Biotyping *Saccharomyces cerevisiae* strains using Matrix Assisted Laser Desorption/Ionization Time of Flight Mass Spectrometry (MALDI-TOF MS)

by

Anushka Moothoo-Padayachie

Dissertation presented for the degree of
Master of Science (Biochemistry)

at
University of KwaZulu-Natal
School of Biochemistry, Genetics and Microbiology
Faculty of Science and Agriculture

Supervisor:	Dr. Patrick Govender
Co-supervisors:	Prof. Hendrik G. Kruger
	Dr Glenn E. M. Maguire
	Dr. Thavendran Govender

April 2011

DECLARATIONS

FACULTY OF SCIENCE AND AGRICULTURE
DECLARATION 1 - PLAGIARISM

I, Ms Anushka Moothoo-Padayachie declare that

1. The research reported in this thesis, except where otherwise indicated, is my original research.
2. This thesis has not been submitted for any degree or examination at any other university.
3. This thesis does not contain other persons' data, pictures, graphs or other information, unless specifically acknowledged as being sourced from other persons.
4. This thesis does not contain other persons' writing, unless specifically acknowledged as being sourced from other researchers. Where other written sources have been quoted, then:
    a. Their words have been re-written but the general information attributed to them has been referenced
    b. Where their exact words have been used, then their writing has been placed in italics and inside quotation marks, and referenced.
5. This thesis does not contain text, graphics or tables copied and pasted from the Internet, unless specifically acknowledged, and the source being detailed in the thesis and in the References sections.

Signed: ……..

Declaration Plagiarism 22/05/08 FHDR Approved

FACULTY OF SCIENCE AND AGRICULTURE
DECLARATION 2 - PUBLICATIONS

DETAILS OF CONTRIBUTION TO PUBLICATIONS that form part and/or include research presented in this thesis (include publications in preparation, submitted, *in press* and published and give details of the contributions of each author to the experimental work and writing of each publication)

Publication 1

- *NOT APPLICABLE*

Signed: ……..

Declaration Publications FHDR 22/05/08 Approved

I, Dr Patrick Govender as supervisor of the MSc study hereby consent to the submission of this MSc Thesis.

_____

_____

Date

SUMMARY

In clinical diagnosis and fermentation industries there is a need for a method that allows for the differentiation of yeast to the strain level (biotyping). The ideal biotyping method should be accurate, simple, rapid and cost-effective, and capable of testing a large number of yeast isolates. Matrix assisted laser desorption/ionization time of flight mass spectrometry has emerged as a powerful biotyping tool for the identification of bacteria and clinical yeast isolates, mainly *Candida*. It has been found that the MALDI-TOF MS signals from yeast are harder to obtain than from bacteria. It has been reported by several research studies that a cell lysis step is required to obtain a mass spectral signal for clinical *Candida* strains. To date an optimized sample preparation protocol has not been devised for the biotyping of *S. cerevisiae* strains. Studies on the identification of yeast using MALDI-TOF MS have focused primarily on clinical *Candida* yeast isolates but have included very few *S. cerevisiae* strains. Furthermore these yeast identification studies using MALDI-TOF MS have only achieved identification to the species and not strain level. A major limiting attribute of MALDI-TOF MS for the accurate identification of microbes, is its dependency on a comprehensive mass spectral database. Bruker Daltonics is a pioneer and leader in providing innovative life science tools based on mass spectrometry thus the Bruker Daltonics mass spectral database and state-of-the-art instruments and accompanying software were selected for this study. The Bruker Daltonics mass spectral database currently holds three thousand seven hundred and forty microorganisms of which only a mere seven are *S. cerevisiae* strains.

Initially in this study, a number of parameters of a generic ethanol/formic acid protein extraction procedure as originally described by Bruker Daltoincs were considered in the development of a sample preparation protocol that yielded characteristic and highly reproducible MALDI-TOF mass spectra. The parameters considered included cell number, alcohol fixation, matrix solution and media. It was found that using the optimized sample preparation protocol unique and highly reproducible mass spectral profiles were obtained for all three *S. cerevisiae* strains. Multivariate analysis confirmed that the differences between all three *S. cerevisiae* strains were statistically significant. For quality assurance, the spectra of the three strains were sent for evaluation by Bruker Daltonics and were deemed suitable for the purpose of biotyping.

The newly created ethanol/formic acid extraction procedure was used to generate an *S. cerevisiae* mass spectral database comprising of forty-five *S. cerevisiae* strains within a local context but also of global significance. The accuracy of the mass spectral database was assessed using blind coded *S. cerevisiae* strains obtained from the Agricultural Research Council Infruitec-Nietvoorbij (Institute for Deciduous Fruit, Vines and Wine), Stellenbosch, South Africa. It was found that *S. cerevisiae* identification to the species and more importantly strain

level was achievable with relatively good accuracy. To determine the potential application of MALDI-TOF MS as an accurate method for *S. cerevisiae* strain identification in industry, blind coded *S. cerevisiae* strains were obtained from Natal Cane Products and subjected to MALDI-TOF MS analysis. It was found that four of the pure cultures submitted were correctly identified to the strain level and the three *S. cerevisiae* strains incorrectly identified may have been contaminants or the result of incorrect optimization conditions for the fermentation. Thus MALDI-TOF MS was shown to be an accurate identification tool, that may also be used to detect contaminants or incorrect environmental conditions which can result in substantial losses.

This dissertation is dedicated to my family and in remembrance of my beloved cousin Desiree Naidoo

BIOGRAPHICAL SKETCH

Anushka Moothoo-Padayachie was born in Port Elizabeth on the 13th of January 1987. Her family moved to Durban where she matriculated with merit at Danville Park Girls High School in 2004. Anushka has always shown a keen interest in science which was demonstrated by her active role in environmental groups and participation in science olympiads from the early age of 10 years old. She went on to pursue a Bachelor of Science degree at the University of KwaZulu-Natal (Westville campus) majoring in Microbial biotechnology. Her interest in Microbiology led her to register for a Bachelor of Science (Honours) degree. She is currently pursuing a Masters degree in Biochemistry and in the interim has been employed by the University of KwaZulu-Natal to run a MALDI-TOF MS biotyping facility. During the course of her masters studies, Anushka has played an active role in academia tutoring Biology part-time at the University of South Africa. Anushka is determined to go further in her academic career and eager to obtain her Doctor of Philosophy (PhD) in Science.

Anushka is the second of three daughters and lives with her family. In her spare time she is an amateur artist and enjoys oil painting.

ACKNOWLEDGEMENTS

I wish to express my sincere gratitude and appreciation to the following persons and institutions:

- Dr Patrick Govender, I thank you for giving me the opportunity to pursue a masters degree in the department of Biochemistry. In addition, for all your guidance, motivation, and the many skills I have acquired under your supervision.

- I also wish to thank Prof. Hendrik G. Kruger, Dr Glenn E. M. Maguire and Dr. Thavendran Govender for your support, guidance and help with my project.

- The GGKM research group and National research foundation (NRF) for your financial support.

- All those at Bruker Daltonik for all their technical support and encouragement.

- NCP Alcohols, Durban, South Africa, South African breweries Ltd. Johannesburg, South Africa and Agricultural Research Council Infruitec-Nietvoorbij (Institute for Deciduous Fruit, Vines and Wine), Stellenbosch, South Africa for your participation in this masters research study.

- Theo, Dinesh and Danny for their technical assistance in the laboratory.

- To all the staff and students at the University of KwaZulu-Natal that assisted me in any manner.

- Brandon, Deshni, Karen, Charlotte, Thashini, Sizwe, Kenny and Mya for your invaluable support, encouragement and friendship.

- Most importantly to my parents, Des and Jeannie Moodley, sisters Nitasha and Karissa, thank you for all your love, advice, support, time, patience, encouragement and understanding.

PREFACE

This dissertation is presented as a compilation of 5 chapters.

Chapter 1 **General Introduction and Project Aims**

Chapter 2 **Literature Review**

Biotyping of yeast strains

Chapter 3 **Research Results 1**

Optimization of MALDI-TOF MS sample preparation parameters

Chapter 4 **Research Results 2**

Biotyping of *Saccharomyces cerevisiae*

Chapter 5 **General Discussion and Conclusion**

CONTENTS

CHAPTER 1. INTRODUCTION AND PROJECT AIMS 1

1.1 INTRODUCTION 1
1.2 AIMS OF STUDY 2
1.3 REFERENCES 3

CHAPTER 2. IDENTIFICATION OF YEAST STRAINS 5

2.1 INTRODUCTION 5
2.2 PHENOTYPIC METHODS 6
2.3 BIOCHEMICAL METHODS 8
2.4 COMMERCIAL IDENTIFICATION KITS 8
2.5 KARYOTYPING 9
2.5.1 Chromosomal gel electrophoretic techniques 9
2.6 MOLECULAR GENETIC METHODS 10
2.6.1 Radio-labelled DNA probes 11
2.6.2 PCR-based identification techniques 11
2.6.2.1 Retrotransposon specific primers 12
2.6.2.2 Intron specific primers 12
2.6.2.3 *COX1*-specific primers 13
2.6.2.4 PCR-amplified ribosomal RNA gene fragments 14
2.6.2.5 PCR-temperature gradient gel electrophoresis (PCR-TGGE) 14
2.6.2.6 Microsatellite marker specific primers 15
2.6.2.7 Real-time PCR 15
2.7 MASS SPECTROMETRY 16
2.7.1 Pyrolysis Mass Spectrometry (PyMS) and Fourier Transform Infrared (FT-IR) spectroscopy 16
2.7.2 Matrix assisted laser desorption ionization (MALDI) spectrometry 18
2.7.3 Time-of-flight (TOF) mass spectrometer 21
2.7.4 Spectra generation using a Time-of-flight (TOF) mass spectrometer 22
2.7.5 Reflectron mass spectrometer 24
2.8 MALDI-TOF MS FACILITATED BIOTYPING OF MICRO-ORGANISMS 25
2.8.1 Bacteria 26
2.8.1.1 Bacterial source tracking (BST) 26
2.8.1.2 Pathogenic non-fermenting bacteria 26
2.8.1.3 *Campylobacter* 27
2.8.1.4 *Helicobacter* 27
2.8.2 Yeasts 28
2.8.2.1 MALDI-TOF analysis of intact yeast cells 28
2.8.2.2 MALDI-TOF analysis of yeast protein extracts 29
2.9 CONCLUSION 31
2.10 REFERENCES 31

CHAPTER 3. OPTIMISATION OF SAMPLE PREPARATION PROTOCOLS FOR MALDI-TOF MS IDENTIFICATION OF *SACCHAROMYCES CEREVISIAE* STRAINS 35

3.1 ABSTRACT 35
3.2 INTRODUCTION 36
3.3 MATERIALS AND METHODS 38
3.3.1 Strains 38
3.3.2 Media and Cultivation conditions 38
3.3.3 Direct transfer whole-yeast cell analysis 38
3.3.4 Enumeration of Yeast cells 38
3.3.5 Optimized ethanol/formic acid extraction procedure 39
3.3.6 Variations of ethanol/formic acid extraction procedure 39
3.3.7 Acquisition of MALDI-TOF MS data 39
3.3.8 Principle component analysis 40
3.4 RESULTS 41
3.4.1 Direct transfer whole-yeast cell analysis 41
3.4.2 The effect of cell number 41
3.4.3 The effect of alcohol fixation 42
3.4.4 Matrix selection 43
3.4.5 The effect of culture media 45
3.4.6 Distinctive mass spectral fingerprints of *S. cerevisiae* strains 46
3.5 DISCUSSION 47
3.6 ACKNOWLEDGEMENTS 49
3.7 REFERENCES 49

CHAPTER 4. BIOTYPING *SACCHAROMYCES CEREVISIAE* STRAINS 51

4.1 ABSTRACT 51
4.2 INTRODUCTION 52
4.3 MATERIALS AND METHODS 54
4.3.1 Strains 54
4.3.2 Media and cultivation conditions 55
4.3.3 Optimized ethanol/formic acid extraction procedure 55
4.3.4 Acquisition of MALDI-TOF MS data 55
4.3.5 Creation of mass spectral database entries 56
4.3.6 Yeast chromosomal DNA isolation 56
4.3.7 PCR-based yeast strain identification 57
4.3.8 Contour-clamped homogenous electric field gel electrophoresis 57
4.3.9 MALDI-TOF-based biotyping of blind-coded *S. cerevisiae* strains 58
4.3.10 Principle component analysis 58

4.4 RESULTS 59
4.4.1 *S. cerevisiae* database creation 59
4.4.2 PCR-based yeast strain identification 61
4.4.3 CHEF-based karyotype strain differentiation 63
4.4.4 MALDI-TOF-based biotyping of blind-coded *S. cerevisiae* strains 64
4.5 DISCUSSION 66
4.6 ACKNOWLEDGEMENTS 68
4.7 REFERENCES 68

CHAPTER 5. GENERAL DISCUSSION AND CONCLUSION 71

5.1 GENERAL DISCUSSION AND CONCLUSION 71
5.2 ACKNOWLEDGEMENTS 73
5.3 REFERENCES 73

Chapter 1

GENERAL INTRODUCTION AND PROJECT AIMS

1. INTODUCTION AND PROJECT AIMS

1.1 INTRODUCTION

The yeast *Saccharomyces cerevisiae* also known as bakers and brewers yeast has been used in food production for thousands of years. The contribution of *S. cerevisiae* to the progression of mankind can be attributed mainly to its rapid and efficient ability to convert sugars (mono- and disaccharides) into alcohol and carbon dioxide. The extensive use of *S. cerevisiae* by mankind over the ages for the preparation of numerous products such as wine, beer and bread, has lead to it being considered a GRAS (generally regarded as safe) organism (Kjeldsen, 2000).

Today, *S. cerevisiae* plays a pivotal role in many industries such as the baking, brewing, bioethanol, wine and other alcoholic beverage industries (Dombek and Ingram, 1987). *S. cerevisiae* now has one of the the best characterised eukaryotic metabolome, proteome and genome. The entire *S. cerevisiae* genome has been elucidated and the current consensus is that the number of genes is close to six thousand (Kessler*, et al.*, 2003). In recent years, the species has been developed as an efficient eukaryotic expression system for biotechnology. Therefore *S. cerevisiae* has become increasingly important in the pharmaceutical and nutraceutical industries. Since 1987, *S. cerevisiae* has been used for the production of genetically engineered human insulin. Although a simple eukaryote, it is remarkably efficient at expressing and secreting insulin, and it offers high productivity (Kjeldsen, 2000). *S. cerevisiae* can also be genetically engineered to produce L-lactic acid of extremely high optical purity (Saitoh*, et al.*, 2005). *S. cerevisiae* appears to be an ideal cell factory for the synthesis of important secondary metabolites. Natural products encompass a wide range of small molecules derived from both eukaryotic and prokaryotic species. They are of commercial interest as aroma and flavour enhancers (terpenes), food colourants and antioxidants (carotenoids), and antiparasitic and anticarcinogenic compounds (taxol) (Chemler*, et al.*, 2006).

In the fermentation industry there is a need for a method that allows for the intraspecies specific differentiation of yeast into specific subgroups (biotypes) or the distinction of individual *S. cerevisiae* strains. Ideally the method should be rapid, accurate and cost effective (Buchaille*, et al.*, 1998; Fenn*, et al.*, 1994; Giusto*, et al.*, 2006; Gonzalez-Techera*, et al.*, 2001; Hierro*, et al.*, 2004; Koehler*, et al.*, 1999; Lopez*, et al.*, 2003; Manzano*, et al.*, 2006; Sherburn and Jenkins, 2003; Verweij*, et al.*, 1999). In industry certain fermentation processes can only be carried out by select strain/s of *S. cerevisiae*. In the fermentation industry, the use of starter cultures has become very common to ensure product quality and consistency. However spoilage yeasts and bacteria if undetected in starter cultures or in the fermentation can result in substantial losses in production, processing, preservation and storage. The accurate identification of *S. cerevisiae* strains is also important in industry for the effective protection of patented strains (Buzzini*, et al.*, 2007)

S. cerevisiae has traditionally been identified based on its morphological, physiological and biochemical properties (Gonzalez-Techera*, et al.*, 2001). These methods are time consuming and often unreliable (Sherburn and Jenkins, 2003). In recent years molecular biology techniques have become the methods of choice for yeast identification. However, these methods are often also laborious, unreliable and require considerable expertise (Giusto*, et al.*, 2006; Sherburn and Jenkins, 2003; Timmins*, et al.*, 1998a). MALDI-TOF MS has been shown to be an excellent tool for the unequivocal identification of undetermined clinical yeast strains at the species level (Qian*, et al.*, 2008; van Veen*, et al.*, 2010). However to date a MALDI-TOF MS study has not been carried out to extensively identify laboratory and industrial *S. cerevisiae* strains.

Previous studies on MALDI-TOF MS have largely focussed on bacteria (Giebel*, et al.*, 2008; Ilina*, et al.*, 2009; Mandrell*, et al.*, 2005; Mellmann*, et al.*, 2008) and less extensively on yeast (Sherburn and Jenkins, 2003). It has been found that characteristic signals are harder to obtain of yeast than bacteria but with optimization to current methods, yeast produce characteristic MALDI-TOF MS signals that can be used for its identification at the genus and species level (van Veen*, et al.*, 2010). The identification of *S. cerevisiae* in industry is important not only at the genus and species level but at the subspecies/strain level as well. Identification software such as the Bruker Daltonics Biotyper 3.0 software (Bruker Daltonics GmbH, Leipzig, Germany) does not identify microbes beyond the species level. The Bruker Daltonics mass spectral database was selected for this study as it is one of the most extensive mass spectral databases commercially available for biotyping microorganisms. Studies on yeast have included limited numbers of *S. cerevisiae* strains as most have focused on clinical strains such as *Candida* strains (Marklein*, et al.*, 2009; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003). Thus despite research studies conducted on the identification of *S. cerevisiae* using MALDI-TOF MS, there has not been an extensive study conducted that fully demonstrates the use of MALDI-TOF MS for the identification of industrially important *S. cerevisiae* at the subspecies/strain level. (Marklein*, et al.*, 2009; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003) .

