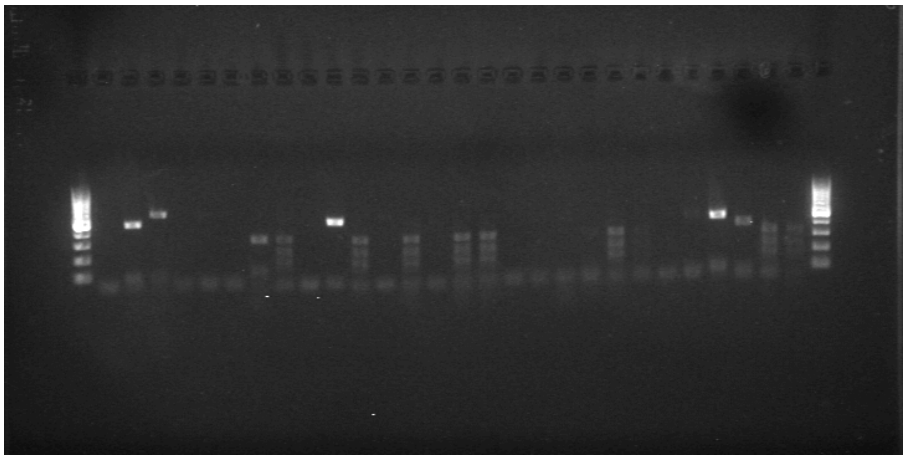
UNIVERSITY OF GOTHENBURG



Identification and characterization of natural yeasts in palm wine

Identifying yeasts in natural West African fermentation

Bachelor's thesis in Biotechnology (15 hc)

LETICIA CASTILLÓN

Department of Chemistry & Molecular Biology
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2018

Bachelor's thesis 2018

Identification and characterization of natural yeasts in palm wine

LETICIA CASTILLON

UNIVERSITY OF GOTHENBURG

Department of Chemistry & Molecular Biology
University of Gothenburg
Gothenburg, Sweden 2018

Identification of yeasts present in natural fermentation
in palm wine from West Africa
Leticia Castillón

Supervisor: Jonas Warringer, Department of Chemistry & Molecular Biology
Examiner: Anders Blomberg, Department of Chemistry & Molecular Biology

Bachelor's Thesis 2018
Department of Chemistry & Molecular Biology
Faculty of Sciences
University of Gothenburg
SE-405 30 Gothenburg
Telephone +46 31 786 0000

Cover: Gel-by analysis of a digestion using the restriction enzyme *Bsu*RI on palm wine samples.

Typeset in LaTeX
Gothenburg, Sweden 2018

Identification of yeasts present in natural fermentation in palm wine from West Africa. Leticia Castillón
Department of Chemistry and Molecular Biology
University of Gothenburg

Abstract

In this project, I identified the yeast species from over 1000 samples originating from Nigerian palm wine, using a combined protocol of PCR amplification, restriction fragment length polymorfism (RFLP) and multiplex PCR.

Most of the identified samples correspond to *Saccharomyces cerevisiae*, usually main yeast participating in fermentation processes, which makes sense with the origin of the studied material. However, for some other results different data was obtained, which could correspond to different *Saccharomyces* sp. as *Saccharomyces mikatae*, *Saccharomyces arboricola* or *Saccharomyces paradoxus*, species that had not been previously identified in Africa.

These unexpected results should be contrasted with the data obtained from the phenotyping analysis and, since the only way to unambiguously identify a species is by sequencing its DNA, the sequencing of the ITS region of a subset of the strains that seem to belong to one of the species mentioned above would be mandatory to confirm the finding.

Keywords: PCR, RFLP, elctrophoresis, yeasts, *Saccharomyces*, palm wine, Nigeria.

Contents

List of Figures

List of Tables

Acronyms and abbreviations

AAB	**Acetic acid bacteria**
ATP	**Adenosine triphosphate**
bp	**base pairs**
CO_2	**Carbon dioxide**
DNA	**Desoxiribonucleic acid**
EtBr	**Ethidium bromide**
H_2O	**Water, oxidane**
ITS	**Internal Transcribed Spacer**
LAB	**Lactic acid bacteria**
NaOH	**Sodium hydroxide**
PCR	**Polymerase chain reaction**
RFLP	**Restriction fragment length polymorphism**
RNA	**Ribonucleic acid**
TAE	**Tris base - acetic acid - EDTA**
UV	**Ultra-violet**
w/v	**Weight/Volume percentage concentration**
YPDA	**Yeast extract - Peptone - Dextrose - Adenine**

1
Introduction

1.1 Yeasts

Yeasts are defined as a group of eukaryotic unicellular fungi. They are encountered in many natural habitats, as plants, soil and salt water, especially in those rich in sugar (since their main source of energy come from fermenting carbohydrates); but they can also be found on the skin or intestines of warm-blood organisms, either as a symbiont or parasite, and some species can be the cause of infection in immunocompromised humans.

About 1500 yeast species have been identified so far, which have been classified in two groups: conventional (*Saccharomyces cerevisiae* and close related species, traditionally used for fermentation and bake) and non-conventional yeasts (the rest). In the genus *Saccharomyces* spp. seven different species are encompassed: *Saccharomyces arboricola*, *Saccharomyces bayanus*, *Saccharomyces cariocanus*, *Saccharomyces cerevisiae*, *Saccharomyces kudriavzevii*, *Saccharomyces mikatae* and *Saccharomyces paradoxus*. All these species have the same mating type system and thus they are able to cross and produce hybrids.[12]

Regarding classification, yeasts are encompassed within the fungi kingdom, and are separated in two phylum: ascomycetes (mostly yeasts and molds, with around 65,000 species described so far) and basidiomycetes (around 30,000 species; mainly mushrooms, but also some yeasts are encompassed within this category (in the *Pucciniomycotina* subphyla)) [3] [2]. This classification is made mostly regarding how the spores are formed.

Humans have been using yeasts since several millennial back for baking, brewing and fermentation of many other beverages.
However, yeasts were not identified as the microorganisms carrying out these processes until the studies of Louis Pasteur during the 19th century [1] [6]. Since then, yeasts have been in the forefront of the development of bio-processes (such as industrial fermentation or production of recombinant proteins) until it reached modern biotechnological processes nowadays, being the subject of important scientific achievements: for example the first eukaryote genome that was sequenced was *Saccharomyces cerevisiae* genome [13], completed in 1998.

Currently yeasts, specially *Saccharomyces cerevisiae* [15], [16], [18], represent a pow-

erful tool in Biotechnology for several reasons: their size makes them suitable for an easy manipulation in laboratory, they are easy to cultivate and store, easy to transform due to its relative simplicity, etc. In addition to this, most species are non-pathogenic to both humans and animals (an outstanding exception would be the *Candida* species, able to cause infections in humans[4]).

There are some conventional yeasts considered key in industry: *Saccharomyces cerevisiae*, *Saccharomyces pastorianus*, *Saccharomyces kudriavzevii* and *Saccharomyces bayanus* [15], although some non-conventional yeasts have also proven an important role in industrial fermentation processes, suchs as *Yarrowia lipolitica* or *Kluyveromyces lactivs* [16] but also in industry go from large-scale production (first molecule produced was citric acid in 1920) of different products. The possibilities of this application broaden with apparition of genetic engineering in the 70s, enabling the production of heterologous proteins, vitamins, etc [14], production of biofuels (mainly bioethanol using *S. cerevisiae* [24]). In addition to this, yeast represent a strong tool in research of systems biology or molecular mechanisms shared with humans (for example mitotic recombination [17]) since their genetics make-up share significant similarities with humans.

The extensive usage of yeasts by humans has provided us with large knowledge about their physiology, genetics, molecular biology, etc. But it has also had an impact on their evolution (and so in their genomes and traits), since an important number of yeasts have evolved in what we could call "human-induced" environments and so have been separated from populations of natural yeasts. In other words, they have been domesticated (in case of yeasts used for wine fermentation a characteristic of this domestication could be the development of resistance to copper or sulfite) [7] [8].
The domestication of these strains has lead to a loss of genetic divergence, specially within the genus *Saccharomyces*, since it has been strongly used in different human activities such as bakery. For this reason, there is an interest in creating new hybrids (either between two domesticated strains or between an artificial strain and a natural one) that would increase this genetic variation and thus produce strains with improved characteristics or more suitable to face new environments. [9]
Although there is a strong interest in producing these hybrids, hybridization can occur in a natural way. However, it is believed that is a process happening seldom in nature and that it is selected in production environments [11] [10].

Another factor affecting genetic variation of yeasts is their way of reproduction: they can reproduce either asexually (trough budding or fission) or sexually, depending on their need of adaptation to the environment (since sexual reproduction introduces genetic variation to a population), which also has a huge impact on their evolutionary profile.

1.2 *Saccharomyces cerevisiae*

S. cerevisiae is the most studied yeast so far, since it plays a key role in different industrial and food related processes (such as bakery or brewing), which makes it the principal yeast utilized in biotechnology.

At the same time, it is the main model eukaryotic organism in basic research and its genetic material was the first eukaryotic genome to be sequenced. Because of this, it is also one of the most suitable organisms whose genome or metabolism we can try to modify in order to optimize different industrial processes or produce molecules of interest.

