaI and AluI.
2. Tester cDNA fragments are divided into 2 portions (sample 1 and 2)
3. The portions are ligated with two different adapters-Adapter 1 and adapter 2R resulting to two populations of tester 1 and 2. The ends of the adapters are designed without phosphate groups such that only the longer strand of each adapter can be covalently attached to the 5′-ends of the cDNA
4. First hybridization- excess of driver is added to each tester sample, then heat-denatured and allowed to anneal – products formed (a) ds tester-tester homohybrids, (b) ds tester-driver heterohybrids, (c) ds driver homohybrids, (d) ss driver fraction and (e) ss cDNA tester fraction are normalized.
5. Second hybridization – the two samples from the first hybridization are mixed together which ensures re-association of subtracted ss tester cDNAs forming products (b), (c) and new (f) hybrids-ds tester mixture from adapter 1 and adapter 2 ligated cDNAs. Addition of excess denatured driver at this point leads to enrichment of product (e) for differentially expressed sequences. Product (f) has different adapter sequences at their 5′-ends, one from sample 1 and the other from sample 2, and this allows for their preferential amplification using PCR and primers P1 and Nested primers P1 and 2R which corresponds to the outer part of the adapter 1 and 2 respectively.
6. Primary PCR –primary PCR is performed with only one primer P1. Here, product (a) has long inverted repeats on the ends and forms stable “panhandle –like” structures after each denaturation-annealing PCR step. The panhandle like structures ensure a more favorable

intramolecular annealing than intermolecular annealing hence the suppression PCR effect, half of (tester sequence) product (b) will amplify linearly while the other half (driver sequence) will not amplify since it lacks primer binding sequences, products (c) and (d) are driver sequences which do not contain primer binding sites hence will not be amplified, product (e) is only amplified linearly since they have only one adaptor sequence on one end and only product (f) which contains different adapter sequences at both ends will allows for exponential amplification.

7. Secondary PCR-this is performed with Nested PCR primers P1 and 2R, usually to reduce the background PCR products and further enrich product (f). From here, the PCR products are purified and directly inserted into T/A cloning vector, followed by subsequence screening. Methods such as Southern blot or dot blot hybridization have been used for this purpose.

   The controls involve monitoring steps that are critical to the success of SSH and include efficiency of digestion with restriction enzymes and ligation efficiency. Incomplete digestion and low ligation efficiency tend to reduce the subtraction efficiency (Diatchenko et al 1996; Huang et al. 2007). The main advantages of SSH over other methods are:- that no prior sequence information is required, it is efficient, it requires smaller sample quantities (2µg cDNA), and the process is relatively simple and fast. It is also able to amplify rare transcript and the suppressive effect that eliminates the need to physically separate the single- and double strands.