1.2 THE AIMS OF THIS STUDY

i. To investigate several parameters of an ethanol/formic acid extraction procedure to optimize the sample preparation protocol for obtaining unique and highly reproducible MALDI-TOF MS spectra from *S. cerevisiae* strains

ii. To develop an improved sample preparation protocol to generate an *S. cerevisiae* database within a local context.

iii. To determine if MALDI-TOF MS allows for further differentiation of closely related *S. cerevisiae* and compare that with a molecular based approach. The polymerase chain reaction (PCR) using delta primers and contour-clamped homogenous electric field gel electrophoresis (CHEF) were employed as reference methods in this study.

iv. To test the accuracy of the created *S. cerevisiae* database, against a trial involving "unknown" industrial samples.

1.3 REFERENCES

Buchaille, L., Freydiere, A. M., Guinet, R. and Gille, Y. (1998). Evaluation of six commercial systems for identification of medically important yeasts. *Eur J Clin Microbiol Infect Dis* **17,** 479-88.

Buzzini, P., Turchetti, B. and Vaughan-Martini, A. E. (2007). The use of killer sensitivity patterns for biotyping yeast strains: the state of the art, potentialities and limitations. *FEMS Yeast Research* **7,** 749-760.

Chemler, J. A., Yan, Y. and Koffas, M. A. (2006). Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. *Microb Cell Fact* **5,** 20.

Dombek, K. M. and Ingram, L. O. (1987). Ethanol production during batch fermentation with *Saccharomyces cerevisiae*: changes in glycolytic enzymes and internal pH. *Appl Environ Microbiol* **53,** 1286-91.

Fenn, J. P., Segal, H., Barland, B., Denton, D., Whisenant, J., Chun, H., Christofferson, K., Hamilton, L. and Carroll, K. (1994). Comparison of updated Vitek Yeast Biochemical Card and API 20C yeast identification systems. *J Clin Microbiol* **32,** 1184-7.

Giebel, R. A., Fredenberg, W. and Sandrin, T. R. (2008). Characterization of environmental isolates of *Enterococcus spp*. by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Water Res* **42,** 931-40.

Giusto, C., Iacumin, L., Comi, G., Buiatti, S. and Manzano, M. (2006). PCR-TTGE and RAPD-PCR Techniques to analyze *Saccharomyces cerevisiae* and *Saccharomyces carlsbergensis* isolated from craft beers. *J. Inst. Brew* **112,** 340-345.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S. G. (1996). Life with 6000 genes. *Science* **274,** 546, 563-7.

Gonzalez-Techera, A., Jubany, S., Carrau, F. M. and Gaggero, C. (2001). Differentiation of industrial wine yeast strains using microsatellite markers. *Letters in Applied Microbiology* **33,** 71-75.

Hierro, N., Gonzalez, A., Mas, A. and Guillamon, J. M. (2004). New PCR-based methods for yeast identification. *J Appl Microbiol* **97,** 792-801.

Ilina, E. N., Borovskaya, A. D., Serebryakova, M. V., Chelysheva, V. V., Momynaliev, K. T., Maier, T., Kostrzewa, M. and Govorun, V. M. (2009). Application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for the study of *Helicobacter pylori*. *Rapid Commun. Mass Spectrom.* **24,** 328-334.

Kjeldsen, T. (2000). Yeast secretory expression of insulin precursors. *Appl Microbiol Biotechnol* **54,** 277-86.

Koehler, A. P., Chu, K. C., Houang, E. T. and Cheng, A. F. (1999). Simple, reliable, and cost-effective yeast identification scheme for the clinical laboratory. *J Clin Microbiol* **37,** 422-6.

Lopez, V., Fernandez-Espinar, M. T., Barrio, E., Ramon, D. and Querol, A. (2003). A new PCR-based method for monitoring inoculated wine fermentations. *Int J Food Microbiol* **81,** 63-71.

Mandrell, R. E., Harden, L. A., Bates, A., Miller, W. G., Haddon, W. F. and Fagerquist, C. K. (2005). Speciation of *Campylobacter coli, C jejuni, C. helveticus, C. lari, C. sputorum* and *C. upsaliensis* by Matrix-Assisted Laser Desorption Ioniztion-Time of Flight Mass Spectrometry. *Appl. Environ. Microbiol.* **71,** 6292-6307.

Manzano, M., Medrala, D., Giusto, C., Bartolomeoli, I., Urso, R. and Comi, G. (2006). Classical and molecular analyses to characterize commercial dry yeasts used in wine fermentation. *Food Microbiol*.

Marklein, G., Josten, M., Klanke, U., Muller, E., Horre, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A. and Sahl, H. G. (2009). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* **47,** 2912-7.

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., Dunn, J., Hall, G., Wilson, D., LaSala, P., Krostrzewa, M. and Harmsen, D. (2008). Evaluation of Matrix-Assisted Laser Desorption Ionization-Time-of-Flight Mass Spectrometry in Comparison to 16S rDNA Gene Sequencing for Species Identification of Nonfermenting Bacteria. *J. Clin. Microbiol.* **46,** 1946-1954.

Qian, J., Cutler, J. E., Cole, R. B. and Cai, Y. (2008). MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers. *Anal Bioanal Chem* **392,** 439-49.

Saitoh, S., Ishida, N., Onishi, T., Tokuhiro, K., Nagamori, E., Kitamoto, K. and Takahashi, H. (2005). Genetically engineered wine yeast produces a high concentration of L-lactic acid of extremely high optical purity. *Appl Environ Microbiol* **71,** 2789-92.

Sherburn, R. E. and Jenkins, R. O. (2003). A novel and rapid approach to yeast differentiation using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Spectroscopy* **17,** 31-38.

Timmins, E. M., Quain, D. E. and Goodacre, R. (1998). Differentiation of brewing yeast strains by pyrolysis mass specrometry and Fourier transform infrared spectroscopy. *Yeast* **14,** 885-893.

van Veen, S. Q., Claas, E. C. and Kuijper, E. J. (2010). High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *J Clin Microbiol* **48,** 900-7.

Verweij, P. E., Breuker, I. M., Rijs, A. J. and Meis, J. F. (1999). Comparative study of seven commercial yeast identification systems. *J Clin Pathol* **52,** 271-3.

Chapter 2

LITERATURE REVIEW

Identification of yeast strains

2 IDENTIFICATION OF YEAST STRAINS

2.1 INTRODUCTION

Yeasts are unicellular eukaryotes that belong to the kingdom fungi. Yeasts do not have photosynthetic abilities therefore live a saprophytic or parasitic life, depending solely on organic carbon sources provided by other organisms. The yeast *Saccharomyces cerevisiae* has been employed by mankind in food production for thousands of years. This has been documented by a series of well-preserved models of a bakery and brewery uncovered in excavations at Thebes on the Nile that date from about 2000 B.C. (Phaff*, et al.*, 1966; Rose and Harrison, 1969). Today *S. cerevisiae* plays a pivotal role in different industries such as the wine, brewing and baking industries. In recent years, *S. cerevisiae* has been developed as an efficient eukaryotic expression system for biotechnology thus is also employed in the pharmaceutical and nutraceutical industries. *S. cerevisiae* although a simple eukaryote has been used for the production of genetically engineered human insulin (Kjeldsen, 2000). *S. cerevisiae* can also be used as cell factories for the production of L-lactic acid of extremely high optical purity and (Saitoh*, et al.*, 2005) secondary metabolites, such as natural products of commercial interest. (Chemler*, et al.*, 2006).

In many instances intraspecific discrimination (biotyping) of yeast into specific subgroups (biotypes) or the distinction of individual strains is needed, especially in industry for the detection of contaminants, verification and for the protection of patented strains. Traditionally phenotypic and biochemical methods have been used for the identification of *S. cerevisiae* strains which are generally time consuming and ambiguous (Gonzalez-Techera*, et al.*, 2001). In recent years progress in molecular biology has provided a large number of DNA-based techniques for identifying and characterising yeast and in particular *S. cerevisiae* (Andrighetto*, et al.*, 2000; Gonzalez-Techera*, et al.*, 2001). Examples of these methods include determination of electrophoretic karyotypes by Pulsed-field gel electrophoresis (PFGE) (Gomes*, et al.*, 2000; Sheehan*, et al.*, 1991), DNA fingerprinting techniques(Wightman*, et al.*, 1996), PCR amplification with delta elements, intron splice site primers, or micro-satellites or simple sequence repeats (SSR) (De Barros Lopes*, et al.*, 1996; Gonzalez-Techera*, et al.*, 2001; Lopez*, et al.*, 2003; Ness*, et al.*, 1993), randomly amplified polymorphic DNA-polymerase chain reaction (RAPD-PCR) (Barszczewski and Robak, 2004; Giusto*, et al.*, 2006), mitochondrial DNA restriction fragment length polymorphism (RFLP) analysis (Barszczewski and Robak, 2004), temporal temperature gradient gel electrophoresis (TTGE) (Giusto*, et al.*, 2006), PCR-temperature gradient gel electrophoresis (PCR-TGGE), restriction enzyme analysis (RE) (Manzano*, et al.*, 2006) and real-time PCR (Logan*, et al.*, 2009). However the application of these techniques in industry can be limited by their high cost and requirement for considerable expertise.

Spectroscopic approaches such as pyrolysis-mass spectrometry (PyMS) and Fourier transform-infrared (FT-IR) spectrometry have also been used for the successful discrimination of *S. cerevisiae* strains but cumbersome data analysis and the need for skilled interpretation of the spectra are the general limitations of the whole-organism fingerprinting approach (Sheehan*, et al.*, 1991; Timmins*, et al.*, 1998b).

Matrix assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) can be used for the identification of various microorganisms (Giebel*, et al.*, 2008; Lee*, et al.*, 2008; Mandrell*, et al.*, 2005; Marklein*, et al.*, 2009; Mellmann*, et al.*, 2008; Qian*, et al.*, 2008). This approach has been used to identify bacteria to the genus, species and strain level (Ilina*, et al.*, 2009; Mandrell*, et al.*, 2005; Mellmann*, et al.*, 2008). Although the identification and characterization of bacteria using MALDI-TOF-MS has been the subject of many recent studies, the application of MALDI-TOF MS for the characterization of yeast has been limited (Qian*, et al.*, 2008; Sherburn and Jenkins, 2003). Furthermore studies on yeast using MALDI-TOF MS have focussed mainly on clinical *Candida*-yeast strains and have included very few *S. cerevisiae* strains (Amiri-Eliasi and Fenselau, 2001; Marklein*, et al.*, 2009; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003; van Veen*, et al.*, 2010). It has been reported that yeast MALDI-TOF MS signals maybe harder to obtain than bacteria and MALDI-TOF MS studies have focused on the optimization of these signals mainly for clinical strains (Qian*, et al.*, 2008). Recent studies have shown that MALDI-TOF MS can be used for the identification of yeast to the genus and species level but have not identified yeast to the strain level (Marklein*, et al.*, 2009; van Veen*, et al.*, 2010).

This literature review attempts to provide an overview of both classical and up to date methods used for *S. cerevisiae* identification.

2.2 PHENOTYPIC METHODS

Traditional methods of yeast identification were based on the shape of the cells (Fig. 2.1), thickness of the wall, lipid content, presence of vacuoles and inclusion bodies, the presence or absence of capsules and formation of ascospores (Phaff*, et al.*, 1966). Information on the cytology of yeast was gathered by:

i) Direct observation under the light microscope.
ii) Staining with specific dyes to determine the localization of certain cell components.
iii) Electron microscopy of isolated cell walls.
iv) Electron microscopy of ultra-thin sections of a yeast cell.

Traditional phenotypic methods based on the morphology of *S. cerevisiae* are still used today as a verification/confirmation test alongside many new techniques but always fail to discriminate strains within a species (Gomes*, et al.*, 2000; Gonzalez-Techera*, et al.*, 2001; Manzano*, et al.*, 2006).

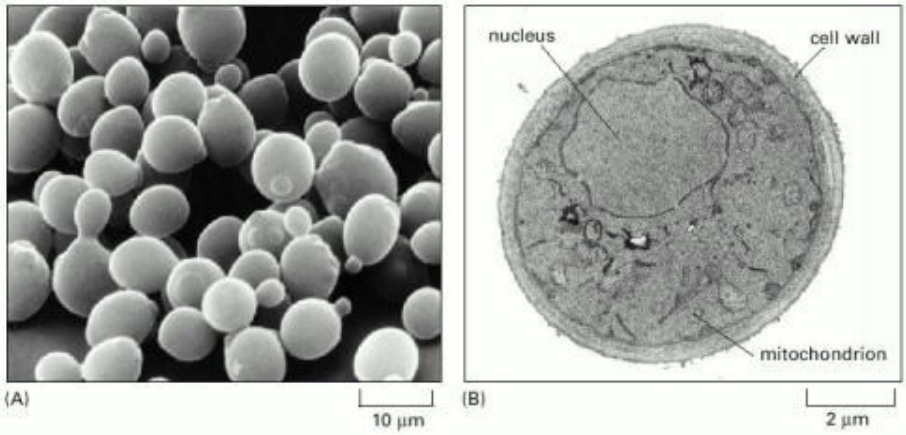


Figure 2.1 The yeast *Saccharomyces cerevisiae*. (A) Scanning electron micrograph of a cluster of the cells. This species is also known as budding yeast: it proliferates by forming a protrusion or bud that enlarges and then separates from the original cell. Many cells with buds are visible in the micrograph. (B) A transmission electron micrograph of a cross section of a yeast cell, showing its nucleus, mitochondrion, and thick cell wall (www.ncbi.nlm.nih.gov/books/NBK26909)

S. cerevisiae can also be identified using a wide range of differential media. Differential media are agar plates that allow growth of different species but contain indicator dyes that cause target organisms to appear as differently coloured colonies. The variation in colour of different species is the result of the cleavage of chromogenic substrates by species specific enzymes. In a study by Gomes *et al*. (2000) ten industrial alcohol and brewing *S. cerevisiae* strains were grown on yeast extract peptone dextrose (YEPD) basal media with the addition of nine different dyes (basic fuchsin, nile blue, bromocresol purple, aniline blue-black, rose bengal, acid fuchsin, methyl violet, crystal violet and brilliant green). This method could not differentiate between the ten industrial *S. cerevisiae* strains and further molecular techniques were required for strain identification. Commercially available chromogenic media can only offer a presumptive identification of yeast at the sub-species level. As seen in Figure 2.2, *S. cerevisiae* strains can be presumptively identified on CHROMagar Candida (CMA) as small purple colonies with a halo and for verification, colony morphology can be compared by growing strains on corn-meal-tween 80 agar (CTA) (Koehler*, et al.*, 1999). On CMA plates *S. cerevisiae* is identified by oval cells with very short multilateral budding

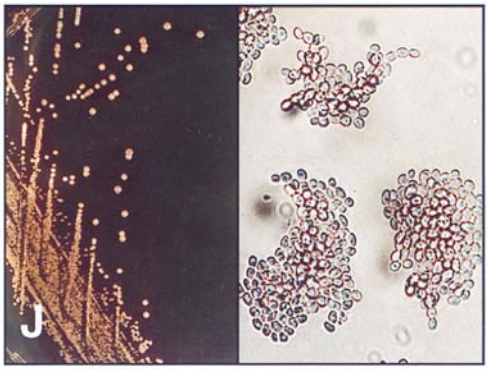


Figure 2.2 Appearance of *S. cerevisiae* on CMA (left; magnification, × 1) and microscopic appearance on CTA (right; magnification, × 400) (Koehler*, et al.*, 1999).

Other confirmation tests may also include growing *S. cerevisiae* strains on bismuth sulphite glucose glycine yeast (BIGGY) agar. *S. cerevisiae* strains can be identified as light brown colonies (Yucesoy and Marol, 2003). Phenotypic methods such as differential media are useful as a presumptive method or part of a combinatorial approach to *S. cerevisiae* identification but cannot be used routinely to differentiate *S. cerevisiae* strains at the species and sub-species level.

2.3 BIOCHEMICAL METHODS

Yeast can be identified using many biochemical assays that are based on carbohydrate fermentation and assimilation, cycloheximide (actidione) resistance, nitrate assimilation, growth at 37 ˚C, urease production on Christensen urea slant agar, and tetrazolium reduction (Buchaille*, et al.*, 1998). Biochemical methods can be used in conjunction with other methods to identify *S. cerevisiae* but like phenotypic methods are tedious and it can take a number of days to reach a conclusive result (Gomes*, et al.*, 2000; Gonzalez-Techera*, et al.*, 2001; Manzano*, et al.*, 2006).

2.4 COMMERCIAL IDENTIFICATION KITS

Classical methods of yeast identification are time-consuming and labour-intensive. Therefore commercially available systems for rapid yeast identification have been developed. The following kit systems will be described:

i) The Vitek system is an automated system designed by National Aeronautics and Space Administration (NASA) to detect and identify pathogenic microorganisms (Stager and Davis, 1992). The Vitek card comprises of 30 wells which contain 26 biochemical broths and four negative control broths. The cards are incubated at 30˚C for 24-48 hours. Colour changes are determined spectrophotometrically and are reported as a nine digit code. The database contains 36 yeast species (Fenn*, et al.*, 1994).
ii) The ApiID 32C and 20C AUX products contain 32 and 20 microtubes, respectively. The turbidity of the cultures are assessed after 24, 48 and 72 hours of incubation at 30 ˚C. A numerical profile of 10 and seven digits respectively can then be deduced. The databases of the ApiID 32 and Api20 AUX include 63 yeast species and 43 yeast species respectively (Fenn*, et al.*, 1994; Verweij*, et al.*, 1999) .
iii) The Yeast Star system relies on the inhibitory effects of specific dyes on yeast growth. A panel of six dyes is placed on top of an inoculated solid growth medium and incubated for 24 to 48 hours at 37˚C. The results of which are recorded as a six digit code. The database contains 16 yeast species (Verweij*, et al.*, 1999).
iv) Auxacolour is an identification system that is based on carbohydrate assimilation, and growth is visualised by colour change of a pH indicator. The microtitre plate contains 16 wells; incubated at 30˚C and read after 24 to 72 hours. Various other characteristics are also taken into account to reach a final identification. These characteristics include the ability to grow at 37˚C, formation of mycelium or artrospores and the formation of a capsule. The database contains 26 yeast species (Buchaille*, et al.*, 1998; Verweij*, et al.*, 1999).

v) The RapID Yeast Plus system is based on carbohydrate substrate utilisation and hydrolysis of fatty acids, glycoside and aryl-amide. A panel of 18 cavities is incubated for four hours at 30 ˚C. A six digit code is derived from the results and compared with those of a yeast database which comprises of 43 yeast species (Buchaille*, et al.*, 1998) .

vi) The Api Candida consists of 10 tubes containing dehydrated substrates and relies on sugar acidification or enzymatic reactions. The strips are assessed after 24 and 48 hours of incubation at 30˚C. A four digit numerical profile is obtained and compared to a yeast database containing 26 yeast species (Buchaille*, et al.*, 1998) .