S. cerevisiae and closely related yeasts (conventional yeasts) usually grow in environments with a high sugar concentrations. The reason of this is that *S. cerevisiae* is specialized to these sugar enriched niches by having a wide range of sugar importers, and a metabolism that allows very fast, but quite inefficient, growth on sugar catabolites. Sugar is initially fermented into ethanol even in the presence of oxygen. The ability to use simultaneously respiration (aerobic) and fermentation (which courses with a lower yield of ATP but does not require oxygen) is known as the Crabtree effect and the yeasts that present this particularity are known as Crabtree positive yeasts.[22]
This oxygen-present fermentation, that can occur when the levels of sugar are high enough (it occurs in presence of high levels of glucose, fructose, manose and galactose [23]), is intentional since it allows fast growth and production of a toxin, ethanol, to which *S. cerevisiae* is more resistant than any other yeast. Once the competition is killed off, the sugar can be respired and efficiently converted into stored energy [28]. However, most environments do not provide so high sugar concentrated conditions, and it is more frequent to find non-conventional yeasts, that are present in greater number in this habitats, leading early stages of fermentation, although they have lower fermentation rates [27] [19]

1.3 Palm wine and Nigeria

In West Africa, palm wine is an alcoholic beverage that constitutes a traditional drink consumed by around 10 million people in different regions. The beverage plays an important role in different ceremonies among certain African regions (such as Nigeria), as well as an important role in the local economy. [32]

In Africa, the wine can be obtained from several wild date palms species. However, in Nigeria specifically, this kind of wine is obtained through fermentation of the sugary sap (with a lot of simple sugars in solution) produced by the male inflorescence (group of flowers clustered on a stem that diverges from the main trunk) from the oil-pal tree and Raphia palms. [33]
To simplify the process: the palms secrete the sap. A tapper collects this sap (usually cutting the flower) and the collected liquid is placed into a recipient where the

fermentation begins immediately due to the natural presence of wild yeasts in the trees (they are present in great amounts around the flowers and so they inoculate the sap naturally). Thus, the process is mostly natural and has not been industrialized yet (palm trees are currently cultivated mostly in family farms and plantations do not exist).

The amount of wine produced is seasonal, the highest quantity being produced during March/April and October/November corresponding to the first period with the highest rate of development of the tree and the latter to the highest amount of flowers that produce the sap. [36]

Sugar composition of the palm wine varies through the fermentation process. The palm sap contains an initial sugar concentration of 14% w/v, mainly sucrose with smaller percentages of glucose and fructose. This sugar concentration drops during fermentation to about 11% or 8%, depending on the duration of the fermentation.[26]. The pH of the sap is close to neutral (7-7,4). It does not contain ethanol at the onset of the fermentation. [38].

During the initial stages of tapping, in Raphia palm wine, sucrose, maltose, glucose and fructose were identified; whereas xylose and cellobiose and raffinose were detected in later stages. [38]

We can classify these sugars attending to several criteria. On one hand, based on the number of sugar monomers we have monosaccharydes (glucose, fructose and xylose), disaccharydes (sucrose (glucose and fructose), maltose (two molecules of glucose bound by a α (1->4) bond), cellobiose (two molecules of glucose bound by a
β(1->4) bond) or trisaccharydes (raffinose (galactose, fructose and glucose)).
We can also classify them as hexoses (glucose), or pentoses (xylose, fructose).
During the fermentation process, the pH drops to 3.5, due to several factors as proton secretion during transport of molecules, production of organic acids during Krebs Cycle (part of the aerobic metabolism of *S. cerevisiae*), such as succinic acid or acetic acid„ the removal of buffering acids and dissolution of CO_2. [34] [35]

The concentration of alcohol (ethanol) can vary depending on the stage of the fermentation, but the range is usually between 3-6% . Fermentation stops not because the consumption of sugar (if the medium runs out of sugar yeast would start to obtain energy from ethanol), but because exhaustion of nitrogen, which is a growth limiting nutrient and necessary for biosynthesis of enzymes and other proteins necessary for fermentation. Thus, there is still sugar in palm wine once fermentation has stopped.
Regarding this, if palm wine is in the future industrially produced, the supplement of nitrogen would prolong fermentation so the percentage of ethanol innit would rise, allowing the commercialization of a stronger wine. [35]

1.4 Fermentation

Fermentation is an anaerobic pathway used by many organisms to obtain ATP (thus, energy) in absence of oxygen. Fermentation uses glycolysis as energy extraction pathway, producing pyruvate from glucose (sugar). From here, through the alcoholic fermentation pathway (the one happening during wine fermentation, carried out by yeasts), pyruvate is transformed into ethanol (alcohol) and CO_2. [21]

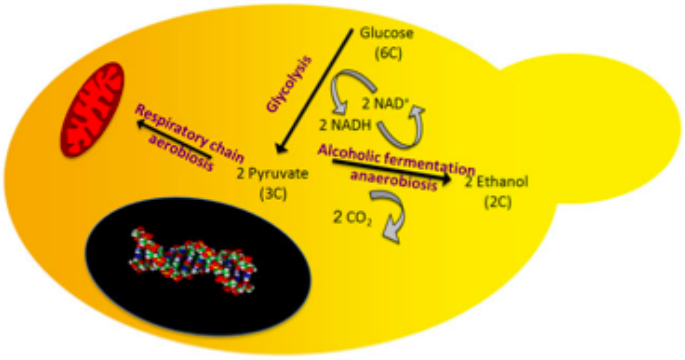


Figure 1.1: Alternative metabolic routes following glycolysis.
A budding yeast cell is shown with the aerobic and anaerobic metabolic pathways following glycolysis. The nucleus (black) and mitochondrion (red) are also shown. 2010 Nature Education. All rights reserved.

Palm wine complete fermentation consists of 3 stages, or 3 different types of fermentation: lactic, alcoholic and acetic, the two former ones involved in wine production. All together these fermentation processes explain the changes in the biochemical composition from the sap to the fermented palm wine.
The complex fermentation process also explains the diverse community of microorganisms that we can find in palm wine sap: besides yeasts (from which *S. cerevisiae* plays the major role in palm wine fermentation according to studies performed on *Elaeis guineensis* in Ghana [38]), comunities of lactic acid bacteria (LAB), such as *Lactobacillus plantarum* or *Leuconostoc mesenteroides*, and acetic acid bacteria (AAB) have been identified [38]. AAB, however, are only present when the wine is no longer consumable. [38]

As yeasts continue growing and metabolizing sugar (while it is available in the media) the concentration of alcohol keeps growing and eventually the high concentration kills the cells. In general, yeast strains tolerate an alcohol concentration within the range of 10-15% of alcohol. However, *S. cerevisiae* has been observed to tolerate higher levels of alcohol without serious effects to its viability (tolerance at 18% and up to 25% v/v concentration of ethanol) [24][25], which provides it an important advantage during fermentation in comparison to less tolerant strains.
The whole process usually takes 48 hours, and it ends once the pH falls to 4.0 due

to the organic acids production by yeasts during fermentation (for example lactic acid). However, fermentation can be affected by seasonal variations that affect palm wine constituents and other factors such as type of soil, time of tapping, etc

1.5 Microorganisms identified in palm wine

Microorganisms	Bandji[1]	Taberna[2]	Palm wine from Cameroon[3]	Palm wine from Ghana[4]	Palm wine from Nigeria[5]	Toddy[6]	Toddy[7]
Saccharomyces cerevisiae	x		x	x	x		x
Saccharomyces ludwigii			x			x	
Saccharomyces bayanus					x		
Saccharomyces uvarum					x		
Saccharomyces bailii						x	
Saccharomyces chevalieri						x	
Candida tropicalis	x				x	x	
Candida pararugosa	x						
Candida quercitrusa	x						
Candida parapsilopsis			x			x	
Candida fermentati			x				
Candida krusei				x			
Candida utilis					x		
Candida guilliermondii						x	
Candida valida						x	
Pichia etchellsii						x	
Pichia farinosa						x	
Pichia membranaefaciens						x	
Pichia ohmeri						x	
Pichia guilliermondii						x	
Pichia fermentans			x				
Zygosaccharomyces bailii			x				
Schizosaccharomyces pombe	x				x	x	x
Issatchenkia orientalis	x						
Arthroascus fermentans	x						
Trichosporon asahii	x						
Hanseniaspora uvarum	x		x				
Kodamaea ohmeri	x						
Trichosporom asteroides	x						
Trigonopsis variabilis	x						
Galactomyces geotrichum	x						
Kloeckera apiculata				x			
Kloeckera javanica						x	
Rhodotorula glutinis						x	
Kluyveromyces lactis					x		
Lactobacillus plantarum	x			x			
Lactobacillus fermentum	x						
Lactobacillus paracasei	x						
Lactobacillus nagelii	x	x					
Lactobacillus sucicola		x					
Lactobacillus sp		x				x	
Leuconostoc mesenteroides	x			x			
Leuconostoc dextranicum							x
Leuconostoc sp						x	
Fructobacillus durionis	x	x					
Fructobacillus fructosus		x					
Streptococcmitis	x						
Acetobacter indonesiensis	x						
Acetobacter tropicalis	x						
Acetobacter estunensis	x						
Acetobacter ghanensis	x						
Acetobacter aceti	x					x	x
Acetobacter lovaniensis	x						
Acetobacter orientalis	x						
Acetobacter pasteurianus	x	x					
Acetobacter cerevisiae	x						
Acetobacter rancens							x
Acetobacter suboxydans							x
Acetobacter sp				x			
Gluconobacter oxydans	x						
Gluconobacter saccharivorans	x						
Gluconobacter sp				x			
Zymomonas mobilis		x					