In a comparative study of commercially available identification kits by Verweij *et al*. (1999) it was found that Auxacolour and Api Candida showed the highest percentage of correctly identified yeast at the species level. Buchaille *et al*. (1998) found that using the Auxacolour and Api Candida systems only 80% of *S. cerevisiae* strains were identified. In contrast, using the RapID-Yeast Plus systems, 90% of *S. cerevisiae* strains in the same study were identifiable to the species level. In a study by Fenn *et al*. (1994), using the Vitek system, *S. cerevisiae* strains were identified with a percentage accuracy of 89% (42 out of 47 strains) and employing the API20C kit, 98% of the strains were identified. Therefore while commercially available identification kits are widely used, they are often expensive and are limited by the sizes of their strain identification databases. Thus identification of uncommon *S. cerevisiae* strains using commercial kits is very difficult and there is no single identification kit available that will differentiate all *S. cerevisiae* strains. Furthermore some of these commercially available kits need morphological assessments to complement final identification (Buchaille*, et al.*, 1998; Verweij*, et al.*, 1999). Commercial kits are designed mostly for clinical yeast strains and are not ideally suited for the identification of laboratory and industrial *S. cerevisiae* strains (Wenning*, et al.*, 2002).

2.5 KARYOTYPING

The advent of electrophoretic techniques for the separation of intact chromosomal DNA molecules of lower eukaryotes has provided a novel means of characterising the chromosome sets of these organisms. Yeast karyotyping can be used for yeast species identification(www.ncbi.nlm.nih.gov/books/NBK26909, accessed on 1 December 2010) and strain identification (Sheehan*, et al.*, 1991).

2.5.1 Chromosomal gel electrophoretic techniques

Pulsed-field gel electrophoresis (PFGE) is employed to determine the genotype of industrial *S. cerevisiae* strains (Dawkins, 1989). There are various configurations of pulsed-field gel electrophoresis systems available. These systems are named according to the positioning of the electrodes relative to the sample on the gel (Fig. 2.3). For example transverse alternating pulsed-field gel electrophoresis (TAFE), orthogonal field alternation gel electrophoresis (OFAGE), field inversion gel electrophoresis (FIGE) and contour-clamped homogenous electric field electrophoresis (CHEF).

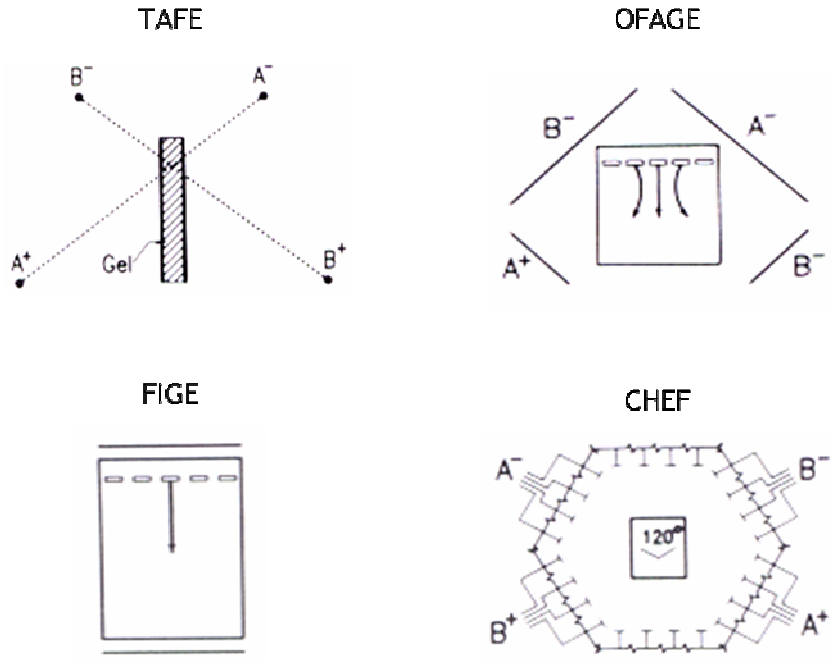


Figure 2.3 Schematic diagram of pulsed field gel systems, transverse alternating pulsed-field gel electrophoresis-TAFE, orthogonal field alternation gel electrophoresis-OFAGE, field inversion gel electrophoresis-FIGE and contour-clamped homogenous electric field electrophoresis-CHEF (www.bio.davidson.edu/Courses/Molbio/MolStudents).

The CHEF electrophoretic system employs a hexagonal electrode array around the gel periphery with a horizontal slab gel. The electrode array consists of 20 electrodes in predetermined locations. The switching system controls the pulse reorientation and sets up the voltage gradient with the remaining electrodes. By combining pulse reorientation, the DNA molecules reorient as they travel across the gel (Dawkins, 1989).

The karyotype of *S. cerevisiae comprises* of sixteen chromosomes which range in size from approximately 200-2500 kb (Goffeau*, et al.*, 1996). Sheehan and co-workers (1991) reported that the electrophoretic karyotype of all sixteen chromosomes could be obtained using CHEF gel electrophoresis. They also found that CHEF gel electrophoresis was able to differentiate between closely related lager strains. Although CHEF gel electrophoresis can be used for the differentiation of industrial *S. cerevisiae* strains, it should be noted that karyotype analysis is difficult, expensive, time consuming and has very low reproducibility (Gomes*, et al.*, 2000).

2.6 MOLECULAR GENETIC METHODS

Recent developments in molecular biology have allowed for the creation of new yeast identification techniques based on DNA analysis. The structure of DNA (sequences) do not vary with physiological states, thus these identification techniques appear to be more reliable than traditional microbiological methods (De Barros Lopes*, et al.*, 1996; Giusto*, et al.*, 2006; Gonzalez-Techera*, et al.*, 2001; Hierro*, et al.*, 2004; Lopez*, et al.*, 2003; Ness*, et al.*, 1993; Wightman*, et al.*, 1996).

2.6.1 Radio-labelled DNA probes

DNA fingerprinting techniques unlike traditional microbiological methods offers a more precise, rapid and objective approach to the differentiation of *S. cerevisiae* strains. Early attempts to use DNA fingerprinting for industrial yeast differentiation failed due to the use of hazardous and unstable nature of the radioactive (^{32}P labelled) probes. However, different strains were distinguishable using this technique. Furthermore, the radioactive probe patterns tended to diffuse thereby affecting resolution and interpretation of data.

The later development of non-radioactive techniques for Southern hybridizations improved the potential of DNA fingerprinting and was employed for the analysis of brewing strains of yeast (Wightman*, et al.*, 1996). *S. cerevisiae* strains used in the production of beer were analysed by DNA fingerprinting, using a Southern blotting and hybridization procedure and employing the TY1-15 transposon as a probe. The technique allowed for the differentiation of all brewing strains used in the study. A vacuum blotting device was used and was found to greatly improve the quality of the Southern blots compared to conventional capillary transfer. The efficiency of DNA transfer was found to be shorter and the resolution of the bands sharper. However a major shortcoming of the technique was that the ability to readily differentiate between strains was very dependent on the restriction enzyme used to digest the DNA prior to Southern blotting and hybridization. Furthermore, the random fragment length polymorphisms seemed minor and were not easily observable (Wightman*, et al.*, 1996).

2.6.2 PCR-based identification techniques

The development of polymerase chain reaction (PCR) based techniques, has opened new avenues for the identification of closely related *S. cerevisiae* strains (Lopez*, et al.*, 2003). The PCR involves the amplification of specific sequences within target DNA. One of the remarkable attributes of the method is that the sequence of the target need not be known, all that is required is knowledge of the flanking sequences. A PCR is carried out using a reaction mixture that consists of a solution containing the target sequence, four deoxyribonucleotide triphosphates, a pair of primers and Taq polymerase (a heat stable DNA polymerase from the bacterium *Thermus aquaticus*). The reaction mixture is then placed into a thermal cycler that is programmed to adjust the temperature to allow for amplification of a particular DNA sequence (Berg*, et al.*, 2003).

The amplification of DNA by PCR occurs in three steps. Strand separation involves separation of the parent DNA by heating the reaction mixture to approximately 94˚C. This is followed by a hybridization step wherein the temperature is rapidly reduced to approximately 54˚C. The single stranded DNA primers hybridize to their complementary target sequences forming double stranded DNA. The next reaction involves DNA synthesis which is initiated when the temperature is increased to 72˚C. Employing the target sequence as a template the Taq polymerase extends the primer to synthesise a complementary DNA strand. The temperature of

the reaction is elevated to the 95˚C so as to separate the newly synthesized and template DNA strands. The specificity of PCR for the identification of *S. cerevisiae* strains depends on both the degree of homology between the primers and target DNA (Berg*, et al.*, 2003). The amplified DNA fragments are then electrophoretically separated on an agarose/polyacrylamide gel followed by staining with ethidium bromide and visualised under ultraviolet irradiation.

2.6.2.1 Retrotransposon specific primers

Ness and co-workers (1993) developed a PCR-based method using specific primers that target the delta (δ) chromosomal elements. Delta elements are direct repeat 300 bp sequences that flank the Ty1 retrotransposon. In laboratory strains the Ty1 element is dispersed throughout the yeast genome and is present in about 35 copies. Ness, *et al*. (1993) used consensus sequences that were previously determined to be complimentary to the 5' termini of δ elements as primers. It was found that the amplification of yeast genomic DNA with the δ1 and δ2 primers yields amplified sequence polymorphisms which are strain specific and stable under classical conditions of cell multiplication. Twenty-six industrial yeast strains were analysed and all strains from different origins were found to display different profiles. The technique was also found to be highly reproducible as identical profiles were obtained for identical strains (Ness*, et al.*, 1993).

2.6.2.2 Intron specific primers

Many eukaryotic mRNA precursors contain intervening sequences or more commonly referred to as introns which are removed or spliced from pre-mRNA to form mature mRNA, a process which requires a large ribonucleoprotein called the spliceosome. Introns are not essential for gene function and the reason for existence still remains unclear. Introns have evolved with minimal constraint, except for the sites that have been conserved for recognition by the spliceosome, their sequences are highly variable (De Barros Lopes*, et al.*, 1996).

De Barros Lopes *et al*. (1996) were able to design intron primers that could be used to differentiate between different commercial wine yeast strains. They designed two primer pairs (Fig. 2.4), one was based on the consensus sequences GTATGT at the 5' exon-intron splice site (EI1 and E12) and the other was based on the lariat branch point consensus sequence TACTAAC (LA1 and LA2). The commercial wine yeast strains used in this study except for one were previously subjected to chromosome karyotyping using transverse alternating field electrophoresis (TAFE). It was found that the PCR-based method using intron splice sites was much simpler, faster and less labour intensive. In some instances polymorphisms observed in certain strains were minor and the use of different primers or a pair of primers was required for greater separation of the strains. Delta sequences are not present in non-*Saccharomyces* yeasts unlike intron splice sites that are conserved within all yeasts thus the intron primers are not limited in that they can be used to differentiate all yeast. A highly osmotolerant fermenting yeast *T. delbrueckii* which is used occasionally for the production of sweet wines was shown to

have a completely different amplification pattern when compared with the other *S. cerevisiae* strains using intron primers which could not have been achieved using delta primers. Therefore intron specific primers are more suited for yeast differentiation than specific primers that target the delta elements of transposons.

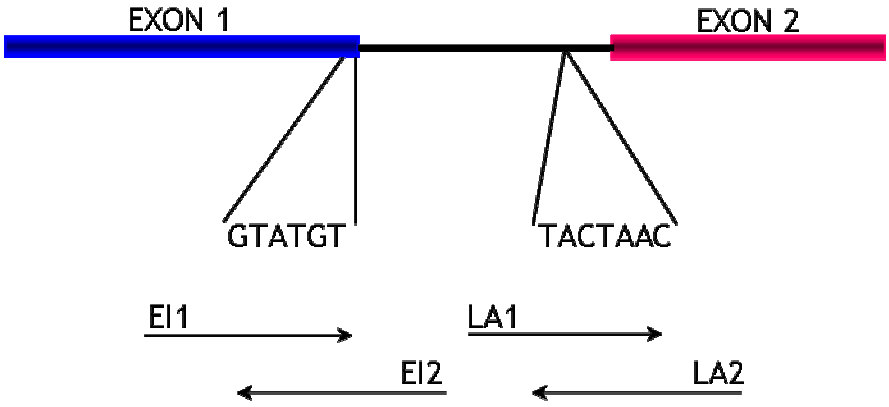


Figure 2.4 Design of intron primers. Primers EI1 and EI2 are based on the consensus sequences GTATGT at the 5' exon-intron splice site and primers LA1 and LA2 are on the lariat branch point consensus sequence TACTAAC (adapted from De Barros Lopes*, et al.*, 1996).

2.6.2.3 *COX1*-specific primers

Polymerase chain reaction-based approaches (using δ and intron specific primers) are fast and simple but a major drawback of both these methods are poor reproducibility of banding patterns due to variations in the quantity and quality of template DNA and the low annealing temperature of the techniques. With this in mind, Lopez et al. (2003) proposed a new PCR-based method for the rapid differentiation of yeast. This technique is based on the extremely intron-rich content of the mitochondrial *COX1* gene which encodes the largest sub-unit of the cytochrome C oxidase. It has been found that the number of introns and their position can vary between yeast species. There is also evidence that strains of *S. cerevisiae* are polymorphic for *COX1*. For example *S. cerevisiae* FY1679 was found to possess seven introns while strain D273-10B possessed six.

Lopez *et al*. (2003), designed primers by comparing available sequences of the *COX1* gene from different yeasts and ascomycetes. It was found that using twelve primers in different combinations could not yield polymorphic patterns for the thirteen wine yeast strains employed in this study. However, it was observed that greater polymorphisms could be achieved by employing four selected primers in multiplex-PCR amplification reactions. This PCR-based method allowed for the differentiation of thirteen *S. cerevisiae* wine yeast strains. Unlike other PCR-based methods which require isolated chromosomal DNA as a template, this technique is especially beneficial as the fermenting grape medium can be used directly in the PCR reaction. The ability to use the must directly in the PCR reaction could allow wine-makers to implement effective combative measures should a problem of contamination arise.

2.6.2.4 PCR-amplified ribosomal RNA gene fragments

Gene sequencing is not an ideal tool for the identification of fermenting yeast because it is very costly, labour intensive and time consuming. Alternative approaches to gene sequencing are randomly amplified polymorphic DNA (RAPD) and restriction fragment length polymorphism (RFLP) patterns of PCR-amplified ribosomal RNA gene fragments (rDNA). In a study by Barszczewski and Robak (2004), RFLP-PCR of rDNA was found to have shortcomings when employed to differentiate wild yeast isolates from a starter culture used in Polish lager breweries. The three restriction enzymes one five-base (*ScrFI*) and two four-base cutting *HaeIII* and *MspI* enzymes, commonly used in the restriction analysis of fungal rDNA did not allow for the differentiation between all the strains tested. Nineteen different restriction patterns were obtained for the 27 strains tested. Single strain determination was only possible after the comparison of PCR-RFLP patterns with profiles obtained in RAPD analysis. Later it was suggested by the same authors that RFLP-PCR rDNA analyses should be used to show similarities in closely related *S. cerevisiae* strains, while RAPD analyses could be used to differentiate these strains (Barszczewski and Robak, 2006).

Giusto *et al*. (2006) have shown that both temporal temperature gradient gel electrophoresis (TTGE) and the RAPD techniques were useful for identifying *S. cerevisiae* strains isolated from craft beer. Furthermore results were obtainable within 8 hours after cell growth on isolation media. The TTGE technique allowed for the differentiation of yeast strains used in the production of numerous types of beer. The data indicates that using the TTGE technique, a single point mutation in a 200-800 bp sequence was recognisable, thereby demonstrating the potential of the technique to be used for microorganism differentiation at the strain level. The TTGE generated data was confirmed using RAPD analysis. The RAPD technique is considered to be simple, rapid and no knowledge about the genome sequence is needed. However, it was mentioned that RAPD-PCR is not always reproducible and interpretation at times was complicated owing to inconsistent band intensities. Due to different thermocyclers being employed difficulties are encountered in inter-laboratory comparisons of RAPD banding patterns.

2.6.2.5 PCR-temperature gradient gel electrophoresis (PCR-TGGE)

Manzano *et al*. (2006) proposed using PCR-temperature gradient gel electrophoresis (PCR-TGGE) and restriction enzyme analysis (RE) to identify commercially available starter cultures of *S. cerevisiae* as well as to detect possible dry contaminant microorganisms (non-*S. cerevisiae* yeasts). In this study an attempt was made to biotype a collection of sixty two commercial enological dry yeasts. The TGGE technique revealed that species names as declared by dry yeast producers were incorrect. This may have been due to dry yeasts producers relying on physiological methods or morphological methods for yeast identification and not confirming their results using molecular techniques. Furthermore strains of *S. cerevisiae* and *S. bayanus var. uvarum* are expected to express different enological properties which may

lead to an overall different wine sensory profile with an overall negative impact. Using restriction enzyme analysis *S. cerevisiae* and *S. bayanus var. uvarum* were successfully differentiated. Although restriction enzyme analysis could not differentiate between *S. paradoxus* and *S. bayanus var. uvarum* the results obtained using this method supported those obtained using the TGGE protocol. The TGGE protocol is a very useful method for the quick identification of dry yeasts as the method does not require the culturing of a pure colony which may reduce the time required for strain identification.

2.6.2.6 Microsatellite marker specific primers

Microsatellite markers or simple sequence repeats comprise of direct tandem repeats of a short DNA motif which is usually less than 10 bp in length. Microsatellites are a major component of DNA of higher organisms and are hypervariable in length due to replication errors such as slipped-strand mispairing. Microsatellites display a substantial level of polymorphism between individuals of the same species. Gonzalez-Techera *et al*. (2001), demonstrated that cultivation of six Uruguayan indigenous winery and three commercial wine *S. cerevisiae* strains on WL Nutrient medium did not allow for adequate phenotypic differentiation of these strains. A PCR-based technique using two microsatellite markers was able to unequivocally differentiate all nine wine yeast strains. The use of specific primers coupled with high annealing temperatures revealed microsatellite polymorphism analysis to be a very reproducible method for the amplification of short DNA fragments. Furthermore the method could potentially be automated for the analysis of multiple samples and multiplexed for the analysis of multiple microsatellite loci. Microsatellite analysis using the DNA isolation technique as described by Gonzalez-Techera *et al*. (2001) is a fairly rapid technique and data is available within 14 hours.

2.6.2.7 Real-time PCR

Real-time PCR is a technique that greatly simplifies amplicon recognition by providing a means of monitoring the accumulation of specific products throughout the amplification process. Real-time PCR instrumentation measures the progress of amplification by monitoring changes in fluorescence within the PCR vessel. There are two methods that are used to obtain a fluorescent signal (Logan*, et al.*, 2009).

i) The first method is based on the property of fluorescent dyes such as sybr green I to bind to double stranded DNA and undergo a conformational change that results in an increase in fluorescence.
ii) The second approach involves the use of fluorescent resonance energy transfer (FRET). A major advantage of this technology lies in its ability to characterise the amplicon *in situ*. This is achieved by analysis of the melting temperature and/or probe hybridization characteristics of the amplicon within the PCR reaction (Logan*, et al.*, 2009).

With the aid of the non-specific fluorescent dye SYBR green, Martorell *et al*. (2005) developed a real-time PCR method to detect and enumerate *S. cerevisiae* with DNA that was extracted directly from wine samples. Data was available within five hours of commencement of this procedure. Real time PCR has an advantage over other PCR-based identification strategies as it avoids restriction endonuclease digestion and electrophoretic steps.

2.7 MASS SPECTROMETRY

The ability to separate charged particles based on mass charge was established in 1897 when J.J Thomson in a pioneering experiment measured the mass to charge ratio of the electron (Tjernberg, 2005). A mass spectrometer is an analytical instrument that is used to determine the molecular mass of free ions and consists of three primary components (Fig. 2.5).

i) The ion source for the vaporization and ionization samples.
ii) A mass analyser for the separation of ions based on their mass to charge ratio.
iii) A detector for measurement of the separated ions.

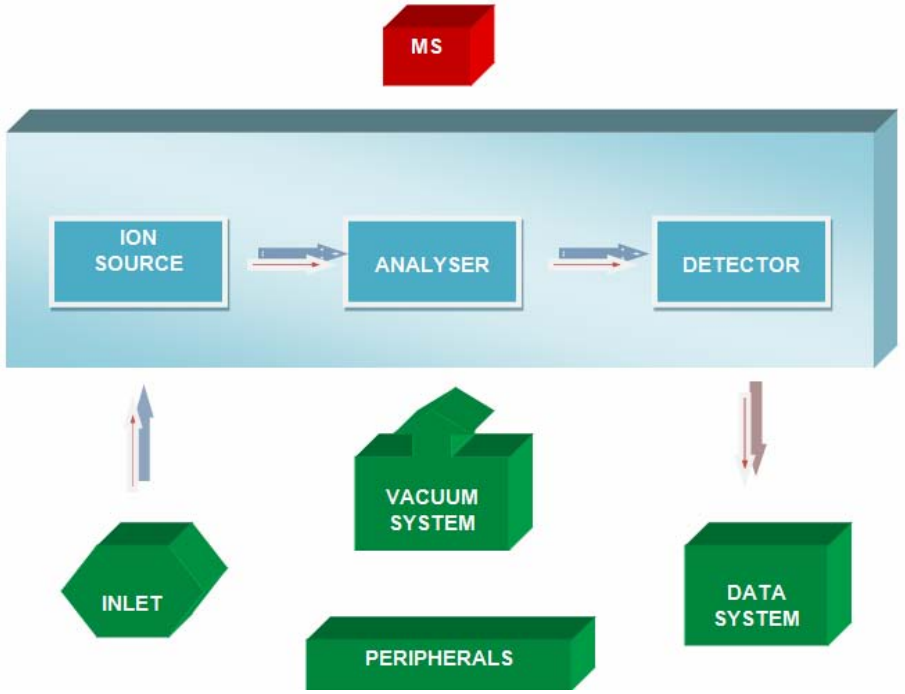


Figure 2.5 Primary components of a mass spectrometer (redrawn from Wilkins and Lay, 2006).

2.7.1 Pyrolysis Mass Spectrometry (PyMS) and Fourier Transform Infrared (FT-IR) spectroscopy

Pyrolysis is a chemical process that involves the thermal degradation of complex organic material such as yeast cells, in an inert atmosphere or vacuum. The thermal energy causes the molecules to cleave at their weakest points to produce smaller, volatile fragments called pyrolysate. The resultant pyrolysate is characteristic of the original material. In Curie-point

pyrolysis the sample is first dried onto a suitable metal and then heated rapidly to the Curie-point of that metal. The constituents of the resultant pyrolysate are ionized and separated by mass spectrometery based on their mass/charge (m/z) ratio (Fig. 2.6). A pyrolysis mass spectrum is produced that represents a chemical profile or "fingerprint" of the analyzed material. The resultant spectra are analyzed using suitable multivariate statistical methods (Wilkins and Lay, 2006).

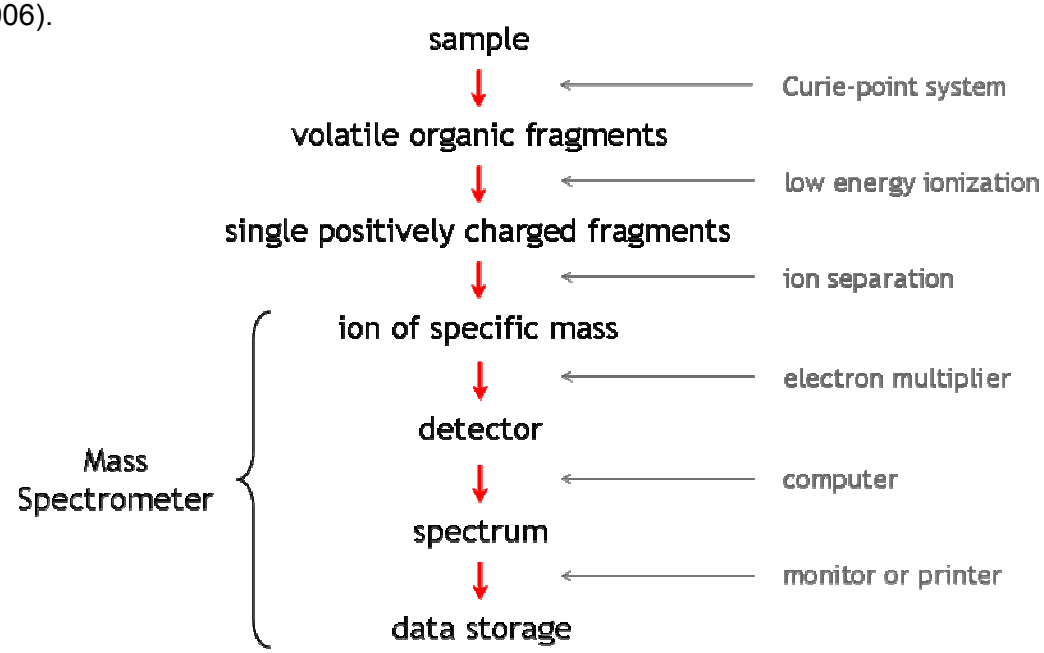


Figure 2.6 The stages of pyrolysis mass spectrometry (adapted from Wilkins and Lay, 2006).

Fourier-transform infrared (FT-IR) spectroscopy involves the absorption of infrared light by cellular compounds which results in a fingerprint-like spectrum that can be identified by comparison to reference spectra. In FT-IR spectroscopy micro-colonies are transferred to an IR-transparent Zinc Selenide (ZnSe) carrier by means of a stamping device, resulting in spatially accurate replicas (Fig. 2.7). The air-dried colonies on the ZnSe carrier are then detected using a computer-driven xy stage and a light microscope. The spectra of single colonies are then recorded by a mid-IR spectrometer which is coupled to a microscope (Wenning, *et al.*, 2002).

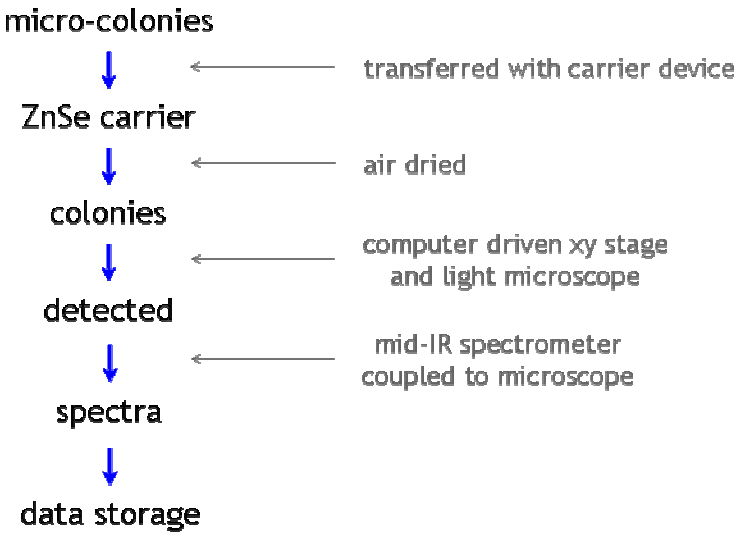


Figure 2.7 Events of Fourier-Transform Infrared (FT-IR) spectroscopy (Wenning, *et al.*, 2002).

The PyMS and FT-IR spectrometric methods were employed in the differentiation of twenty two brewery yeast strains (Timmins*, et al.*, 1998b). The result from this investigation was compared to previous genotypic investigations of the same isolates. In addition, the yeast strains were grown on different media types to assess the effect of growth media on spectra. It was observed that both PyMS and FT-IR spectrometric spectra for ale and lager yeast strains grown on different media were very similar. Therefore multivariate statistical techniques were employed for the analysis of both the PyMS and FT-IR spectral data. The pseudo-3-D discriminant function analysis (DFA plot) of PyMS spectral data clearly revealed that cultivating yeast on different media does indeed cause a change in the PyMS spectra. Two clusters were seen on the DFA plot which did not mirror one another. The one cluster consisted of all yeast strains cultivated in liquid media and the other reflected yeast strains cultivated on solid media. It was also observed that dendograms yielded by either PyMS or FT-IR spectrometry of all twenty two strains when cultivated in liquid medium closely correlated with the dendrogram obtained by DNA analysis from a previous study.

The dendrogram created from PyMS data showed four main clusters. The ale strains showed more diversity than the lager strains. This was expected as lager yeasts represent a comparatively homogenous group of yeast strains. The data showed good differentiation of the yeasts, although there was no clear separation between the lager and the ale strains which was also the case in DNA analysis. The DFA plot from the infrared data also showed good discrimination. Most of the results obtained through DNA analysis were in agreement with the results obtained from both the PyMS and FT-IR methods. Therefore PyMS and FT-IR are useful for strain discrimination and in practice could be concurrently used with other procedures to confirm that the correct strain is being used in a brewery. Both these techniques have many advantages in terms of sensitivity and identification can be completed within 24 hours. The general limitations of these methods are the cumbersome data analysis and the need for skilled interpretation of spectra (Timmins*, et al.*, 1998b).

2.7.2 Matrix assisted laser desorption ionization (MALDI) spectrometry

In 1988 the inventors of MALDI, Michael Karas and Franz Hillenkamp presented their first MALDI spectrum (Fig. 2.8) of a protein with a mass greater than 1000 Daltons at the International Mass spectrometry Conference in Bordeaux (Hillenkamp and Karas, 2000).

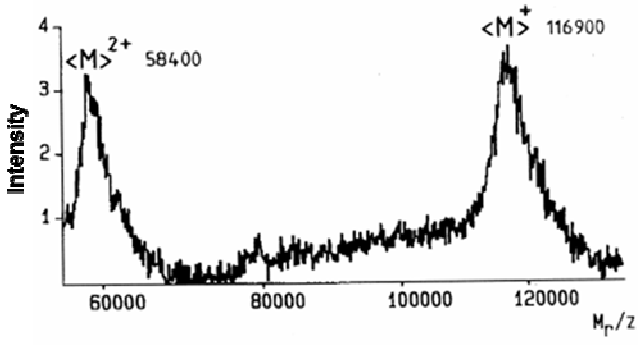


Figure 2.8 MALDI spectrum at 266 nm of β-D-galactosidase employing nicotinic acid as the matrix and a sum of 100 shots (adapted from Hillenkamp and Karas, 2000).

Matrix assisted laser desorption/ionization spectrometry has the advantage that both the vaporization and ionization procedures can be accomplished in a single step. Of significance and a major achievement in mass spectrometry, the MALDI technique can be employed in the analysis of fragile biomolecules. To obtain a mass spectrum, the solid phase (or liquid phase) must first be converted to gas phase ions. Thus the first step in the MALDI process involves desorption (vaporization) and ionization (Tjernberg, 2005).

Traditionally vaporization was accomplished by heat-induced evaporation of the sample. The MALDI process is different as many samples in the high vacuum environment of the mass spectrometer (approximately 10^{-6} torr) sublime directly into the source with little or no heating necessary. Conventionally ionization has been achieved by bombarding vaporized samples with electrons in a process called electron bombardment (EB) or electron impact (EI) ionization. The kinetic energy of the 70 eV electron is much higher than typical ionization potentials (10-15 eV) which invariably results in the ejection of a secondary electron and the formation of a radical cation. The radical cation (molecular ion) is extremely unstable and loses its excess energy by fragmentation giving rise to lower mass fragments. The fingerprints produced are only suitable for the identification of smaller organic molecules (<1000 Da) and cannot be used to identify larger molecules such as proteins. However, the MALDI process has an advantage over these methods as it provides a “soft” or non-destructive (at the molecular level) ionization mechanism of converting solid phase analytes into gas phase ions with minimal fragmentation (Wilkins and Lay, 2006). Fragmentation can lead to considerable loss of spectra quality such as loss of resolution or even complete loss of the parent ion (Hillenkamp and Peter-Katalinic, 2007).

In the MALDI process (Fig. 2.9), the analyte is embedded in a very large excess of matrix material. The matrix is a weak acid that absorbs strongly at the wavelength of the laser thus there is a strong interaction between the matrix and analyte components. Three commonly employed matrix compounds include 2, 5-dihydroxy-benzoic acid (DHBA), sinapinic acid (SA) and α-cyano-4-hydroxy-cinnamic acid (HCCA). As illustrated in Table 2.1, these compounds all have an aromatic ring structure and contain a carboxylic acid moiety (Hillenkamp and Peter-Katalinic, 2007).

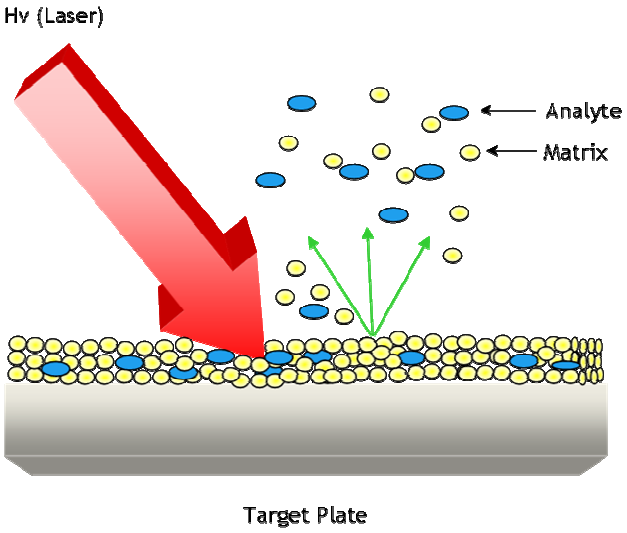
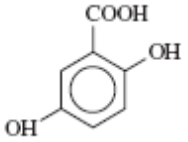
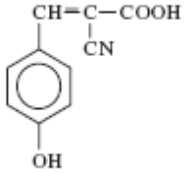
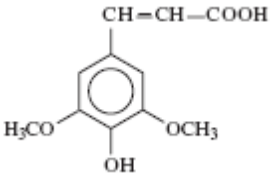


Figure 2.9 Principle of the MALDI process (redrawn from Wilkins and Lay, 2006)

Table 2.1 Matrices used in MALDI (adapted from Hillenkamp and Peter-Katalinic, 2007).

Matrix	Structure	Wavelength	Major applications
2,5-Dihydroxybenzoic acid		UV 337 nm, 353 nm	Proteins, peptides, carbohydrates, synthetic polymers
α- Cyano-4-hydroxycinnamic acid		UV 337 nm, 353 nm	Peptides, fragmentation
Sinapinic acid		UV 337 nm, 353 nm	Proteins and peptides

Approximately 1 µl of sample solution is applied onto a MALDI target plate which is then dried and overlaid with an equal volume of matrix solution. The target spot is then allowed to dry and placed into the mass spectrometer's ion source. After a short laser pulse, the irradiated spot is rapidly heated and becomes vibrationally excited. The laser energy is absorbed by the matrix, and the vibrational energy results in the expulsion of material in a plume just above the irradiated site. Desorption is achieved under high vacuum. The expanding plume of matrix compound "carries" the analyte into the gas phase with very little heating of the sample. Therefore intact proteins can be transferred into the gas phase without decomposition which occurs in heat-induced evaporation. The MALDI process does result in the transfer of both ionic and neutral species into the vapour state. The matrix as mentioned is a weak acid and some of the polar analytes (like proteins) attract protons from the carboxylic acid groups of the matrix in an acid-base equilibrium process. This process can occur in both the solid phase or by gas phase-proton transfer. In both instances gas phase-protonated analytes are formed. The mass spectrometer does not detect abundant desorbed neutrals; furthermore the region that corresponds to the matrix is typically much lower in mass to charge ratio (m/z) value than that of the analyte. Therefore the abundant protonated matrix species can be omitted by not sampling ions within this m/z range (Wilkins and Lay, 2006).

The MALDI process is ideally suited for the analysis of proteins because they are good proton acceptors and are much higher in mass than protonated matrix, or protonated matrix cluster ions. MALDI is also well suited for the analysis of mixtures of proteins (including enzymatic digests) because the technique produces one large ion (protonated molecule) per component.