1: Ouoba et al., 2012; 2: Alcántara-Hernández et al., 2010; 3: Stringini et al., 2009; 4: Amoa-Awua et al., 2007; 5: Ezeronye and Ekerentugba, 2001; 6: Atputharajah et al., 1986; 7: Shamala and Sreekantiah, 1988. x: Identified microorganisms. Surrounded in yellow microorganisms found in palm wine from Nigeria

Table in the previous page summarizes the different microorganisms identified so far in palm wine from different regions from Africa.[38]

Strains identified in palm wine from Nigeria have been surrounded with a yellow rectangle. Seven different species were identified, belonging to *Saccharomyces* spp., *Candida* spp and some to *Schizosaccharomyces* and *Kluyveromyces* spp. [31] These identifications were performed through mating and hybridization studies, genetic screening (searching for presence or absence of genes such as rho (structural protein present in yeasts' mitochondria playing an important role in respiratory capacity [39])) or if they are able to use ammonium as a nitrate source; and growth studies in which ethanol production and tolerance were measured.

2

Aim of the project

The aim of the project is to identify and characterize the microorganisms present in the samples extracted from palm wine coming from different regions of Nigeria, using molecular techniques such as PCR and RFLP.
Of special interest is the identification of *Saccharomyces cerevisiae* due to its importance in fermentation of high sugar substrates into ethanol, and thus try to gain a better understanding of the process of palm wine fermentation and provide data that could help to a future optimization of this fermentation; but also to detect putative species that are not expected to be found in this source or area.

3

Materials and Methods

3.1 Sample extraction

The samples of yeasts were obtained from different regions of Nigeria.

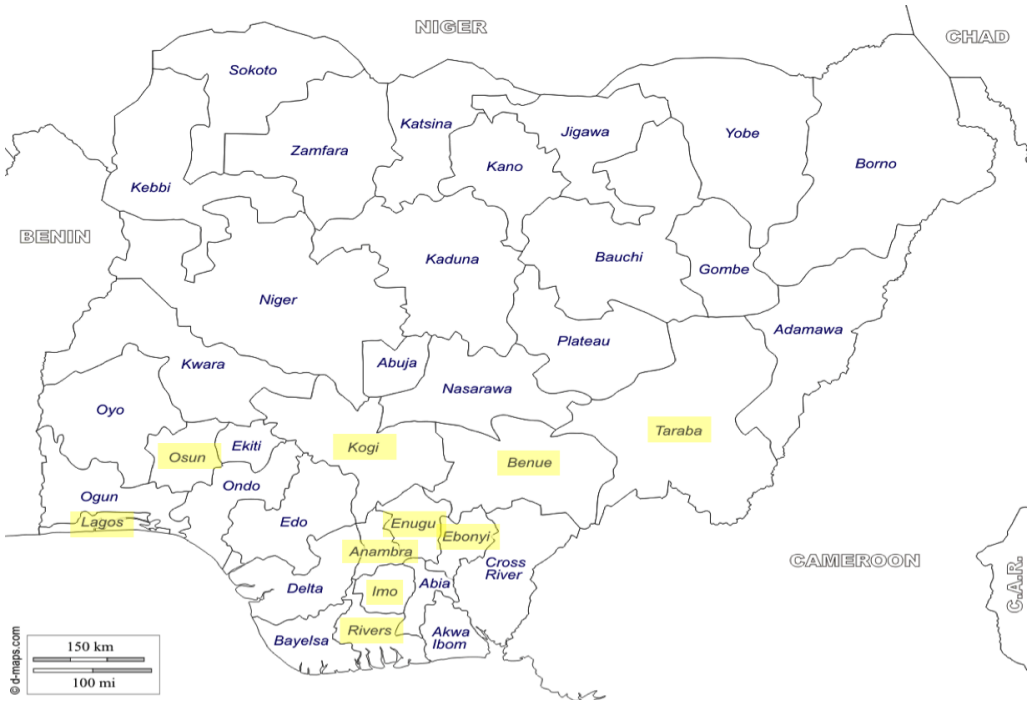


Figure 3.1: Map of Nigeria. The highlighted names correspond to the states from where the strains were isolated.

The yeasts samples were obtained from two different species of Raffia tree: *Raphia hookeri* and *Raphia vinifera* and one species of oil palm tree: *Elaensis guineensis*. To obtain the samples the help of palm wine tappers was necessary. The tapping process lasted within 30-60 minutes and the sap was recollected in sterile 500 ml screw-capped bottles immediately after the tapping. Once samples are collected in bottles the fermentation initiates. The bottles are aerated. If they were not isolated immediately, the samples were stored at 4°C.

After 2-48 hours (depending on the source of the sample), yeasts were cultured in YPDA medium plates with the following composition: 1% (w/v) yeast extract, 2%

(w/v) peptone, and 2% (w/v) glucose, 2% (w/v), and 20 mg/ml of chloramphenicol (antibiotic) and incubated at 28°C for 24-48 hours. Since yeast have a fast growth the incubation lasted 24 h. for most of them.
Once the growth was observed the strains were stored at 4°C, and kept at this temperature during the experiment. Since the optimal temperature for fermentation is around 25°C (it is strongly strain dependent), the storage at such low temperature inhibits the growth of the strains significantly and thus, their metabolism, so the level of fermentation carried out at these conditions is negligible. [20]

Table 3.1: Number of samples from the different regions and total.

Region	Number of samples
Anambara	74
Benue	90
Ebonyi	70
Enugu	156
Imo	68
Kogi	45
Lagos	79
Osun	56
Rivers	77
Taraba	124
`Total`	842

3.2 Identification of yeasts using RFLP of ITS region.

Identification of yeasts in palm wine has been usually performed using poor identification techniques: traditional morphological and physiological tests (Gram staining, catalase test, gas production and growth in differential culture media), or trying to determine the pattern of fermentation. Unfortunately, these methods only work properly on easily culturable microorganisms and they are prone to inaccuracy.
Here we use a molecular approach based on PCR amplification and RFLP analysis of the ITS regions and 5.8S rRNA gene. Altogether, we expect that the size of the PCR product and the RFLP pattern lead to a good profile for each yeast species (as already done in previous work [29]), which permit to give a reasonably accurate identity to each sample. Such as techniques have already been used to identify fungus phylogenetic relationships and even in yeast identification in palm wine from Cameroon [40], but never applied to microorganism identification in palm wine from Nigeria (the only techniques used to study this have been previously described in the section 1.5).
Finally, we chose among those whose profile matched with the expected for *Saccharomyces* spp. and performed a PCR multiplex on them, searching for a clear differentiation of *S. cerevisiae* from the rest of the *Saccharomyces* species [41, 43]

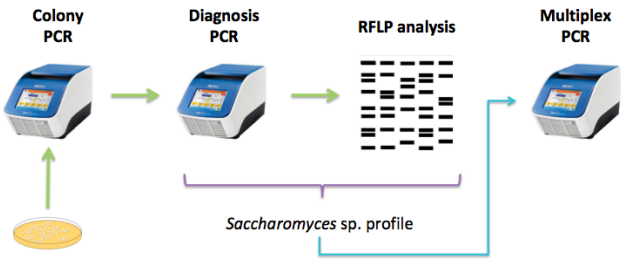


Figure 3.2: Scheme of the yeasts identification process

3.3 PCR reactions

The aim of the PCR reaction is to amplify a particular region comprising the well-conserved gene 5.8S rRNA and two other highly variable structures (ITS1 and ITS2) that have been proven to be useful to classify fungus genealogical relationships.

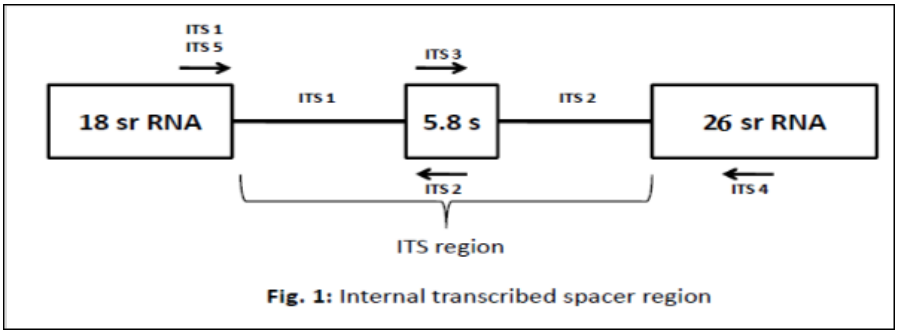


Figure 3.3: Representation of the 5.8S rRNA and ITS region of interest for the amplification.
Although the image represents the mentioned genetic region in plants, we have chosen it because of its simplicity and because the distribution of the genes in yeasts for this region is the same.[37]

In order to do this, first we need enough DNA material to serve as template for this reaction. To achieve this we carried out a colony PCR followed by a diagnosis PCR.