2.7.3 Time-of-flight (TOF) mass spectrometer

Experimental data obtained using a TOF MS instrument called the Velocitron was initially reported in 1948. The early Velocitron and modern linear TOF instruments have a simple design and are very similar (Wilkins and Lay, 2006). The three principle regions (Fig. 2.10) of these instruments are as follows:

i) An ion source
ii) A drift region (mass analyser)
iii) A detector (Fig. 2.10)

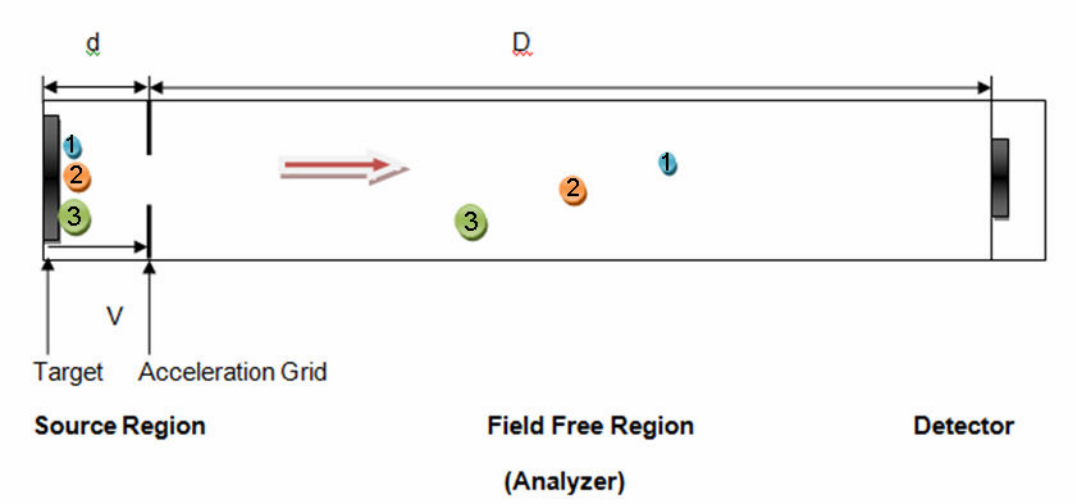


Figure 2.10 The principle of a time-of-flight mass spectrometer where: ion source region (d), distance (D) between the ion source and detector and voltage (V). The ions of lower mass, ion 1 will travel and strike the detector first followed by larger ion 2 and lastly ion 3 (redrawn from Wilkins and Lay, 2006).

There are a range of mass spectrometers that can be used to analyse mass ions, these range from inexpensive ion trap mass, which generate poor quality spectra, to Fourier transform mass spectrometers (FT-MS), which generally produce extremely high quality spectra but are very expensive. The most common type of analyzer for MALDI ions, is the time of flight mass spectrometer (Hillenkamp and Peter-Katalinic, 2007).

The TOF mass spectrometer is operated in pulsed mode which means each spectrum is generated from a discrete “event” whereby a packet of ions is formed in a plume and accelerated (Wilkins and Lay, 2006). A typical MALDI ion source generates a pulse of ions at most a few nanoseconds in duration at a rate of 1 to 100 pulses per second (Hillenkamp and Peter-Katalinic, 2007). All of the ions of the same sort (positively- or negatively-charged) in the plume are then separated based on their mass to charge ratio and strike the detector at a specific time interval thereby yielding a recordable high sensitivity mass spectrum. This is called “separation in time”. An alternative to “separation in time” is the “separation in space”. Separation in space is used in magnetic and quadrupole-based mass analysers to differentiate one m/z value (in a certain location or space) from all others by discarding all the non-selected

values. The difference between separation in “time” and “space” is one of the major differences between TOF mass spectrometers and other mass spectrometers. Using TOF mass spectrometers the entire process can be rapidly repeated. In a few seconds many mass spectra can be achieved, which can then be averaged to produce a final recorded mass spectrum with an enhanced signal to noise ratio (Wilkins and Lay, 2006).

2.7.4 Spectra generation using a Time-of-flight (TOF) mass spectrometer

Ions need to be generated in a very specific region of the ion source as ions formed in different regions would have a differing accelerating potential. Therefore desorption/ionization coupled to TOF-MS, usually occurs on a stainless steel target plate onto which a sample/s have been deposited. Within the ion source, the ions are accelerated across a potential difference (V) created in this region. The voltage difference is the result of different potentials applied to the target and an acceleration grid at opposite ends of the source region. The voltage creates an electric field that imparts to the ions energy of motion equal to the product of the number of charges on each ion (z), the magnitude of the charge of a single electron (e) and the overall voltage difference over the region designated d in (Fig. 2.10). If no collisions occur within the drift region, ion flight times can be directly related to the mass to charge value (Wilkins and Lay, 2006). Therefore the following equations can be implemented:

$$\text{Energy} = zeV \tag{1}$$

This energy of motion is known as kinetic energy (KE) and is given by the equation

$$\text{KE} = \frac{mv^2}{2} \tag{2}$$

Therefore;

$$\frac{mv^2}{2} = zeV \tag{3}$$

In equation (3) mass and velocity are represented as m and v respectively. Mass spectrometry data (mass spectra) are usually plotted with the m/z ratio on the x-axis and the ion abundance on the y-axis. Thus solving for the mass to charge value the following equation can be obtained;

$$\frac{m}{z} = \frac{2eV}{v^2} \tag{4}$$

Since the value for the charge of an electron is a constant and in many experiments the voltage applied across the ion source region is also a constant. Equation (4) can be simplified further to yield;

$$\frac{m}{z} = \frac{2eV}{v^2} \tag{5}$$

or

$$v = \sqrt{\frac{kz}{m}} = \frac{k}{\sqrt{m/z}} \tag{6}$$

Equation (4) shows that ions leave the source region with velocity values inversely related to the square root of their m/z values. If no collisions occur, the ions will maintain a constant velocity across the field free region right to the detector. A very convenient way of discerning mass to charge values is to measure the velocity values indirectly. Since the detector is at a fixed location within the mass spectrometer at distance D from the ion source (Fig. 2.5) the time taken by ions to reach the detector can be related to velocity to give the equation

$$t = \frac{D}{v} \tag{7}$$

Substituting for v, D is also recognised as a constant;

$$t = \frac{D}{v} = \frac{D\sqrt{m}}{\sqrt{k'z}} = k'\sqrt{\frac{m}{z}} \tag{8}$$

From equation (6) the ion arrival times at the detector can be related to the square root of their m/z values. In order to obtain the m/z values we can simply rearrange the equation to give;

$$\frac{m}{z} = k't^2 \tag{9}$$

It is possible to substitute the constant with the appropriate values for constituent parameters $(2eV/D^2)$, it is more practical to simply determine the value empirically. A well known calibration equation for TOF mass spectrometry is a little more complicated (with the addition of an "offset" m/z value, b), and can be written as;

$$\frac{m}{z} = kt^2 + b \tag{10}$$

The values of k (sometimes denoted as a) are determined using spectra obtained with known mass standards (Wilkins and Lay, 2006). Mass accuracy is limited by the quality of the calibration and how accurately the centre of the peak can be determined (Hillenkamp and Peter-Katalinic, 2007). An example (Table 2.2) for the purpose of biotyping a Bruker bacterial standard with known protein mass standards is used for calibration.

Table 2.2 Bruker Daltonics bacterial standard incorporating known mass standards

Protein	Mass (Daltons)
RL 36 $[M+H]^+$	**4365.30000**
RS32 $[M+H]^+$	**5096.80000**
RS34 $[M+H]^+$	**5381.40000**
RS33 meth $[M+H]^+$	**6255.40000**
RL29 $[M+H]^+$	**7274.50000**
RS19 $[M+H]^+$	**10300.10000**
RNase A $[M+H]^+$	**13683.20000**
Myoglobin $[M+H]^+$	**16952.30000**

Mass resolution can be defined as the mass spacing at which spectral features (peaks) can be separated, and is reported in Daltons but most often the term “resolution” refers to the “resolving power”. The resolving power is measured as $m/\delta m$ where m is the mass of the peak centre and δm the resolution (which is often defined as the width of a peak at half height or Full Width at Half Maximum (FWHM). The resolving power is a unit-less parameter. The typical resolving power of a linear TOF mass spectrometer is around 300 to 400. The linear TOF mass spectrometer with its simple design can easily produce resolved spectra of small molecules and can also yield “spectral fingerprints” for bacteria, but cannot resolve unit mass differences such as the masses of carbon isotopes much above 300 to 500. The resolving power can be increased by as much as an order of magnitude but it requires modifications to be made to the simple design of the linear TOF mass spectrometer (Hillenkamp and Peter-Katalinic, 2007; Wilkins and Lay, 2006).

One major concern for TOF mass spectrometers is pressure and ion collisions. During desorption and ionization the density of matter within the resulting plume can be relatively high for a very short time. The brief laser pulse desorbs both ions and neutral species at the same time and due to the high density of material in the plume collisions result in errors in velocity and hence measured mass values. The error causes a noticeable reduction in the resolving power of the mass measurement. Ions that have the same m/z value reach the detector at different times, giving rise to very broad peaks. As peak width increases, the number of discrete peaks that can be measured in a mass spectrum decreases. If the ions accelerating from the target out of the ion source were delayed for a brief period of time, collisions would less likely occur, and the resolving power of the mass measurement would be greatly improved. This technique, of inserting a brief time delay between ion formation and ion acceleration, is called delayed extraction. Delayed extraction has allowed MALDI ionization to be coupled with TOF mass spectrometers but does not significantly increase resolution. However, a slightly more complicated design for TOF mass spectrometers has been found to significantly increase their resolution (Hillenkamp and Peter-Katalinic, 2007; Wilkins and Lay, 2006).

2.7.5 Reflectron mass spectrometer

Within the same vacuum system of a linear TOF, it is possible to essentially double the field-free region (Fig. 2.11). This is achieved by “reflecting” the ions back toward the source (field-free region 1 + field-free region 2). Employing this design, ions travel from the source region towards the conventional detector (Detector 1) but are not detected. Instead the ions are reflected by an opposing field back toward the source region (V_2). This effectively doubles the length of the field-free region thereby increasing the maximum achievable resolving power by a factor of approximately 2. The ions enter a region with a retarding potential that is greater than the acceleration potential in the ion source region. In this region the ions “turn around” and move back towards the second detector. Since ions can be reflected at an angle of slightly less than 180˚, the second detector is situated adjacent to the ion source region. This design or arrangement produces what is called a reflectron TOF mass spectrometer (Hillenkamp and Peter-Katalinic, 2007; Wilkins and Lay, 2006).

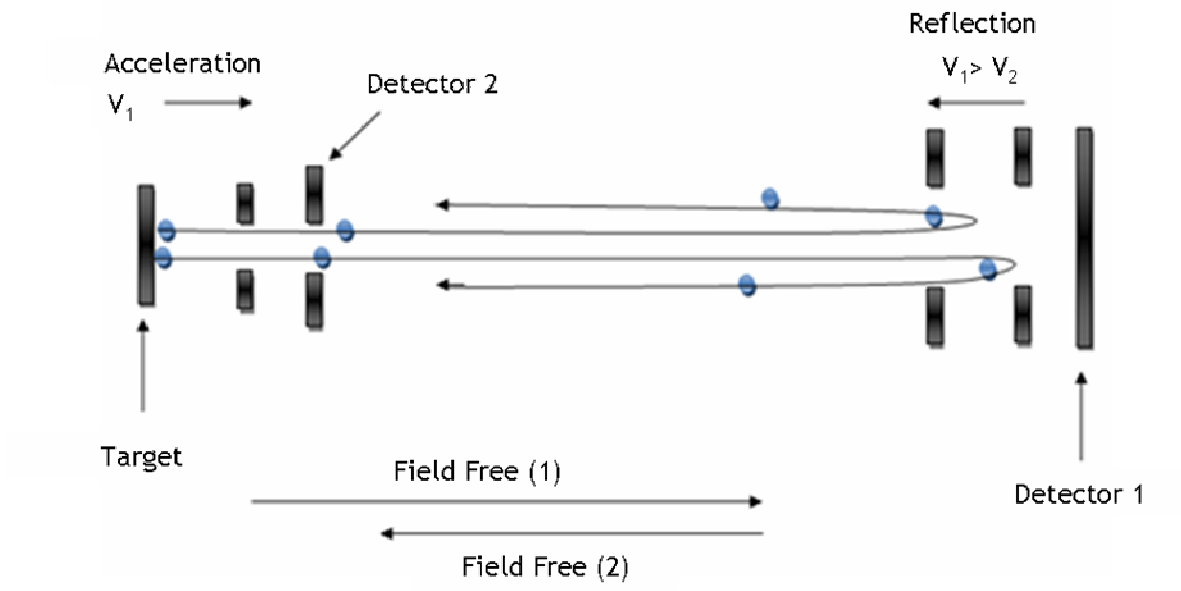


Figure 2.11 Principle of a reflectron time-of-flight mass spectrometer (adapted from Wilkins and Lay, 2006).

The reflectron mass spectrometer design does not only have an advantage over the linear TOF in terms of doubling the length of the field-free region, the retarding field may also act to somewhat refocus the ion beam. Ions that have a small amount of excess velocity for a given mass penetrate slightly more into the field before they are reflected, whereas ions with slightly less velocity with identical mass do not penetrate as much. The velocity imparted in the reverse direction (leaving the reflectron) depends on the depth of penetration of the ions at the exact moment they cease their forward motion when they briefly stop. Thus they are reflected with slightly different energies and it is these differences that cancel out the small differences in velocities that they had when entering the reflecting region. The refocusing capability of the reflectron mass spectrometer allows it to have resolving powers of greater than 5000 (Wilkins and Lay, 2006).

Traditional mass spectrometers (magnetic quadruple and ion trap) differentiate only one m/z value from others by simply discarding all ions with non-selected m/z values. TOF mass spectrometers are more sensitive as ions are separated in time and all are eventually detected. The abundances of proteins in whole cells are not as great as in other components thus sensitivity is an important factor. Furthermore, TOF is also capable of detecting high mass ions which makes it a more superior mass spectrometer for the purpose of biotyping (Wilkins and Lay, 2006).

2.8 MALDI-TOF MS-FACILITATED BIOTYPING OF MICRO-ORGANISMS

Since the introduction of MALDI-TOF MS in the late 1980's, MALDI-TOF MS has been used for a variety of different applications(Amiri-Eliasi and Fenselau, 2001; Chaurand, *et al.*, 2002; Deininger, *et al.*, 2008; Marklein, *et al.*, 2009; Qian, *et al.*, 2008). In recent years a number of research groups have demonstrated the use of MALDI-TOF MS to identify and characterize

bacteria (Giebel, *et al.*, 2008; Ilina, *et al.*, 2009; Mandrell, *et al.*, 2005; Mellmann, *et al.*, 2008). The application of this powerful tool to characterise yeast has received far less attention (Amiri-Eliasi and Fenselau, 2001; Qian, *et al.*, 2008; Sherburn and Jenkins, 2003). In the following review, the identification of bacteria and yeast species identification will be discussed.

2.8.1 Bacteria

MALDI-TOF MS has emerged as a technology that can be used to analyze the protein composition of a bacterial cell and differentiate bacteria at the species level. This is achieved by measuring the exact sizes of peptides and proteins that are characteristic of a particular bacterial species. MALDI-TOF MS biotyping of bacteria can be used in a wide range of applications, a few of which will be explained briefly.

2.8.1.1 Bacterial source tracking (BST)

When recreational waters are contaminated it is up to public health officials to determine the source of the contamination and resolve the problem early enough to avoid exposure or continual exposure of bathers to pathogenic microorganisms. Bacterial source tracking (BST) is often used to determine the source of faecal contamination of recreational waters. There are several methods such as phenotypic (e.g. antibiotic sensitivity testing) and genotypic (e.g. ribotyping, repetitive sequence PCR, 16S rDNA analysis, pulse field gel electrophoresis) methods that can be used as BST tools however most of these methods are complex, costly, and require days to complete.

MALDI-TOF MS was examined as a potential method for the characterization of environmental isolates of *Enterococcus* which is an indicator organism of faecal contamination. Although requiring further development, it was found that MALDI-TOF MS-based fingerprinting of *Enterococcus* had enormous potential as a rapid and accurate BST tool. The method allowed for the grouping of isolates from sources including humans, gulls, chickens and canines. The proposed technique includes a 20 hour lysozyme digestion step that contributes to an overall BST diagnosis time of two days. This is a major improvement on several other PCR-based BST approaches which can be completed in 3-4 days (Giebel, *et al.*, 2008).

2.8.1.2 Pathogenic non-fermenting bacteria

Although non-fermenting bacteria are ubiquitous in the environment; they can cause severe infections especially in immuno-compromised patients. Accurate species identification is extremely critical because the prognosis of an infected patient depends on exact species identification. Moreover, in certain cases such as patients with cystic fibrosis, strict infection control measures have to be implemented as non-fermenting bacteria are the main cause of morbidity and mortality in these patients. MALDI-TOF was evaluated as a method to identify non-fermenting bacteria. Using MALDI-TOF MS spectra it was found that 82.5% of blind coded

non-fermenting bacterial isolates from clinical specimens were identifiable to the species level when compared to a spectral database of well-characterized non-fermenting culture collection. In addition 95.2% of the isolates were identified to the genus level. In a previous study that employed the commercial API 20NE and Vitek 2 phenotypic identification systems, only 62% (API 20NE) and 54 % (Vitek 2) of the strains were correctly identified. The above data clearly indicates that MALDI-TOF MS is a more accurate identification method when compared to commercially available identification kits. In addition MALDI-TOF MS was shown to identify more non-fermenting bacteria to the species level than the reference partial 16S rRNA gene sequencing method. Furthermore the MALDI-TOF MS method provided a fast and accurate means of identification for non-fermenting bacteria (Mellmann*, et al.*, 2008).

2.8.1.3 *Campylobacter*

Campylobacter strains are thermophilic bacteria that account for a high percentage of food-borne illnesses. Most illnesses occur sporadically and are caused primarily by *C. jejuni*. Other gastrointestinal pathogens include *C. coli*, *C. lari*, *C. fetus*, *C. upsaliensis*, *C. sputorum*, *C. concisus* and *C. curvus*. There are numerous assays that can be used to verify *Campylobacter* by genus and species but these methods have various limitations in terms of being labour intensive and sometimes unreliable. It has been demonstrated that *Campylobacter* species grown on solid media under a variety of conditions can be analyzed using MALDI-TOF mass spectrometry to yield high-intensity intact proteins (9-14 kDa range) that are diagnostic of *Campylobacter* species. Using MALDI-TOF MS, pure cultures of *Campylobacter* species were identified rapidly and with less ambiguous results than many other methods used currently for the confirmation of this species. An additional advantage of MALDI-TOF MS is that multiple species of *Campylobacter* that are present in mixed cultures can be identified more easily than conventional methods (Mandrell*, et al.*, 2005).