3.3.1 Colony PCR

Colony PCR is a fast technique for DNA amplification. By using it we aimed to obtain enough DNA to act as a template in the posterior diagnosis PCR, avoiding DNA purification steps.

The PCR was performed obtaining the genetic material directly from the colonies growing in the YPD media plates. Using toothpicks, we poked carefully in the colony and add it to the PCR tubes. In the PCR tubes we had previously dispensed 50 μL of 10 mM NaOH to break down the cell walls and have access to the DNA. We were

careful to not add too many yeasts material so we would not introduce too many inhibitors that would prevent the PCR reaction to occur; but making sure that the solution was cloudy so we have enough cells in it.

After the incubation in the PCR machine, we added glass beads to the tubes and vortex it for 15" to lyse the cells and ensure the access of the PCR components to the DNA. Finally we span it down to get the genetic material in the supernatant.

The protocol designed for the DNA extraction using colony PCR consists in the following thermal cycler: Stage 1 (x1) heating at 95°C during 1', stage 2 (x1) is divided in 2 steps. Step 1: 99°C during 20'; step 2 cooling at 4°C until we retrieve the samples from the machine.

3.3.2 Diagnosis PCR

The diagnosis PCR aimed to amplify the region we wanted to analyze (5.8S rRNA and ITS regions).

The diagnosis PCR was performed both in the samples obtained from Nigeria and in typed cultured strains (known *Saccharomyces cerevisiae* strains). This way we obtained positive controls of what the PCR amplification should look like for the strains we are interested in.
The PCR components were mixed together preparing a master mix containing the following:

Receipe for 25 μL PCR mix	Volume (μ)
dH_2O	17.3
Phusion HF buffer (5x)	5
dNTPs (10mM)	0.5
Primer F (10μM)	0.5
Primer R (10μM)	0.5
Template (2ng/μL)	1
Phusion enzyme (2U/μL)	0.2
TOTAL	25

Table 3.2: Composition of the PCR mix for the diangnosis PCR

We dispensed 24l of the mix in PCR tubes and added here 1μL from the colony PCR (described in the section 3.3.1).

Regarding the polymerases used to perform the reaction, we used two different ones: Phusion DNA polymerase (2 U/μL) and Phusion hot-start II DNA polymerase (2 U/μL). Using the last one we managed to avoid the start of the reaction at room temperature.
When the phusion enzyme was used we kept the mix on ice in order to prevent the reaction to start before the tubes were placed in the PCR machine.

Regarding the primers, we used ITS1 (5'-TCCGTAGGTGAACCTGCGG-3') as forward primer and ITS4 (5'-TCCTCCGCTTATTGATATGC-3') as reverse primer.

The PCR protocol was set up trying two different annealing temperatures: the one that proved to be more suitable for this reaction was 60°C. The temperature of 62°C gave weak bands, and it was rejected.

The run method for the PCR amplification consisted of 3 stages corresponding to initial denaturation, several cycles of denaturation-annealing-extension and a final extension:

An initial denaturation (x1) 98°C for 30 seconds, 35x [98°C for 10 seconds, 60°C for 30 seconds and 72°C for 1 min], followed by a final extension step of 72°C for 5 min.

3.3.3 Multiplex PCR

In order to differentiate among the putative *Saccharomyces* spp. strains, we performed a multiplex PCR using the primers designed by *Muir et al.*[41]

This protocol amplifies different genes for each species. A total of six genes were chosen, all of them ubiquitous and orthologues for the different species, but with differences in their sequence for the species they were designed for (for example, amino acids deletions or insertions) [41] [44].

For this PCR we selected from the samples those ones that had given a profile likely to belong to *Saccharomyces* species (having a PCR amplicon equal or above 650 bp).

The multiplex PCR is specially important to distinguish between *S. cerevisiae*, *S. bayanus* and *S. pastorianus* because the length of their PCR product is exactly the same (880) and their RFLP profile (see Results, section 4.2) is quite similar except for the size of the fragment produced by *Bsu*RI, which is 320 for *S. cerevisiae* and 500 for *S. bayanus* and *S. pastorianus*.
However, the band pattern that is expected to result from the multiplex for these species is unique for each of them, therefore allowing their differentiation.

We also performed the multiplex PCR to *Saccharomyces cerevisiae* typed strains (strains that were already known to belong to a particular species) to have a positive control to compare our results with.

Spacers	Primer	Primer segments (5' -> 3'	Amplicon size (bp)
Saccharomyces arboricolus	Sarb F1	GGC ACG CCC TTA CAG CAG CAA	349
	Sarb R2	TCG TCG TAC AGA TGC TGG TAG GGC	
Saccharomyces bayanus	Sbay F1	GCT GAC TGC TGC TGC TGC CCC CG	275
	Sbay R1	TGT TAT GAG TAC TTG GTT TGT CG	
Saccharomyces cerevisiae	Scer F2	GCG CTT TAC ATT CAG ATC CCG AG	150
	Scer R2	TAA GTT GGT TGT CAG CAA GAT TG	
Saccharomyces kudriavzevii	Skud F2	ATC TAT AAC AAA CCG CCA AGG GAG	660
	Skud R1	CGT AAC CTA CCT ATA TGA GGG CCT	
Saccharomyces mikatae	Smik F1	ACA AGC AAT TGA TTT GAG GAA AAG	508
	Smik R1	CCA GTC TTC TTT GTC AAC GTT G	
Saccharomyces paradoxus	Spar F7	CTT TCT ACC CCT TCT CCA TGT TGG	739
	Spar R7	CAA TTT CAG GGC GTT GTC CAA CAG	

Table 3.3: Species specific primers used [41]

The primers needed for the multiplex PCR were lyophilized. First we want to make a 100X stock solution by adding to the primers 10 times the amount of water than the concentration indicated in the side of the tube.
After this we have to make another dilution 1:10 so the primers have the adequate concentration to be used in the reaction.
All the primers are mixed up together with the rest of the components of the PCR reaction, the same as used for the PCR amplification (described in the section 3.3.2).

Recipe 25µL PCR	Volume (μ)
dH_2O	11.3
Phusion HF buffer (5x)	5
dNTPs (10 mM)	0.5
Primer F (10 μM) (x6)	0.5
Primer R (10 μM) (x6)	0.5
Template (2ng/μL)	1
Phusion enzyme (2U/μL)	0.2
`Total`	25

Table 3.4: Composition of the PCR mix for the multiplex PCR.

In this case we have 6 forward primers and 6 reverse primers so in total we added 6μL of primers for each sample. To maintain the volume of the mix, we reduce the amount of distilled water.
From here, the protocol follows the thermal cycle described in section 3.3.2.

3.4 RFLP analysis

The restriction fragment length polymorphism analysis consisted in this particular case of the digestion of the genetic material using three different restriction enzymes: *Hha*I, *Bsu*RI and *Hinf*I.

*Hha*I is a restriction endonuclease that recognizes GCG|C sites.
*Bsu*RI recognizes GG|CC sites.
*Hinf*I recognized G|ANTC sites.

The amplified DNA resulting from the diagnosis PCR (see section 3.3.2) were digested without further purification.
The digestions were performed in the same way using the three different restriction enzymes: in an eppendorf tube we placed 0,34 μL of enzyme, 0,66 μL of buffer (already provided by the commercial house together with the enzyme; the buffer provides the optimal conditions for the enzyme to cut) and 5,66 μL of H_2O. The DNA coming from the PCR amplification was added to this mix (3,34 μL), completing the 10 μL that were set in the digestion protocol. The digestion protocol consisted in 15 seconds incubation at 37°C.

The pattern these enzymes are expected to produce for the species that have been identified so far in palm wine from Nigeria are the following:

Species	*Hha*I	*Bsu*RI	*Hinf*I
S. cerevisiae	385+365	320+230+180+150	365+155
S. bayanus	385+365	500+220+145	365+155
S. pastorianus	385+365	500+220+145	365+155
C. tropicalis	280+250	450+90	270+270
S. pombe	600+400	1050	600+450
K. lactis	285+190+165+90	655+80	290+180+120+80+65

Table 3.5: Band size expected for the different yeasts species that have been identified so far in nigerian palm wine. [29]

3.5 Electrophoresis

In order to visualize the products of the different PCR reactions (diagnosis PCR and multiplex PCR) and the RFLP analysis, after the pertinent incubating protocol in the PCR machine we run the result in electrophoresis gels.

For the diagnosis PCR, we run the amplified product in a 1% agarose gel (dimensions 7x10 cm). The agarose gel was made 5 g. of agarose diluted in 500 ml of 0,5X TAE and heating the mixture until the agarose was completely melted.

Two combs of 30 "spikes" were used to make the wells, thus having 56 wells ready for the samples in each gel (leaving 4 free to place the ladder).
Once the gel was polymerized, it was placed in the trail and covered completely with 1X TAE as running buffer. In order to visualize the samples in the gel, we mixed 10 μl of the product of the amplification with 2 μl of dye before placing them in the wells. The electrophoresis was run at 100W during 30' and a 1 kb DNA standard (2μl in the wells) was used to compare the weight of the different DNA amplicons.

The different 1X and 0,5X TAE buffers were made from dilutions of stock TAE buffer (50X) by mixing either 20 mL of the concentrated buffer with 980 ml of distilled water for the first one, or 10 mL of the concentrated buffer with 990 mL of distilled water for the later (0,5X TAE).
The running buffer was changed every 3 electrophoresis to avoid exhaustion of the buffer.