2.8.1.4 *Helicobacter*

Helicobacter pylori which possesses a very high natural variability causes a chronic low-level inflammation of the stomach lining and is strongly linked to the development of duodenal and gastric ulcers, and stomach cancer. Finding a rapid accurate method for the identification and epidemiological characterization of these bacteria may aid researchers to better understand *H. pylori's* transmission pathways and virulence mechanisms. Recently MALDI-TOF MS was employed to characterize *Helicobacter pylori* (Ilina*, et al.*, 2009). From the collected mass spectra, seven to thirteen significant peaks were observed and six protein signals were identical for almost half of the analysed strains. Four of these protein signals were identified as ribosomal proteins (RL32, RL33, RL34, and RL36). In addition the MALDI Biotyper 2.0 software programme was implemented for the identification of *H. pylori*. Characteristic main spectra (20 spectra per sample) were created for each of the clinical *H. pylori* strains and compared separately to the MALDI Biotyper library containing 3287 main spectra of bacterial strains from

more than 1200 different species that included two *H. pylori* strains (26695 and J99). Even though differences in MALDI profiles were observed for the clinical isolates, all tested clinical isolates were correctly identified. Thus it was concluded that MALDI-TOF MS could be used for the rapid and accurate identification of highly variable *H. pylori* strains.

2.8.2 Yeasts

Although MALDI-TOF MS has been shown to be a rapid method used to identify and characterize bacteria, its application as a tool to identify and differentiate fungal cells has been largely neglected. Yeast do appear to yield characteristic MALDI-TOF MS signals but the signals may be more difficult to obtain than from bacteria (Qian*, et al.*, 2008; Sherburn and Jenkins, 2003). This may be attributed to the highly complex cell wall structure of yeast which apparently limits intimate contact of surface components with the MALDI matrix crystals, which results in low-count ion signals and therefore poor quality mass signatures (Qian*, et al.*, 2008). To enhance this contact many researchers have attempted using intact cells while others have developed optimized cell lysis techniques and use the cell extract for MALDI-TOF MS analysis which will be discussed further below.

2.8.2.1 MALDI-TOF analysis of intact yeast cells

Sherburn and Jenkins (2003) conducted a MALDI-TOF MS study to develop a method for generating genus, species and strain specific biomarkers from yeast cell wall components. Pre-treatment of yeast samples with lyticase (β1, 3-glucanase) was introduced to partially digest the glucan cell wall so as to release other cell wall components that could be ionised by MALDI. The yeast cell wall proteins are covalently linked to β (1,6)-linked glucan, which in turn is linked to β (1,3)-linked glucan and to a lesser extent chitin as shown in Figure 2.12. The lyticase enzyme is a β (1,3)-glucanase with residual β (1,6)-glucanase activity . Several other parameters were also investigated and optimized which included matrix selection, matrix solvent and ratio of sample to matrix. The matrix chemical SA was selected over DHB which confirms the earlier findings of Amiri-Eliasi and Fenselau, (2001). The spectra generated in this study were reproducible, showing that with the aid of an optimized procedure both *Candida* species and *S. cerevisiae* strains could be differentiated at the species and strain level using MALDI-TOF MS. It should be noted that 100 units of lyticase was required and the yeast cells had to be pre-treated with lyticase for 15 minutes. Therefore although the method may produce characteristic MALDI-TOF MS spectra, lyticase would make the MALDI procedure more expensive. Furthermore only three *S. cerevisiae* strains (NCYC 22118, NCYC 19201 and RR1) were evaluated in this study. The method did not show whether *Candida* species or *S. cerevisiae* strains could be identified to the species or strain level.

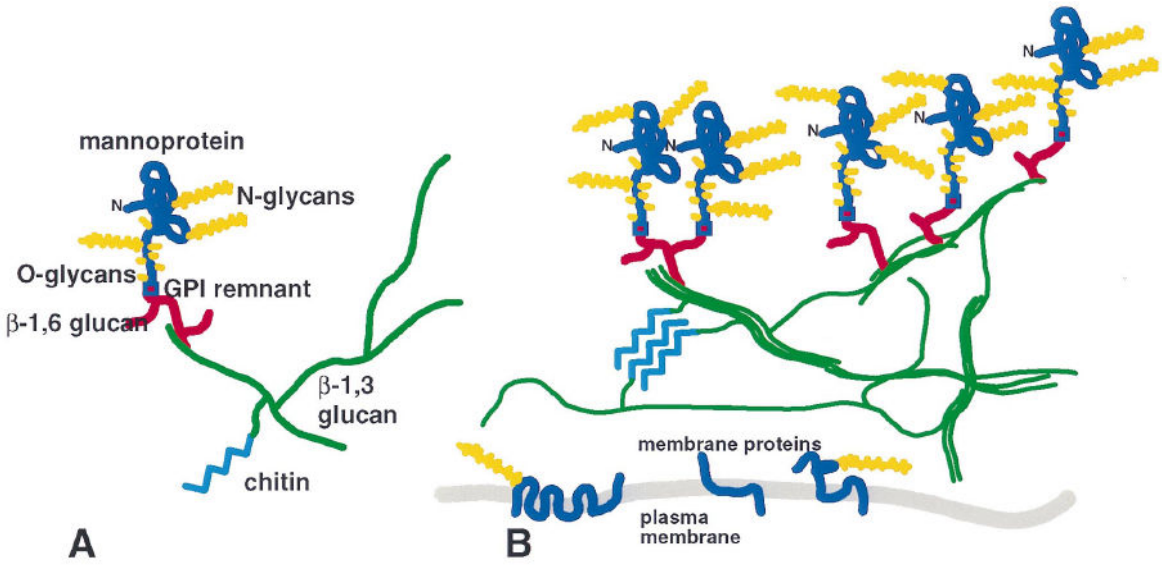


Figure 2.12 Components of an *S. cerevisiae* cell wall (A) the mannoprotein polypeptide is blue, and oligosaccharides are shown in yellow, labelled N- or O-linked. Two branch points of glucans are shown. Chitin can also be linked to β (1, 6)-linked glucan. (B) Association of modules to form a wall lattice, the colours are the same as A. The crosslinking of mannoproteins are not shown (adapted from Lipke and Ovalle, 1998).

van Veen *et al*. (2010) found that yeast could not be identified directly by MALDI-TOF MS and had to undergo pre-treatment. Sherburn and Jenkins,(2003) also found that they could not obtain spectra unless first pre-treating yeast cells with lyticase and optimizing certain conditions.

2.8.2.2 MALDI-TOF analysis of yeast protein extracts

Amiri-Eliasi and Fenselau (2001) employed ultraviolet-MALDI-TOF MS in an attempt to differentiate *Candida albicans* (ATCC 14053), *Epidermophyton floccosum* (ATCC 9646) and *S. cerevisiae* (ATCC 26108). In this study it was observed that the choice of cell wall lysis technique played an important role in determining the number of proteins desorbed, as well as the signal-to-noise ratio of the spectra. Several cell wall lysis techniques were explored which included exposing the cells to high concentrations of acids, zymolase, protein extraction reagent, ultrasonication, glass beads and corona discharge. The best spectra were yielded on treatment of all three isolates with 25 % formic acid for 5 minutes. In addition three different matrices were evaluated as a suitable MALDI matrix, these included DHB, HCCA and SA. It was observed that higher quality spectra were obtained when using sinapinic acid in comparison to the other two matrices.

Bioinformatics software was used to further compare the MALDI-TOF MS acquired protein masses for *S. cerevisiae* to protein masses generated *in silico* from protein databases. Most of the proteins that were tentatively identified were basic, four were ribosomal and five

were mitochondrial proteins. The detected ribosomal proteins were not unexpected because ribosomes constitute a large portion of the mass of a growing cell and most of the proteins are strongly basic. The detection of mitochondrial proteins was significant as any observation of biomarkers from cytosolic organelles indicates that the cell wall is effectively being lysed. This study showed that *S. cerevisiae* (ATCC 26108) using MALDI-TOF MS could be differentiated from two different species using a 25 % formic acid cell wall lysis method. The study did not confirm if MALDI-TOF MS could be used to differentiate *S. cerevisiae* strains or be used for its identification at the genus or species as only one *S. cerevisiae* strain was employed.

Marklein *et al*. (2009) studied the use of MALDI-TOF MS for the identification of clinical yeasts and yeast-like fungi. For the first time the technique was used to identify over 250 clinical isolates, with part of the study running under routine conditions. MALDI-TOF MS proved to be a rapid reliable procedure for identification of pathogenic yeast but only to the species level. Unlike the study by Qian *et al*. (2008) the amount of biological material used was not standardised, cells from up to five representative single colonies were transferred to a 1.5 screw-cap extraction tube and subjected to the standard Bruker ethanol/formic acid extraction procedure which is also used on bacterial cells. Qian *et al*. (2008) have reported that ethanol is more suited for the extraction of protein form bacterial cells and that methanol should be used in the pre-treatment of yeast for MALDI-TOFMS analysis. It was found that the procedure took 10 minutes per isolate and 3 hours per 96 samples, starting from a single colony on the agar plate. Furthermore overall there were no misidentifications to the species level as long as the reference strains were present in the database. It was observed that incorrect species assignments as “best hits” were only obtained when the strain was absent from the database which could clearly be seen by the low score values according to the MALDI Biotyper software. Log (score) values range from 0 to 3. Values of > 1.7 generally indicate a relationship at the genus level; > 2.0 show a relationship at the species level. The highest log (score) of a match against the database was used for species identification. Marklein *et al*. (2009) reported that the MALDI software was user-friendly and no specialist mass spectrometry training was required, after a few days of training technicians were able to run the system. Furthermore the investment and maintenance of the instrument was balanced by the low consumable costs.

In a recent study by van Veen *et al*. (2010) it was found that the application of MALDI-TOF MS was dependent on the reference strains included in the database. This was in agreement with an earlier study by Marklein *et al*. (2009). van Veen *et al*. (2010) found that the pre-treatment of yeast was essential for obtaining a correct identification and also used an ethanol/formic acid extraction procedure. Like Marklein *et al*. (2009) the amount of biological material used was not standardised; a single colony was used for MALDI-TOF MS analysis. It was found that neither culture medium or incubation temperature, nor length of incubation affected the accuracy of MALDI-TOF MS identifications. This was contrary to what had been previously reported by Qian *et al*. (2008) which claim that the effects of cell incubation conditions are strain dependent. In this study a total of 980 clinical isolates of bacteria and yeast were tested by conventional identification systems and MALDI-TOF MS. It was found that

MALDI-TOF MS performs better than conventional biochemical methods. van Veen *et al*. (2010) correctly identified 96.7% of yeasts to the genus level and identified 85.2 % of yeasts to the species level but 0% to the strain level.

2.9 CONCLUSION

It can thus be concluded that yeast do produce characteristic MALDI-TOF MS signals however due to the structure of the yeast cell wall, these signals may be more difficult to obtain than from bacteria (Qian*, et al.*, 2008). For the technique to be successful, the spectra obtained have to be both diagnostic and reproducible. Therefore certain standard operating procedures need to be formulated (Marklein*, et al.*, 2009; Qian*, et al.*, 2008). Studies have shown that identification of *S. cerevisiae* to the genus and species level is possible (Amiri-Eliasi and Fenselau, 2001; Marklein*, et al.*, 2009; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003) but further studies need to be conducted to determine the level of accuracy achievable at the strain level as well. Studies on the identification of yeast using MALDI-TOF MS have largely concentrated on clinical yeast strains and very few *S. cerevisiae* strains were included. Thus there is a need for a more extensive study on *S. cerevisiae* using MALDI-TOF MS which plays such a pivotal role in industry (Chemler*, et al.*, 2006; Dombek and Ingram, 1987; Kjeldsen, 2000; Saitoh*, et al.*, 2005). There is one major limiting aspect of MALDI-TOF MS for *S. cerevisiae* identification and it is the need for a comprehensive database (Marklein*, et al.*, 2009; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003) . The Bruker Daltonics mass spectral database currently stands at 3740 microorganisms of which only seven are *S. cerevisiae* strains thus a more comprehensive *S. cerevisiae* database is critical in an endeavour to efficiently biotype these industrially important microbes.

2.10 REFERENCES

.

Amiri-Eliasi, B. and Fenselau, C. (2001). Characterization of Protein Biomarkers by MALDI from whole fungal cells. *Anal. Chem* **73,** 5228-5231.

Andrighetto, C., Psomas, E., Tzanetakis, N., Suzzi, G. and Lombardi, A. (2000). Randomly amplified polymorphic DNA (RAPD) PCR for the identification of yeasts isolated from dairy products. *Lett Appl Microbiol* **30,** 5-9.

Barszczewski, W. and Robak, M. (2004). Differentiation of contaminating yeasts in brewery by PCR-based techniques. *Food Microbiology* **21,** 227-231.

Barszczewski, W. and Robak, M. (2006). PCR-Based Differentiation and Homology of Brewing and Type strains of Genus *Saccharomyces*. *J. Inst. Brew* **112,** 165-172.

Berg, J. M., Tymoczko, J. L. and Stryer, L. (2003). *Biochemistry*, Fifth edn. W.H Freeman and Company.

Buchaille, L., Freydiere, A. M., Guinet, R. and Gille, Y. (1998). Evaluation of six commercial systems for identification of medically important yeasts. *Eur J Clin Microbiol Infect Dis* **17,** 479-88.

Chaurand, P., Schwartz, S. A. and Caprioli, R. M. (2002). Imaging mass spectrometry: a new tool to investigate the spatial organization of peptides and proteins in mammalian tissue sections. *Curr Opin Chem Biol* **6,** 676-81.

Chemler, J. A., Yan, Y. and Koffas, M. A. (2006). Biosynthesis of isoprenoids, polyunsaturated fatty acids and flavonoids in *Saccharomyces cerevisiae*. *Microb Cell Fact* **5,** 20.

Dawkins, H. J. (1989). Large DNA separation using field alternation agar gel electrophoresis. *J Chromatogr* **492,** 615-39.

De Barros Lopes, M., Soden, A., Henschke, P. A. and Lang, P. (1996). PCR Differentiation of Commercial Yeast Strains Using Intron Splice Site Primers. *Appl. Environ. Microbiol.* **62,** 4514-4520.

Deininger, S. O., Ebert, M. P., Futterer, A., Gerhard, M. and Rocken, C. (2008). MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *J Proteome Res* **7,** 5230-6.

Dombek, K. M. and Ingram, L. O. (1987). Ethanol production during batch fermentation with *Saccharomyces cerevisiae*: changes in glycolytic enzymes and internal pH. *Appl Environ Microbiol* **53,** 1286-91.

Fenn, J. P., Segal, H., Barland, B., Denton, D., Whisenant, J., Chun, H., Christofferson, K., Hamilton, L. and Carroll, K. (1994). Comparison of updated Vitek Yeast Biochemical Card and API 20C yeast identification systems. *J Clin Microbiol* **32,** 1184-7.

Giebel, R. A., Fredenberg, W. and Sandrin, T. R. (2008). Characterization of environmental isolates of *Enterococcus spp*. by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Water Res* **42,** 931-40.

Giusto, C., Iacumin, L., Comi, G., Buiatti, S. and Manzano, M. (2006). PCR-TTGE and RAPD-PCR Techniques to analyze *Saccharomyces cerevisiae* and *Saccharomyces carlsbergensis* isolated from craft beers. *J. Inst. Brew* **112,** 340-345.

Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. and Oliver, S. G. (1996). Life with 6000 genes. *Science* **274,** 546, 563-7.

Gomes, L. H., Duarte, K. M. R., Argueso, J. L., Echeverrigaray, S. and Taveres, F. C. A. (2000). Methods of yeast characterization from industrial products. *Food Microbiology* **17,** 217-223.

Gonzalez-Techera, A., Jubany, S., Carrau, F. M. and Gaggero, C. (2001). Differentiation of industrial wine yeast strains using microsatellite markers. *Letters in Applied Microbiology* **33,** 71-75.

Hierro, N., Gonzalez, A., Mas, A. and Guillamon, J. M. (2004). New PCR-based methods for yeast identification. *J Appl Microbiol* **97,** 792-801.

Hillenkamp, F. and Karas, M. (2000). Matrix-assisted laser desorption/ionisation, an experience. *Int. J. Mass Spectrom.* **200,** 71-77.

Hillenkamp, F. and Peter-Katalinic, J. (2007). A Practical guide to instrumentation, methods and applications, Wiley-VCH.

Ilina, E. N., Borovskaya, A. D., Serebryakova, M. V., Chelysheva, V. V., Momynaliev, K. T., Maier, T., Kostrzewa, M. and Govorun, V. M. (2009). Application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for the study of *Helicobacter pylori*. *Rapid Commun. Mass Spectrom.* **24,** 328-334.

Kjeldsen, T. (2000). Yeast secretory expression of insulin precursors. *Appl Microbiol Biotechnol* **54,** 277-86.

Koehler, A. P., Chu, K. C., Houang, E. T. and Cheng, A. F. (1999). Simple, reliable, and cost-effective yeast identification scheme for the clinical laboratory. *J Clin Microbiol* **37,** 422-6.

Lee, F. W., Ho, K. C. and Lo, S. C. (2008). Rapid identification of dinoflagellates using protein profiling with matrix-assisted laser desorption/ionization mass spectrometry. *Harmful Algae* **7,** 551-559.

Lipke, P. N. and Ovalle, R. (1998). Cell wall architecture in yeast: new structure and new challenges. *J Bacteriol* **180,** 3735-40.

Logan, J., Edwards, K. and Saunders, N. (2009). Real-time PCR Current Technology and applications, Caister Academic press.

Lopez, V., Fernandez-Espinar, M. T., Barrio, E., Ramon, D. and Querol, A. (2003). A new PCR-based method for monitoring inoculated wine fermentations. *Int J Food Microbiol* **81,** 63-71.