However, while running the electrophoresis under these conditions for the RFLP analysis we observed that the samples seemed to be running off the gel (even the dye-front was not visible). We understood that the fragments resulting from the digestion were smaller than the PCR amplicon and thus their migration was faster. To solve this problem we decided to change the composition of the gel, trying a denser 1,7% agarose gel (prepared in the same conditions but using 8,5g. of agarose

instead of 5g.). The dimensions of the gel were the same.
Using the new concentration we avoided the problem, but as an additional precaution we decided to use just one row of wells for each run, having 28 samples in each electrophoresis gel.

In this case we did not need to mix the product of the digestion with the dye because the buffer already contained a dye that permitted the visualization of the samples. We also changed the ladder and used a 100 bp DNA standard more suitable for the size of the fragments.

Both for the 1% and the 1,7% agarose gels we pre-stained them using 5 μL of EtBr.

3.6 Sequencing

Some of the samples gave results for the multiplex PCR that seem to correspond to *Saccharomyces* spp. species that were not expected to be found in Africa, such as *S. mikatae*, *S. paradoxus* or *S. arboricola*.
For this reason we decided to send some of these samples to sequence the PCR amplicon.

We prepared the samples for the sequencing using a PCR purification kit. Once the purification protocol was done, we measured the DNA concentration in the purificated product using the nano spectrophotometer (NanoDrop^{TM}), and we diluted the samples to an appropriate concentration (5 ng/μl) for the length of the purified PCR products (which is around 700 bp).
The primer used to sequence was ITS4, the reverse primer used in the PCR amplification as described in section 3.3.2.

The samples sent to sequence were the following:

State	Strain ID	Length multiplex PCR product
BN	146	500
OS	811	
TB2	1008	
KG	604	800
OS	790	
LG	750	
KG	680	349
LG	749	
OS	813	

Table 3.6: List of strains sent to sequence

3.7 Visualization & Software

The electrophoresis gels were visualized under UV light and photographed and saved via computer assisted transmittance using Image Analysis Software: Quantity One© 1-D Analysis Software, provided by Bio-Rad.®

4

Results

In order to show the results obtained for the different procedures we will just include some pictures to illustrate what we wanted to obtain and for those cases we had anomalies.

The rest of the images can be seen in the appendixes. A table collecting the results of each experiment for every strain can be found in the appendix.

4.1 PCR amplification

Previous studies found out the estimated length for the amplified ITS-5.8S rRNA region of *Saccharomyces*.
The expected weight is around 800 bp. However, in some cases it is difficult to accurately state the exact size of the band (if the bands are not too sharp, for example), and for that reason we estimated as candidates to be *S. cerevisiae* all those fragments above 700 bp.

We performed the PCR amplification protocol (see section 3.3.2) to the total of 843 samples coming from the different regions of Nigeria.

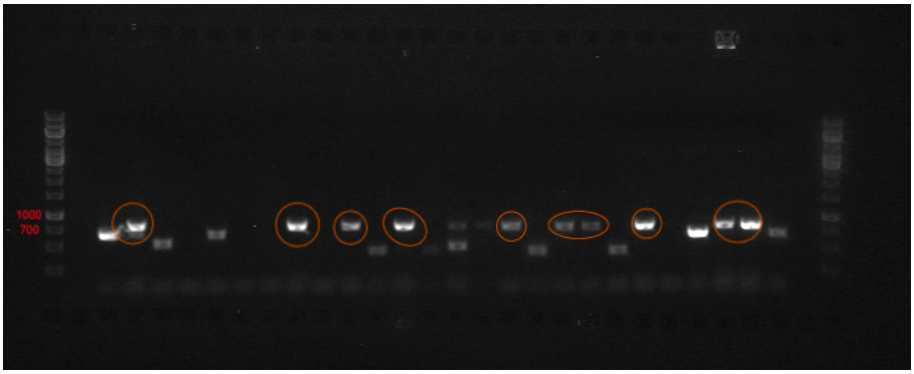


Figure 4.1: Image of the electrophoresis for the PCR amplification results of the first 28 samples coming from Kogi. *The circled bands are those ones that have the right size to be Saccharomyces.*
1kb ladder

In some of the wells we cannot observe any band. For these cases we repeated the PCR amplification later on. In some of the samples we got to obtain a product

when repeating the experiment. However, in other cases it was not possible.
Possible reasons for this might be any mistake when gathering the sample or the presence of inhibitors for the PCR in the sample.
There exist also the possibility of the absence of the region encompassed by the primer in some of the species, for example if some of them would actually be bacteria. ITS1 and ITS4 are primers designed for the identification of fungi so it is possible that if the sample is something different we could not get any amplification.

Table 4.1: Number of samples with an amplicon of 700 bp or above by region.

Anambara	40/74
Benue	67/91
Ebonyi	30/69
Enugu	68/82
Imo	17/27
Kogi	34/44
Lagos	9/78
Osun	32/55
Rivers	18/75
Taraba 1	25/81
Taraba 2	38/42
Total	378/718

In this stage we only focused in those ones that gave an amplicon size compatible with the size of *S. cerevisiae*, so this classification just divides the data in two groups: amplicon that could belong to *S. cerevisiae* and amplicons that do not have the expected size for this. This information is useful to confirm an identity to the strains later on.

4.2 RFLP

Using as reference the article published by Belloch et al. [29], we know the pattern we expect from each enzyme for *S. cerevisiae* and so, the profile we should obtain from the three of them to suspect a strain to be *S. cerevisiae*.

We did the RFLP analysis to the same 843 strains for which we had previously done the PCR amplification. For each sample we performed a digestion with each one of the three different restriction enzymes.

Enzyme	Restriction fragments
*Hha*I	385 + 365
*Bsu*RI	320 + 230 + 180 + 150
*Hinf*I	365 + 155

Table 4.2: Size in bp of the restriction fragments expected for *S. cerevisiae* [29]

Within the genus *Saccharomyces*, the profiles that result from the RFLP analysis can be very similar (the pattern for *Hha*I is the same for *S. cerevisiae*, *S. bayanus*, *S. paradoxus* and *S. pastorianus*, for example) and do not permit to distinguish between species in those cases from which the three different digestions are not available.

However, it is possible to disambiguate this by doing a multiplex PCR to all the putative *Saccharomyces* spp. In this way the combination of the RFLP profile and the appearance of a band of the expected size (as discussed in the next section, see 4.3) would permit to narrow the species the sample could belong to.

In spite of this, it is necessary to keep in mind that the only way to unquestionably assign an identity to a strain is by sequencing its DNA.

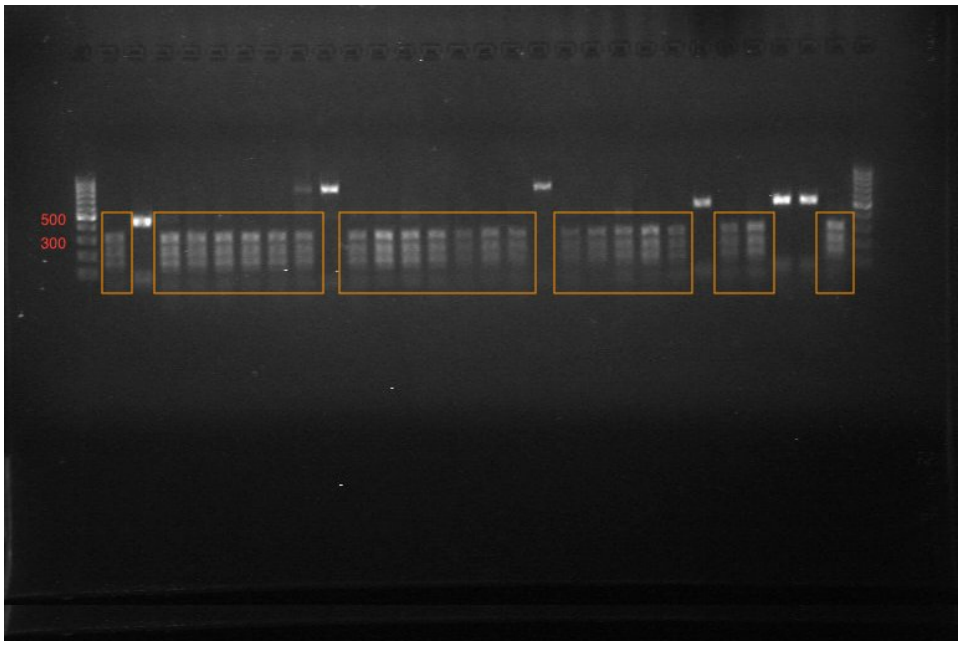


Figure 4.2: Image showing the digestion using *Bsu*RI in samples coming from Enugu (277-322). *The bands surrounded by a circle are those showing the expected pattern for S. cerevisiae*
100 bp ladder

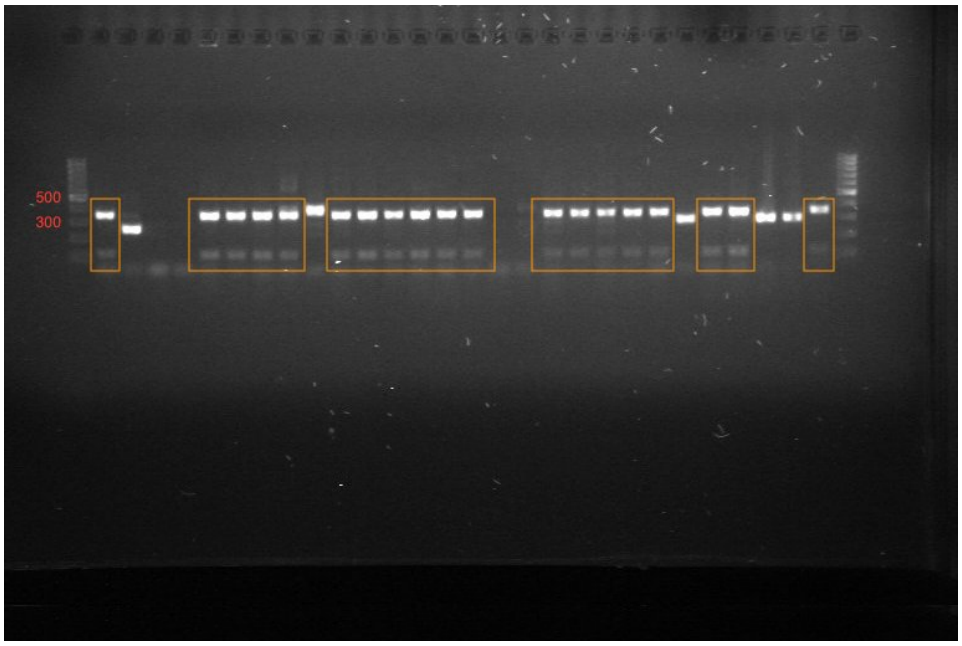


Figure 4.3: Image showing the digestion using *Hinf*I in samples coming from Enugu (277-322). *The bands surrounded by a circle are those showing the expected pattern for S. cerevisiae*
100 bp ladder

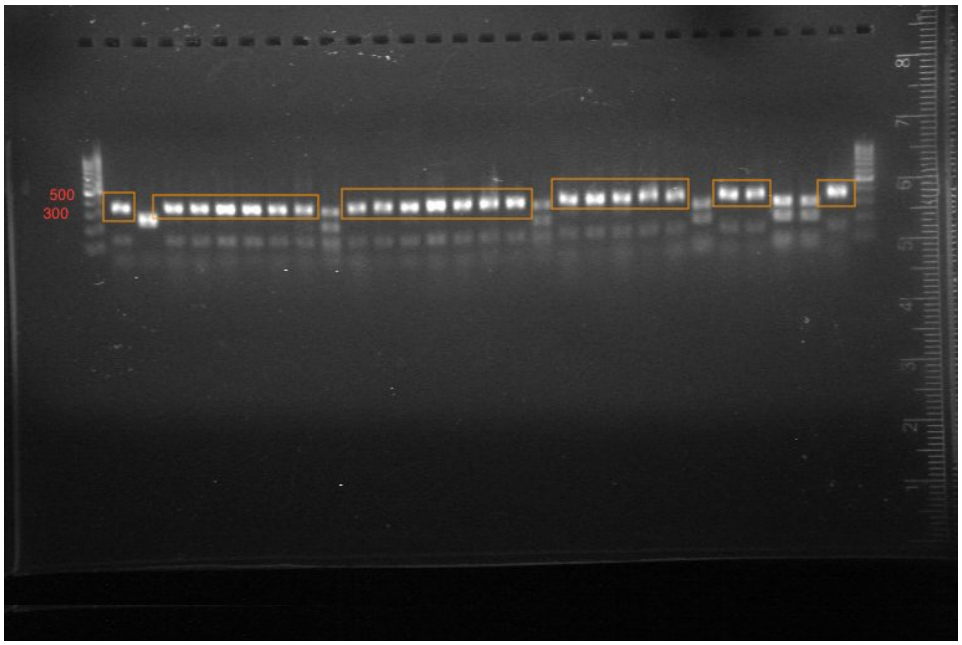


Figure 4.4: Image showing the digestion using *Hha*I in samples coming from Enugu (277-322). *The bands surrounded by a circle are those showing the expected pattern for S. cerevisiae*
100 bp ladder

In some cases we could not observe any band at all. An example of this would be samples from Benue where we could see the digestion by *Bsu*RI but no clear band appeared in the digestions by *Hinf*I and *Hna*I. The procedure was repeated three times (in identical conditions) with the same result.

Something similar happens with the samples from Imo, where everything seem to work fine with the first half of the samples but no bands can be observed in the last half for any of the enzymes.

Table 4.3: Results by region attending to RFLP results (regardless multiplex PCR

Region	Species	Number
Anambara	*Saccharomyces cerevisiae*	16
	Candida species	9
	Schizosaccharomyces pombe	1
	Saccharomyces capsularis	3
	Uniidentified	15
	No visible	29
Benue	*Saccharomyces cerevisiae*	51
	Candida species (no *tropicalis*	5
	Candida tropicalis	11
	H. mrakii/*Pichia* species	4
	Unidentified	0
	No visible	20
Ebonyi	*Saccharomyces cerevisiae*	18
	Candida species	22
	Pichia species	7
	Unidentified	17
	No visible	4
Enugu	*Saccharomyces cerevisiae*	63
	Candida species	13
	Pichia species	3
	Torulaspora debrueckii	2
	Unidentified	0
	No visible	0
Imo	*Saccharomyces cerevisiae*	7
	Pichia species	2
	Unidentified	13
	No visible	6
Kogi	*Saccharomyces cerevisiae*	23
	Candida species	6
	Pichia species	5
	Other *Saccharomyces*	3
	Unidentified	0
	No visible	8
Lagos	*Saccharomyces cerevisiae*	9
	Candida species	14
	Pichia species	40
	Other *Saccharomyces*	1
	Unidentified	3
	No visible	12

Table 4.4: Results by region attending to RFLP results (continuation)

Region	Species	Number
Osun	*Saccharomyces cerevisiae*	25
	Candida species	12
	Other *Schizosaccharomyces*	4
	Pichia species	1
	Kluyveromyces species	4
	Uniidentified	3
	No visible	5
Rivers	*Saccharomyces cerevisiae*	16
	Candida species	36
	Pichia species	7
	Other *Saccharomyces*	3
	Unidentified	0
	No visible	13
Taraba 1	*Saccharomyces cerevisiae*	4
	Candida species	33
	Pichia species	7
	Unidentified	17
	No visible	4
Taraba 2	*Saccharomyces cerevisiae*	27
	Other *Saccharomyces*	1
	Unidentified	0
	No visible	13

While analyzing the samples from Benue, we encountered the problem that the only restriction enzyme that apparently was able to cut the DNA was *Bsu*RI, so no pattern was available for any of the other two restriction enzymes. The digestion was repeated three different times under the same conditions without success. A possibility could be that in the samples there is something that acts as an inhibitor for these enzymes, since we have only faced this situation in this region.

The lack of information coming from two of the three restriction enzyme makes more complicated to identify the supposed identity of the strains, so the results of this region should be handled with additional caution.

We would like to highlight the case of Taraba 1, where there is a lack of *Saccharomyces cerevisiae*. In this region it was also very difficult to try to assign an identity to many of the samples because in many cases the information coming from the restriction enzymes was missing. However, it was possible to see that the variety among strains (although we were not able to give them a name) was greater than for other regions, where there are usually 3 or 4 species predominating.
However, in Taraba 2 (both clusters come from the same region) the presence of *Saccharomyces cerevisiae* is predominant. It is possible that the samples were distributed this way on purpose, but otherwise it would be interesting to determine

whether this difference can be related to the origin of the samples (where in the region where they extracted), or the tree species, etc.

4.3 Multiplex PCR

We performed the multiplex PCR protocol (see section 3.3.3) to all the samples that have a PCR product (resulting from the PCR amplification) of 650 bp or above (up to 1000 bp). The total of strains fulfilling this criterion were 462.

The results revealed that not all the species that were thought to be *Saccharomyces* attending to the RFLP analysis belong to this species. However, most of the bands seem to correspond to the genus *Saccharomyces* spp. except for some samples that gave multiple bands and did not match with any of the expectations from previous experiments. Probably these strains are actually bacteria resistant to chloramphenicol, the antibiotic used in the plates to grow the samples.

Based on previous experiments [41] we know that the expected amplification product for *S. cerevisiae* in the multiplex reaction should be a single band around 150 bp.