Mandrell, R. E., Harden, L. A., Bates, A., Miller, W. G., Haddon, W. F. and Fagerquist, C. K. (2005). Speciation of *Campylobacter coli, C jejuni, C. helveticus, C. lari, C. sputorum* and *C. upsaliensis* by Matrix-Assisted Laser Desorption Ioniztion-Time of Flight Mass Spectrometry. *Appl. Environ. Microbiol.* **71,** 6292-6307.

Manzano, M., Medrala, D., Giusto, C., Bartolomeoli, I., Urso, R. and Comi, G. (2006). Classical and molecular analyses to characterize commercial dry yeasts used in wine fermentation. *Food Microbiol*.

Marklein, G., Josten, M., Klanke, U., Muller, E., Horre, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A. and Sahl, H. G. (2009). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* **47,** 2912-7.

Martorell, P., Querol, A. and Fernandez-Espinar, M. T. (2005). Rapid identification and enumeration of *Saccharomyces cerevisiae* cells in wine by real-time PCR. *Appl Environ Microbiol* **71,** 6823-30.

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., Dunn, J., Hall, G., Wilson, D., LaSala, P., Krostrzewa, M. and Harmsen, D. (2008). Evaluation of Matrix-Assisted Laser Desorption Ionization-Time-of-Flight Mass Spectrometry in Comparison to 16S rDNA Gene Sequencing for Species Identification of Nonfermenting Bacteria. *J. Clin. Microbiol.* **46,** 1946-1954.

Ness, F., Lavellee, F., Dubourdieu, D., Aigle, M. and Dulau, L. (1993). Identification of Yeast strains using the polymerase chain reaction. *J Sci Food Agric* **62,** 89-94.

Phaff, H. J., Miller, C. M. and Mrak, E. M. (1966). *The Life of Yeasts*, 5 edn. Harvard University Press.

Qian, J., Cutler, J. E., Cole, R. B. and Cai, Y. (2008). MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers. *Anal Bioanal Chem* **392,** 439-49.

Rose, A. H. and Harrison, J. S. (1969). *The Yeasts*. Academic Press.

Saitoh, S., Ishida, N., Onishi, T., Tokuhiro, K., Nagamori, E., Kitamoto, K. and Takahashi, H. (2005). Genetically engineered wine yeast produces a high concentration of L-lactic acid of extremely high optical purity. *Appl Environ Microbiol* **71,** 2789-92.

Sheehan, C. A., Weiss, A. S., Newsom, I. A., Flint, V. and O'Donnell, D. C. (1991). Brewing yeast identification and chromosome analysis using high resolution CHEF gel electrophoresis. *J. Inst. Brew* **97,** 163-167.

Sherburn, R. E. and Jenkins, R. O. (2003). A novel and rapid approach to yeast differentiation using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Spectroscopy* **17,** 31-38.

Stager, C. E. and Davis, J. R. (1992). Automated systems for identification of microorganisms. *Clin Microbiol Rev* **5,** 302-27.

Timmins, E. M., Quain, D. E. and Goodacre, R. (1998). Differentiation of brewing yeast strains by pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. *Yeast* **14,** 885-93.

Tjernberg, A. (2005). Protein mass spectrometry in the drug discovery process Karolinska University Press.

van Veen, S. Q., Claas, E. C. and Kuijper, E. J. (2010). High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *J Clin Microbiol* **48,** 900-7.

Verweij, P. E., Breuker, I. M., Rijs, A. J. and Meis, J. F. (1999). Comparative study of seven commercial yeast identification systems. *J Clin Pathol* **52,** 271-3.

Wenning, M., Seiler, H. and Scherer, S. (2002). Fourier-transform infrared microspectroscopy, a novel and rapid tool for identification of yeasts. *Appl Environ Microbiol* **68,** 4717-21.

Wightman, P., Quain, D. E. and Meaden, P. G. (1996). Analysis of production brewing strains of yeast by DNA fingerprinting. *Lett Appl Microbiol* **22,** 90-4.

Wilkins, C. L. and Lay, J. O. (2006). *Identification of Microorganisms by mass spectrometry*. John Wiley and sons.

www.bio.davidson.edu/Courses/Molbio/MolStudents (accessed on 28 November 2010).

www.ncbi.nlm.nih.gov/books/NBK26909 (accessed on 1 December 2010).

Yucesoy, M. and Marol, S. (2003). Performance of CHROMAGAR candida and BIGGY agar for identification of yeast species. *Ann Clin Microbiol Antimicrob* **2,** 8.

Chapter 3

RESEARCH RESULTS I

Optimisation of sample preparation protocols for MALDI-TOF MS identification of *Saccharomyces cerevisiae* strains

Optimisation of sample preparation protocols for MALDI-TOF MS identification of *Saccharomyces cerevisiae* strains

Anushka Moothoo-Padayachie [1], Thavendran Govender [2], Hendrik G. Kruger,[3] Glenn E. M. Maguire[3] and Patrick Govender.[1]

Department of Biochemistry [1], School of Pharmacy [2], School of Chemistry [3], University of KwaZulu-Natal, Private Bag X54001, Durban, 4000, South Africa.

3.1 ABSTRACT

Matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) has been used for the characterisation of various microorganisms. It has been reported that yeast do produce characteristic protein fingerprint spectra but these signals may be harder to obtain than for protein extracts from bacteria. Although the identification and characterization of yeast using MALDI-TOF-MS has been the subject of recent studies, these studies have primarily focused on the optimization of signals mainly for clinical *Candida* yeast strains. The main objective of the present study was to develop a sample preparation protocol to obtain characteristic and highly reproducible MALDI-TOF mass spectra for *Saccharomyces cerevisiae* strains. In the present study it was found that ethanol/formic acid protein extraction, of a cell density of 1×10^8 cells co-crystallized with α-cyano-4-hydroxycinnmaic acid was required to yield mass spectral profiles that were diagnostic of each strain type. The mass spectral profiles of all three *S. cerevisiae* strains were unaffected by the culture medium, length of incubation and alcohol fixation. The optimized sample preparation protocol was simple and provided mass spectra that were unique and highly reproducible for individual *S. cerevisiae* strains that were suitable for the purpose of biotyping.

3.2 INTRODUCTION

Matrix assisted laser desorption time-of-flight mass spectrometry (MALDI-TOF MS) is an emerging tool for microbe identification at the species level (Marklein*, et al.*, 2009; Martorell*, et al.*, 2005; Mellmann*, et al.*, 2008; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003; van Veen*, et al.*, 2010). It was reported by Qian *et al*. (2008) and more recently by van Veen *et al*. (2010) that MALDI-TOF MS signals are more difficult to obtain for proteins extracted from yeast than for bacteria. It has been found that unlike bacteria, the direct approach that involves depositing a thin layer of yeast cells overlaid with matrix on a MALDI target plate does not yield sufficient diagnostic signals (van Veen*, et al.*, 2010). This may be attributed to the cell wall structure of yeast cells (Qian*, et al.*, 2008) which makes effective extraction of proteins problematic. Thus the optimization of sample preparation protocols has been considered crucial for obtaining unique and highly reproducible mass spectra (Marklein*, et al.*, 2009; Qian*, et al.*, 2008). Most studies based on the characterization and differentiation of yeast, have focused primarily on clinical *Candida* yeast isolates with the inclusion of very few *S. cerevisiae* strains.

In a study by Qian *et al*. (2008) it was demonstrated that MALDI-TOF mass spectrometry of whole-cell samples (dried-droplet method) could be used as a method for the differentiation of *Candida* species and other yeast in general. Notably only two *S. cerevisiae* strains (ATCC 9763 and ATCC 24903) were employed in this study. Furthermore it was shown that *C. albicans* strains based on MALDI-TOF MS signals could be grouped within a species. A large set of parameters were investigated which included the concentration of the yeast suspension, methanol fixation, the type of matrix employed, sample spotting techniques and the effect of cell incubation conditions. A cell number of 5×10^5 cells was found to produce the best mass spectral result for *C. albicans* (ATCC 6284) as fewer satisfactory mass spectral peaks were achieved above and below this cell concentration.

The fixation of yeast cells with methanol was first employed as a measure of safety when handling infectious organisms as methanol inactivates potentially harmful organisms (Qian*, et al.*, 2008). However, it was found that fixing yeast cells using a 50% methanol/ water solution improved the mass spectral signals in terms of increasing the number of characteristic ions. The analyte molecules in contact (co-crystallized) with the matrix molecules are those that can be desorbed/ionised and finally detected thus the medium sized molecules (4000-15000 Da) detected in this whole yeast cell study originated from the yeast cell surface. A possible explanation for the increased signals following alcohol fixation is that the surface macromolecules unfolded, thus exposing medium-sized molecules to the matrix. A second possibility is that the fixation may have promoted the release of cell membrane or

cytosolic/ribosomal materials to the cell surface, placing them in contact with the matrix. Other cell lysis methods including sonication, boiling and exposure to high concentrations of acids, such as trifluoroacetic acid, formic acid, nitric acid, and acetic acid which were used as a pre-treatment step for yeast failed to yield satisfactory results (Amiri-Eliasi and Fenselau, 2001). In addition the fixation procedure increased reproducibility in the sample spotting technique as methanol is thought to reduce yeast cell aggregation (Qian*, et al.*, 2008).

Sinapinic acid (SA) was previously shown to be an ideal MALDI matrix compound that is best suited for the analysis of peptides and small proteins (Beavis and Chait, 1990). Thus Qian and co-workers (2008) selected SA as the preferred matrix over compounds 2, 5-dihydroxy-benzoic acid (DHBA) and α-cyano-4-hydroxy-cinnamic acid (HCCA) which is consistent with the findings of Amiri-Eliasi and Fenselau, (2001). A two-layer method of sample preparation was also investigated by Qian *et al*. (2008). However this procedure provided no advantages over the dried droplet method. Production of diagnostic signals and reproducibility are critical for the approach to be successful. Thus the final parameter investigated by Qian *et al.* (2008) was the effect cell incubation conditions have on the spectra generated for different *Candida* strains. The effect of cell incubation conditions was found to be strain dependent. Therefore it was suggested that *Candida* yeast strains used for MALDI-TOF MS analysis should ideally be grown under identical conditions to avoid differences in the mass spectral signatures of the same strain.

Qian and co-workers (2008) found that by optimization of parameters mentioned above, differentiation of clinically important *Candida* strains was indeed possible. Furthermore a multivariate approach such as principle component analysis was employed in this study and was effective at clustering similar *C. albicans* strains but this was not performed on the two *S. cerevisiae* strains. This study proved that the yeast strains could be differentiated but importantly did not conclude that yeast strains could be identified by MALDI-TOF MS.

In the present study several parameters of an ethanol/formic acid protein extraction procedure as originally described by (Bruker, 2007) were assessed to develop a sample preparation protocol that yields characteristic and highly reproducible MALDI-TOF mass spectra. These parameters included cell density, length of incubation, culture medium, alcohol fixation and matrix selection and concentration. It was found that using the optimized sample preparation protocol, unique and highly reproducible mass spectral profiles were achievable for all three *S. cerevisiae* strains selected suitable for the purpose of biotyping.

3.3 MATERIALS AND METHODS

3.3.1 Strains

The yeast strains employed in this study are listed in Table 3.1

TABLE 3.1 *S. cerevisiae* strains employed in this study.

Strain	Use	Source
ATCC 9763	Ethanol	American Type Culture Collection
BJ2168	Laboratory	American Type Culture Collection
WS-NEW	Brewing	South African Breweries Ltd.

3.3.2 Media and cultivation conditions

Strains of *S. cerevisiae* were routinely cultivated at 30ºC on nutrient-rich YEPD medium (pH 6.5) containing 1% (wt/vol) yeast extract, 2% (wt/vol) peptone and 2% (wt/vol) glucose. Yeast strains were also cultivated on MYGP medium (pH 6.2), containing 0.3% (wt/vol) malt extract, 0.3% (wt/vol) yeast extract, 0.5% (wt/vol) peptone and 2% (wt/vol) glucose. For liquid cultures, single yeast colonies from 2 day old YEPD plates were used to inoculate 20 ml YEPD broth contained in 100 ml Erlenmeyer flasks, which were incubated at 30ºC with shaking (160 rpm) for 24 hours. In this study 2% (wt/vol) agar (Difco) was used for all solid media. For long-term storage, yeast strains were cryopreserved at -80ºC in YEPD-supplemented with 15% (vol/vol) glycerol (Ausubel*, et al.*, 1995).

3.3.3 Direct transfer whole-yeast cell analysis

A thin smear of a single colony of a *S. cerevisiae* strain grown aerobically on YEPD agar plates was deposited onto a ground-steel MALDI target plate (Bruker, Germany) as suggested by Bruker Daltonics (2007). The culture was then overlaid with 1 µl of α-cyano-4-hydroxycinnamic acid matrix solution [HCCA, (20 μg μl^{-1})] and air-dried in a laminar-flow hood at room temperature for 5 minutes prior to MALDI-TOF MS analysis.

3.3.4 Enumeration of yeast cells

The cell density of homogenized suitably diluted yeast cell suspensions in distilled water was determined either by manually counting cells using a hemocytometer or employing the Countess™ automated cell counter (Invitrogen) according to the instructions of the manufacturer.

3.3.5 Optimised ethanol/formic acid extraction procedure

Unless otherwise stated, all *S. cerevisiae* strains were cultivated on YEPD agar medium at 30°C for 48 h. Yeast cells were collected from the agar plate and suspended in 1 ml distilled water. A cell population of 1×10^8 cells was enumerated and the required volume was transferred to a 2 ml microcentrifuge tube. Yeast cells were harvested by centrifugation at 16000 rpm for 2 minutes (Eppendorf, centrifuge 5417R, Germany). The yeast cell pellet was resuspended in 300 µl of distilled water by vigorous vortexing and 900 µl of absolute ethanol was added to the cell suspension. The suspension was homogenized and centrifuged at 16000 rpm for 2 minutes and the supernatant was discarded. Centrifugation was repeated to completely remove all residual liquid. The ethanol-exposed cell pellet was air-dried in a laminar-flow hood at room temperature for 5 minutes. For yeast protein extraction, 50 µl of 70% formic acid (v/v) was added to the pellet and the contents were thoroughly mixed by vortexing. Thereafter 50 µl of 100% acetonitrile was added to the solution and mixed thoroughly by vortexing. The suspension was centrifuged at 16000 rpm for 2 minutes. Thereafter 1 µl of supernatant containing extracted yeast proteins was applied to a specified spot on a ground-steel MALDI target plate and air-dried in a laminar-flow hood at room temperature for 5 minutes. The applied sample was then overlaid with 1 µl of HCCA solution (20 $\mu g\ \mu l^{-1}$) and air dried. Once the sample was dry the target plate was analyzed by MALDI-TOF MS. All samples were prepared in triplicate.

3.3.6 Variations of ethanol/formic acid extraction procedure

In an attempt to develop an ethanol/formic acid extraction procedure that would yield distinctive and highly reproducible mass spectral profiles of the three *S. cerevisiae* strains employed in this study, the following parameters of the above procedure were investigated. Firstly, the influence of cell number on mass spectral profiles was evaluated using cell numbers ranging from 1×10^5 to 1×10^9 cells. Secondly, the effect of alcohol fixation was assessed by suspending *S. cerevisiae* strains in 95% methanol (v/v) as apposed to absolute ethanol. Thirdly, HCCA was evaluated as a suitable matrix solution by overlaying dried extracted samples on ground-steel MALDI target plate with 1µl HCCA solution containing HCCA that ranged from 2 to 20 µg. In a similar manner, the suitability of sinapinic acid as a matrix solution was also investigated. Fourthly, mass spectral analysis of *S. cerevisiae* strains cultivated on different nutrient-rich standard laboratory growth media (YEPD agar, YEPD broth, MYGP agar and MYGP broth) was also evaluated. All investigations were performed in triplicate.

3.3.7 Acquisition of MALDI-TOF MS data

The MALDI-TOF mass spectra were acquired on an AutoFlex III Smartbeam MALDI-TOF MS instrument (Bruker, Germany) using the instrument's pre-programmed FlexControl 3.0 method MBT_FC.par (Bruker, Germany). The flex-control method stipulated the following parameter settings. The intensity of the nitrogen laser (at 337 nm) was set significantly above the threshold

for desorption/ionization. The voltage of ion source one and two were set at 20 kV and 25 kV respectively. The pulsed-ion extraction time was set at 100 ns. For each sample spot (approximately 5 mm in diameter on the ground-steel MALDI target plate), an average of 600 shots was delivered at one point, and the final spectrum was an average accumulation of all the spectra gathered from at least six different points on a given sample spot. The protein molecular weight detection limit was set in the medium range between 2000 and 20000 Da. To validate the authenticity of mass spectral data generated by the MALDI-TOF MS instrument, each batch of samples contained the Bruker bacterial standard (BTS) (Bruker, Germany). The mass spectral profile of the BTS sample was aquired using the instrument's pre-programmed FlexControl 3.0 method, MBT_autoX.axe (Bruker, Germany) and it was routinely employed to calibrate the instrument. The MALDI Bruker Daltonics BioTyper software applies pattern matching to compare unknown mass spectra with reference data stored in the Bruker Daltonics mass spectral database. The Bruker Daltonics database currently holds reference spectra for 3740 microorganisms. The Bruker Daltonics Biotyper 3.0 software stipulates log (score) values for identification to species and genus levels as shown in Table 3.2.

Table 3.2 Interpretation of score values (Bruker, 2007).

Range	Description	Colour
2.300 to 3.000	Highly probable species identification	Green
2.000 to 2.299	Secure genus and probable species identification	Green
1.700 to 1.999	Probable species identification	Yellow
0.000 to 1.699	No reliable Identification	Red

3.3.8 Principle component analysis (PCA)

Principle component analysis (PCA) was employed to statistically evaluate MALDI-TOF mass spectra. PCA is a broadly used mathematical technique for reducing the dimensionality of multivariate data whilst preserving most of the variance (Kettaneh*, et al.*, 2005). PCA was performed using ClinProTools software 3.0 (Bruker, Germany). The PCA is managed by an external MATLAB software tool, which is integrated in ClinProTools. All spectra were normalised before PCA analysis. Complex relationships between samples were explored and illustrated with the use of PCA plots. These models allow one to determine similarities or dissimilarities between samples. For example samples that cluster closer together are more similar than samples that are further apart.