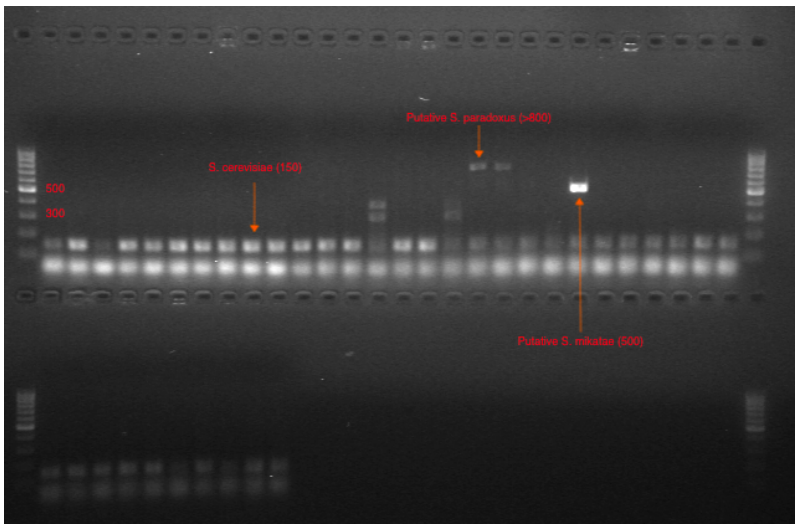


Figure 4.5: Results of multiplex PCR from samples 502-693.
100 bp ladder

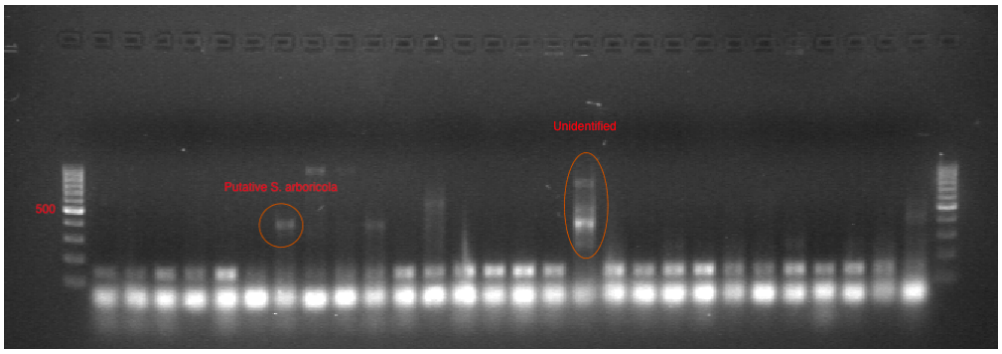


Figure 4.6: Results of multiplex PCR from samples 696-945, first row 100 bp ladder

However, we also found some bands (31 in total) that seem to correspond to 349 bp. According to *Muir et al.* [41] this band could correspond to *S. arboricola*. This species has not been found so far in palm wine in Nigeria. Unfortunately, it was not considered in the RFLP analysis performed by *Belloch et al.* [29] so it is not possible to compare this profile.
In order to confirm the identity of these strains, the solution would be to sequence, for example, their ITS region.

Something similar occurs with bands that have a size of 1000 bp or similar (in total there are 15 strains with this characteristics, that have been gathered in the table below within the category of 800 bp). This size, according to the paper we are using as reference, would correspond to *S. paradoxus*. This species has not been found in previous studies in Nigerian palm wine and thus it could be interesting to sequence these samples to see if indeed they correspond to *S. paradoxus* strains. However, the size of the bands could have been overestimated in some cases and they could correspond to bands slightly smaller, which could correspond to *S. kuvriadzevii* or maybe a hybrid species (which would be more unlikely).

Another deviation from the 150 bp we expect for *S. cerevisiae* is a band under 100 bp (this band was found in 15 samples). This amplicon is smaller than the *S. cerevisiae* amplicon and none of the *Sacharomyces* species expected to be found in palm wine could (in theory) have produced it, so these ones would as well worth a further study. Some of this bands below 100 bp could be *Schizosacharomyces pombe*, attending to their RFLP profile and their PCR amplicon, but for many it is hard to determine the species by using the RFLP profile (sometimes incomplete) and the PCR amplicon, so it is difficult to assign a solid identity. A solution for this would be to sequence some of these samples.

Corresponding to the region Ebonyi, there are some multiplex profiles that do not correspond to what we would expect from the RFLP analysis and also give the strange pattern we have mentioned above (several bands), that could belong to bacteria.

Size (bp)	Putative species	Number
150	*Saccharomyces cerevisiae*	372
90	*Schizosaccharomyces pombe*	15
349	*Saccharomyces arboricola*	31
500	*Saccharomyces mikatae*	5
800	*Saccharomyces paradoxus*	15
Unidentified	-	11
No visible	-	11

Table 4.5: Table summarizing the results obtained from the multiplex PCR

In the table summarizing the results, we have included in each species all the strains that gave a result corresponding to those bp (approximately) for the multiplex PCR. The anomalies mentioned before (multiple bands pattern that did not correspond to any of the species studied in the paper used as reference, [41]) have been included in the category "Unidentified".

4.4 Sequencing

We obtained the results for the sequencing of the PCR amplicon for the samples indicated in the section 3.6 and performed a BLAST to determine their identity.

Strain ID	BLAST match
811	*S. cerevisiae*
604	*S. cerevisiae*
790	*S. cerevisiae*
750	*Meyerozyma caribbica*
680	*Meyerozyma caribbica*
749	*Cryptococcus curvatus*
813	*S. cerevisiae*
146	*S. cerevisiae*
1008	*S. cerevisiae*

Table 4.6: Sequencing results

5
Conclusion

The use of molecular techniques seem to be more accurate than identification regarding morphological or metabolic features (such as the methods that had been used in the past to identify the microorganisms present in Nigerian palm wine).

The microorganisms identified in palm wine originating from Nigeria in this study correspond to the microorganisms found previously in studies analyzing palm wine from different regions in Africa. However, with the protocol established it was also possible to identify some strains that are likely to belong to *Saccharomyces species* that we were not expecting in this source from previous studies, such as *S. arboricola* or *S. mikatae* or some possible strains of *Sch. pombe*.

Another possibility to explain these "abnormalities" could be the presence of *Saccharomyces* hybrids that would result in a different pattern both for the RFLP analysis and the multiplex PCR (giving for one sample bands that correspond to two different species).

S. arboricola is endemic of East Asia (China) and it has only been identified in this geographical region, except for another possible finding in Taiwan [12].
Something similar is observed for *S. mikatae*, endemic of Japan, which also has not been identified anywhere else.
S. paradoxus is the closest relative of *S. cerevisiae*. It has been identified in northern hemisphere forests, but has not been isolated in the south hemisphere so far. [42]
If these finding correspond indeed to these yeast species, we think it is safe to say that they are not endemic of Africa (since many other regions have been studied applying similar methods than the ones used in this project) but they could have been introduced.
Since the biodiversity of yeasts in Africa has not been as studied (nor exploited) as in European or North American niches, it is possible that some of these species could have reached Nigeria because of human activity (trade or other reasons) and that its presence has been unnoticed due to the lack of study with accurate enough techniques.

Since these results were surprising, we sent 9 strains (3 of them corresponding to putative *S. mikatae*, 3 corresponding to putative *S. arboricola* and 3 corresponding to putative *S. paradoxus* that were found during the project) to sequence so we could confirm its identity.

There is a strong possibility that some of these findings are actually bacteria (although the majority of the samples corresponded to yeasts, probably due to the use of enriched media), since some of the strains we have been working with were already sent to sequence and they turned out to be bacteria resistant to chloramphenicol, which is the antibiotic that the YPD plates were treated with.
A way to confirm this would be checking its morphology using the microscopy.
Probably the strains that gave multiple band patterns (classified as "Unidentified") for the multiplex PCR correspond to bacteria contamination as well.

The results of the sequencing actually reject these theories, since most of the strains that were sent to sequence turned out to be *S. cerevisiae*, except for two strains corresponding to *Meyerozyma caribbica* and *Cryptococcus curvatus*.
The region sequenced was the ITS region amplified in the diagnosis PCR, whereas the PCR multiplex amplifies regions of different genes for each strain. There is the possibility that some strains could be hybrids between two *Saccharomyces* spp. that would present this particularity. More likely, there has been some error during the procedure.
Still the strains identified with the sequencing were just a small subset of the putative strains suggested by the multiplex PCR results, so the multiplex PCR results cannot be fully rejected.

Regarding the other species identified, both of them are yeasts.
Meyerozyma caribbica belongs to the phylum of Ascomycetes, was previously recognized as part of *Candida* spp. and it is usually linked to fermented foods and widely distributed since it has been isolated in many positions in the West Hemisphere. [45]. However, to the best of my knowledge it has never been reported in Nigerian palm wine before.

As to *Cryptococcus curvatus*, previously known as *Candida curvata*, it is classified in the phylum of Basidiomycetes. It is an oleaginous yeast fairly studied for its ability to produce lipids from low-cost substrates (as sugar) under nitrogen limiting conditions. It is as well widely distributed and it had not been identified in Nigerian palm wine previously. [46]

It is possible that these species were unnoticed before due to the inaccuracy of the methods previously used to study biodiversity in palm wine coming from Nigeria, and the difficulty to distinguish them from other species such as *Candida* spp.

From the results we can conclude that the composition of yeasts in palm wine largely differs from one region to another, which would affect to the production of palm wine. Since the percentage of sugar present in the sap varies seasonally, it would also be interesting to study if this composition changes depending on this.

It is quite interesting that in some regions the predominant species is not *Saccharomyces cerevisiae* as we would expect due to its role in fermentation. This could mean that there are other species (for example members of *Candida* spp.) driving

the anaerobic process, not allowing *Saccharomyces* spp. to take the lead, maybe just because the number of non-*Saccharomyces* yeasts greatly overtakes the amount of *Saccharomyces* spp.

In order to industrialize palm wine production, a standardized starter culture of yeasts that would start fermentation would be necessary. Next steps of the project, which include phenotyping using around 60 different environments (including different carbon and nitrogen sources) will give more information that could be used to standardize an optimal starter culture to the production of this beverage.
The results obtained in this part of the project, however, should be taken just as illustrative since the techniques used present the disadvantage that they are, sometimes, quite subjective when assigning the size of the digested fragments, which could lead to mistakes when assigning the identity.
In order to partially avoid this problem, we could have included a digested control in all the digestions, ideally of all the species that had been previously found so the comparison of the patterns would have been easier.

In spite of this, since the main goal of the project is to identify *S. cevrevisiae*, and the pattern that results for this yeast is reasonably good (only difficult to distinguish from *S. bayanus*, *S. pastorianus* and *S. paradoxus*) and we can also confirm its identity with the PCR multiplex, these concrete results can be trusted. However, we should always keep in mind that the only way to know the identity of a species completely sure is to sequence its genome (or a region with enough divergence among species, such as the ITS region).

Bibliography

[1] S. Marsit, J. Leducq, E. Durand et al. *Evolutionary biology through the lens of budding yeast comparative genomics* Nature Review Genetics, 2017

[2] B. Bowman, J. Taylor et. al. *Molecular Evolution of the Fungi: Relationship of the Basidiomycetes, Ascomycetes, and Chytridiomycetes* Molecular Biology Evol., 1992

[3] A. MC, M. PB et. al. *An overview of the higher level classification of Pucciniomycotina based on combined analyses of nuclear large and small subunit rDNA sequences.* Mycologia, 2006

[4] J. Ropars, C. Maufrais et al. *Gene flow contributes to diversification of the major fungal pathogen Candida albicans* Nature Communications, 2018

[5] S. H. Mohd Azhar, R. Abdulla, S.A. Jambo, H. Marbawi, et al. *Yeasts in sustainable bioethanol production: A review* Biochemistry and Biophysics Reports, 2017

[6] R. K. Mortimer *Evolution and Variation of the Yeast (Saccharomyces) Genome* Genome Research, 2000

[7] D. Libkind, C. T. Hittinger, E. Valério et. al. *Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast* PNAS, 2011

[8] B. Gallone, J. Steensels et. al. *Domestication and Divergence of Saccharomyces cerevisiae Beer Yeasts* Cell, 2016

[9] J. R. Bellon, F. Schmid et. al. *Introducing a New Breed of Wine Yeast: Interspecific Hybridisation between a Commercial Saccharomyces cerevisiae Wine Yeast and Saccharomyces mikatae* PLOS One, 2013

[10] S. González, E. Barrio, J. Gafner, A. Querol. *Natural hybrids from Saccharomyces cerevisiae, Saccharomyces bayanus and Saccharomyces kudriavzevii in wine fermentations* FEMS, 2006

[11] I. Masneuf, J. Hansen, C. Groth, J. Piskur, D. Dubourdieu *New Hybrids between Saccharomyces Sensu Stricto Yeast Species Found among Wine and Cider Production Strains* Applied and environmental microbiology, 1998

[12] G. Naumov, C-F. Lee, E. S. Naumova *Molecular genetic diversity of the Saccharomyces yeasts in Taiwan: Saccharomyces arboricola, Saccharomyces cerevisiae and Saccharomyces kudriavzevii* Antonie van Leeuwenhoek, 2013

[13] M. Johnston *Genome sequencing: The complete code for a eukaryotic cell* Current Biology, 1996

[14] K-K. Hong, J. Nielsen *Metabolic engineering of Saccharomyces cerevisiae: a key cell factory platform for future biorefineries* Cellular and molecular life sciences, 2012

[15] G. G. Stewart *Industrial uses of yeast – brewing and distilling* New Food, 2016

[16] J. Steensels, T. Snoek, E. Meersman et. al. *Improving industrial yeast strains: exploiting natural and artificial diversity* FEMS Microbiology Review, 2014

[17] A. Aguilera, S. Chávez, F. Malagón *Mitotic recombination in yeast: elements controllingits incidence* Yeast, 2000

[18] H. Li, M. Wu, L. Xu, et. al. *Evaluation of industrial Saccharomyces cerevisiae strains as the chassis cell for second-generation bioethanol production* Microbial Biotechnology, 2015

[19] D. Mattanovich, M. Sauer, B. Gasser. *Yeast biotechnology: teaching the old dog new tricks*. Microbial Cell Factories, 2014.

[20] A. Fiener, A. Canbafi, M. Ümit Ünal *The Effect of Fermentation Temperature on the Growth Kinetics of Wine Yeast Species* 2007

[21] L. Alba-Lois, C. Segal-Kischinevzky. *Beer & Wine Makers*. Nature Education, 2010.

[22] T. Pfeiffer, A. Morley *An evolutionary perspective on the Crabtree effect* Frontiers in molecular biosciences, 2014

[23] J. Thierie, M. Penninckx *Crabtree Effect* Research Gate, 2010

[24] Cristina Pina*, Cristina Santos, Jose! Anto!nio Couto, Tim Hogg *Ethanol tolerance of five non-Saccharomyces wine yeasts in comparison with a strain of Saccharomyces cerevisiae—influence of different culture conditions* Food Microbiology, 2003

[25] Z. Chi and N. Arneborg *Saccharomyces cerevisiae strains with different degrees of ethanol tolerance exhibit different adaptive responses to produced ethanol* Journal of Industrial Microbiology Biotechnology, 2000

[26] W.K. Amoa-Awua E. Sampson K. Tano-Debrah *Growth of yeasts, lactic and acetic acid bacteria in palm wine during tapping and fermentation from felled oil palm (Elaeis guineensis) in Ghana* Journal of Applied Microbiology, 2016

[27] M Tristezza, C. Veltrano et. al. *Biodiversity and safety aspects of yeast strains characterized from vineyards and spontaneous fermentations in the Apulia Region, Italy* Food Microbiology, 2013

[28] F.F. Bauer, I. S. Pretorius *Yeast Stress Response and Fermentation Efficiency: How to Survive the Making of Wine - A Review* S. Afr. J. Enol. Vitic., 2000

[29] C. Belloch, F. Uruburul, A. Queroi *Identification of yeasts by RFLP analysis ofthe 5.8S rRNA gene and the two ribosomal internal transcribed spacers*. 2018.

[30] I. Ezeagu, M. Fafunso, F. E. Ejezie. *Biochemical constituents of palm wine* Ecology of Food and Nutrition, 2003.

[31] O. Ezeronye, P. Okerentugba. *Genetic and physiological variants of yeast selected from palm wine*. 2001.

[32] J. Uchenna *The Value of Palm Wine Tapping in the Food Production Practices of Igbo-Land: A Case Study of Idemili South Local Government Area, Anambra State* Research on Humanities and Social Sciences, 2014

[33] F. Morah *The Nigerian palm-wine: science and socio-economic importance* 2010

[34] G. C. Whiting *Organic acid metabolism of yeasts during fermentation of alcoholic beverages* J. Inst. Brew, 1976

[35] G. M. Walker, G. G. Stewart *Saccharomyces cerevisiae in the Production of Fermented Beverages* Beverages, 2016

[36] F. Morah. *The nigerian palm-wine: science and socioeconomic importance.* 2010.

[37] P. Sahare *EcoRV digestion patterns in ITS region of medicinal plants.* International Journal of Scientific Engineering Research, 2013

[38] C. Universitaria. *Mini Review: Microbiology and biochemistry of traditional palm wine produced around the world.* 2014

[39] H. Tuppy, P. Swetly *Structural Protein of Wild-Type (ρ^+) and Respiration-Deficient (ρ^-) yeast mitochondria. A comparative study using electrophoretic and immunological methods.* 1968

[40] M. Stringini, F. Comitini, M. Taccari et al. *Yeast diversity during tapping and fermentation of palm wine from Cameroon* Food microbiology, 2009

[41] A. Muir, E. Harrison A. Wheals. *A multiplex set of species-specific primers for rapid identification of members of the genus Saccharomyces.* Yeast Research, 2011

[42] Velimir Gayevskiy and Matthew R. Goddard *Saccharomyces eubayanus and Saccharomyces arboricola reside in North Island native New Zealand forests* Environmental Microbiology, 2016

[43] G. Suzzi, F. Dellaglio *Rapid identification and differentiation of Saccharomyces cerevisiae, Saccharomyces bayanus and their hybrids by multiplex PCR* 2004

[44] A. Rokas et. al. *Genome-scale approaches to resolving incongruence in molecular phylogenies* Nature, 2003

[45] Wahengbam Romi, Santosh Keisam, Giasuddin Ahmed, and Kumaraswamy Jeyaram *Reliable differentiation of Meyerozyma guilliermondii from Meyerozyma caribbica by internal transcribed spacer restriction fingerprinting* BMC Microbiology, 2014

[46] Sidoine Sadjeu Tchakouteu, Afroditi Chatzifragkou, Ourania Kalantzi, Apostolis A. Koutinas, George Aggelis and Seraphim Papanikolaou *Oleaginous yeast Cryptococcus curvatus exhibits interplay between biosynthesis of intracellular sugars and lipids* Eur. J. Lipid Sci. Technol. 2015