3.4 RESULTS

3.4.1 Direct transfer whole-yeast cell analysis

Irrespective of the matrix solution employed, the direct spotting of intact whole cells to the ground-steel MALDI target plate for identification of *S. cerevisiae* BJ2168 yielded poor signals that were not reproducible and were representative of red score values indicating no reliable identification (Fig. 3.1). All the signals fell into a narrow mass/charge range from 2000 to 8000 Daltons. Similar data was generated for the ATCC 9763 and WS-NEW strains (data not shown). In effect this approach did not produce sufficient diagnostic ions in the higher mass/charge range. Although this method is simpler than the ethanol/formic acid extraction protocol, from the above findings of no reliable identification it was decided not to proceed with further optimization of this protocol.

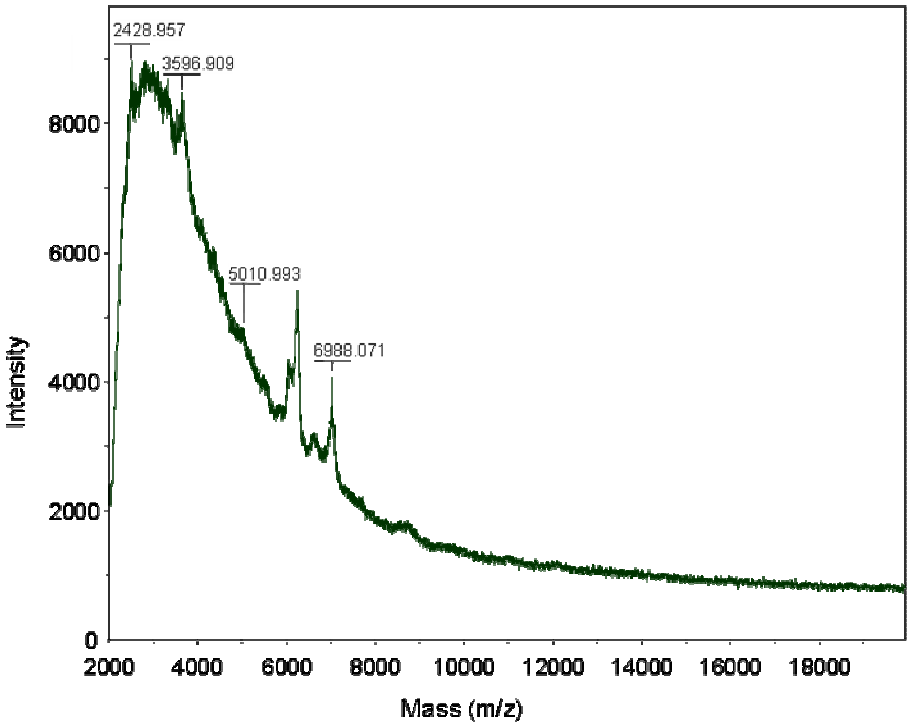


Figure 3.1 MALDI-TOF mass spectral fingerprint generated from direct whole cell analysis of *S. cerevisiae* strain BJ2168. The BJ2168 strain was grown on YEPD agar medium at 30°C for 48 hours and α-cyano-4-hydroxycinnamic acid was employed as the matrix solution.

3.4.2 The effect of cell number

Ethanol/formic acid extraction of samples containing from 1×10^5 to 1×10^9 cells was performed to determine the effect of initial cell number on MALDI-TOF mass spectral signatures (Fig. 3.2). It was found that when the typed *S. cerevisiae* BJ2168 strain was run against its previous acquisition on the Bruker Daltonics mass spectral database at varying cell densities, it was clearly observed that a reliable identification could not be obtained at a cell number of less than 1×10^7 cells. Extraction of proteins from increasing cell numbers was shown to have a

progressively positive effect in terms of the BioTyper software generated score values that reliably identified *Saccharomyces* to the species level. For all three strains investigated, the data seems to clearly indicate that a cell number of 1 x 10^8 cells is optimal for assignment to the species level. In addition the triplicate score values associated with samples of this cell number displayed the lowest variance. Interestingly, higher cell numbers for all three were observed to generally decrease the score values generated by the BioTyper software programme. The reproducibility of these mass spectral signatures was confirmed by principle component analysis (data not shown). Thus in all further studies, ethanol/formic acid extracted protein samples were prepared using 1 x 10^8 cells so as to ensure secure identification to the *Saccharomyces* strain level.

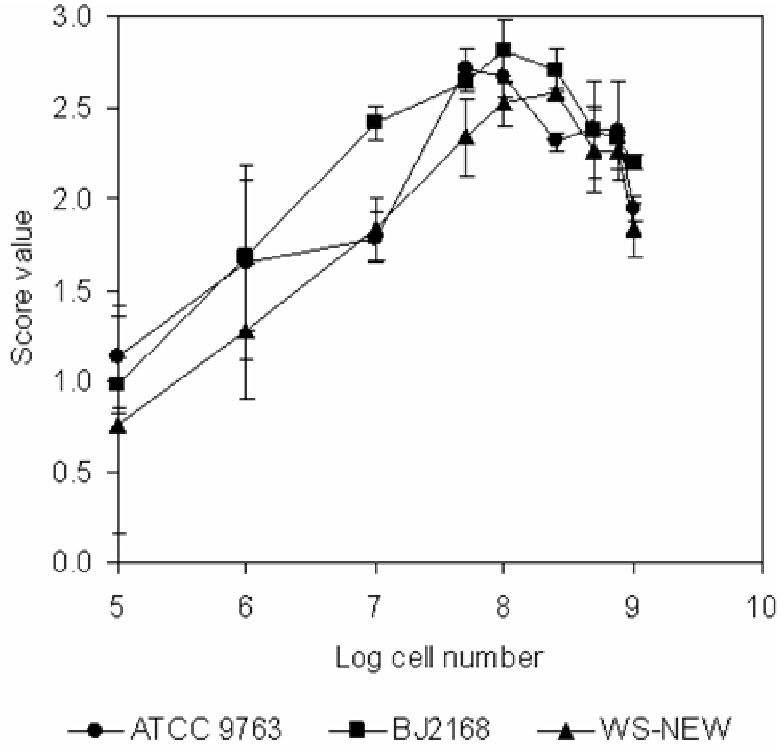


Figure 3.2 The effect of cell number on BioTyper software generated score values using the ethanol/formic extraction procedure for *S. cerevisiae* strain ATCC 9763, BJ2168 and WS-NEW. Alpha-cyano-4-hydroxycinnamic acid was employed as matrix the solution. All *S. cerevisiae* strains were grown on YEPD agar medium at 30°C for 48 hours.

3.4.3 The effect of alcohol fixation

The effect of substituting 95% methanol for absolute ethanol in the alcohol/formic acid extraction procedure on mass spectral profiles was investigated. The ethanol- or methanol-based mass spectral profiles for the same *S. cerevisiae* strain produced strong signals and sufficient diagnostic ions in the range of mass/charge 2000 to 14000 Daltons (spectra not shown). Comparison of spectral fingerprints of either ethanol- or methanol-treated *S. cerevisiae* BJ2168 samples to the Bruker Daltonics BioTyper database resulted in score values 2.810 ±

0.169 and 2.431 ± 0.243 respectively. This indicates that MALDI-TOF spectral signatures yielded with either ethanol or methanol were capable of securely identifying to both the species and strain level (green score value). However, following principle component analysis of the MALDI-TOF spectra obtained from both sample preparation strategies, it was found that spectra grouped separately according to the alcohol fixative employed (Fig. 3.3). Importantly it was also observed that mass spectral differences between ethanol-treated samples were smaller than the mass spectral differences between methanol-treated samples. This is clearly evidenced by the broader clustering of methanol-generated spectra. From these findings absolute ethanol was selected for the alcohol/formic acid extraction procedure for the purpose of biotyping *S. cerevisiae* strains.

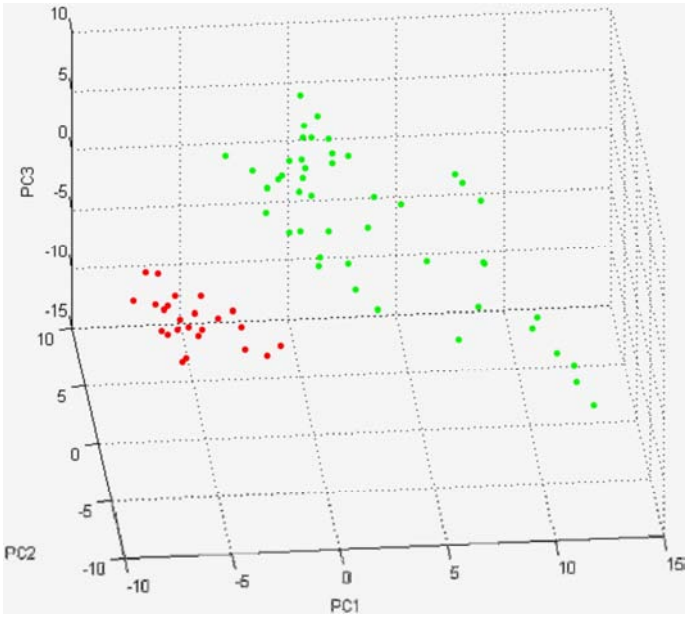


Figure 3.3 Principal component analysis of the mass spectral fingerprints of *S. cerevisiae* BJ2168 samples treated with either absolute ethanol (red) or 95% methanol (green). All *S. cerevisiae* BJ2168 cultures were grown on YEPD agar at 30°C for 48 hours. Samples were prepared from suspensions containing 1 x 10^8 cells and α-cyano-4-hydroxycinnamic acid was employed as the matrix solution. The scores plot concerns the first three principal components (PC), PC1, PC2 and PC3 which explain most of the variance within the data set. The scores plot shows the relatedness of the samples which may cluster close together or separately depending on their similarity or dissimilarity respectively. The axes of the scores plot record arbitrary units.

3.4.4 Matrix selection

The MALDI matrix compounds sinapinic acid (SA) and α-cyano-4-hydroxy-cinnamic acid (HCCA) were evaluated for their ability to facilitate reproducible mass spectral signatures for the purpose of biotyping *S. cerevisiae* strains. The SA or HCCA overlaid samples mass spectral profiles for the same *S. cerevisiae* strain produced different spectral qualities depending on the choice of matrix employed (Fig. 3.4). Both matrices produced a number of diagnostic ions in the range of mass/charge 2000 to 14000 Daltons. However, in terms of reproducibility the mass spectral fingerprints using sinapinic acid were not as consistent as those generated when

employing HCCA. Moreover decreased peak intensities were associated with the use of SA. Principle component analysis also confirmed the suitability of the HCCA matrix for MALDI-TOF-based differentiation of *S. cerevisiae* strains (data not shown).

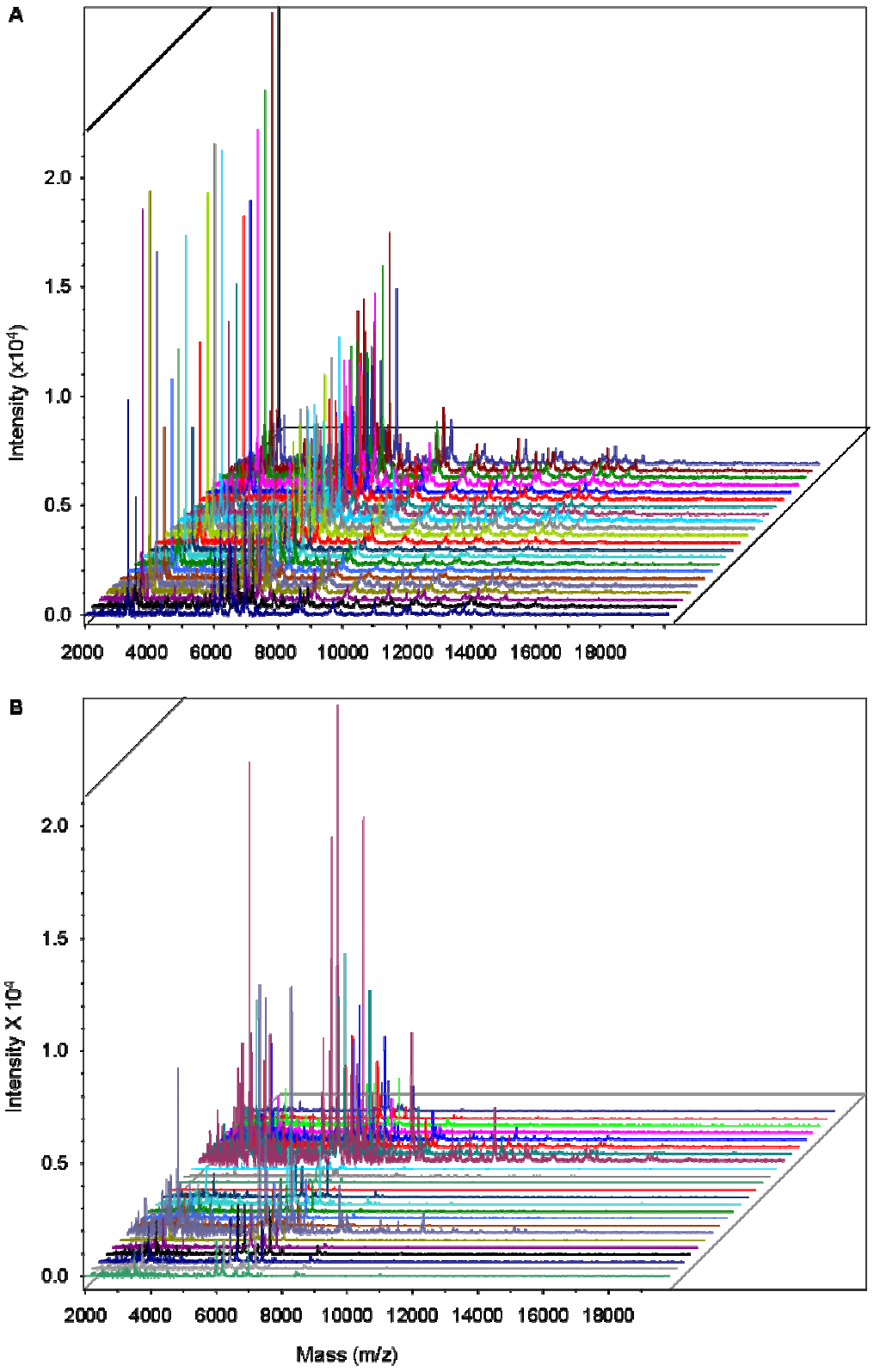


Figure 3.4 Stacked view of MALDI-TOF mass spectral fingerprints of *S. cerevisiae* WS-NEW when (A) α-cyano-4-hydroxycinnamic acid (HCCA) and (B) sinapinic acid was employed as the matrix solution.

3.4.5 The effect of culture media

Investigation of three *S. cerevisiae* strains cultivated in MYGP agar, YEPD agar and YEPD broth media revealed that culture media and conditions had an influence on MALDI-TOF mass spectral profiles. It was observed that irrespective of the culture medium employed the mass spectral signatures for the same *S. cerevisiae* strain produced strong signals and adequate diagnostic ions in the range of mass/charge 2000 to 14000 Daltons (spectra not shown). Comparison of spectral fingerprints of *S. cerevisiae* BJ2168 samples cultivated on different media types to the Bruker Daltonics BioTyper database spectral fingerprint for the same strain resulted in score values that securely differentiated to both the species and strain level (green score values). Principle component analysis of the MALDI-TOF spectral profiles grouped spectra separately according to the different media types (Fig. 3.5). Similar trends were observed for the ATCC 9763 and WS-NEW strains (data not shown). Although the type of cultivation media seems to clearly influence the mass spectral profiles, it does not compromise *S. cerevisiae* strain identification. Considering the Bruker BioTyper mass spectral database recommends growth on solid agar medium for biotyping purposes, all *S. cerevisiae* were subsequently grown on solid YEPD agar medium containing 2% glucose.

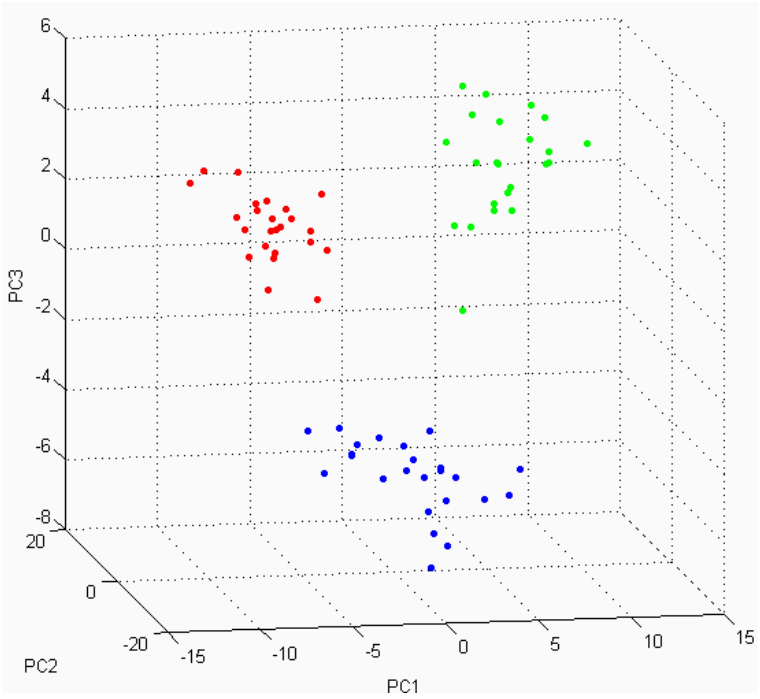


Figure 3.5 Principle component analysis of mass spectral signatures of *S. cerevisiae* BJ2168 cultivated in YEPD broth (blue), YEPD (red) and MYGP (green) agar plates. Samples were prepared from suspensions containing 1 x 10^8 cells and α-cyano-4-hydroxycinnamic acid was employed as the matrix solution.

3.4.6 Distinctive mass spectral fingerprints of *S. cerevisiae* strains

Employing the optimized ethanol/formic acid extraction procedure on cells of three *S. cerevisiae* strains (BJ2168, ATCC 9763 and WS-NEW) that were aerobically cultivated on YEPD agar medium at 30°C for 48 hours yielded characteristic and highly reproducible mass spectral profiles for replicates of the same strain grown on different days (Fig. 3.6). Whilst a small proportion of diagnostic ions were evident in mass spectral signatures of all three strains, the signatures also displayed significantly different ions in the mass/charge range of 2000 to 14000 Daltons.

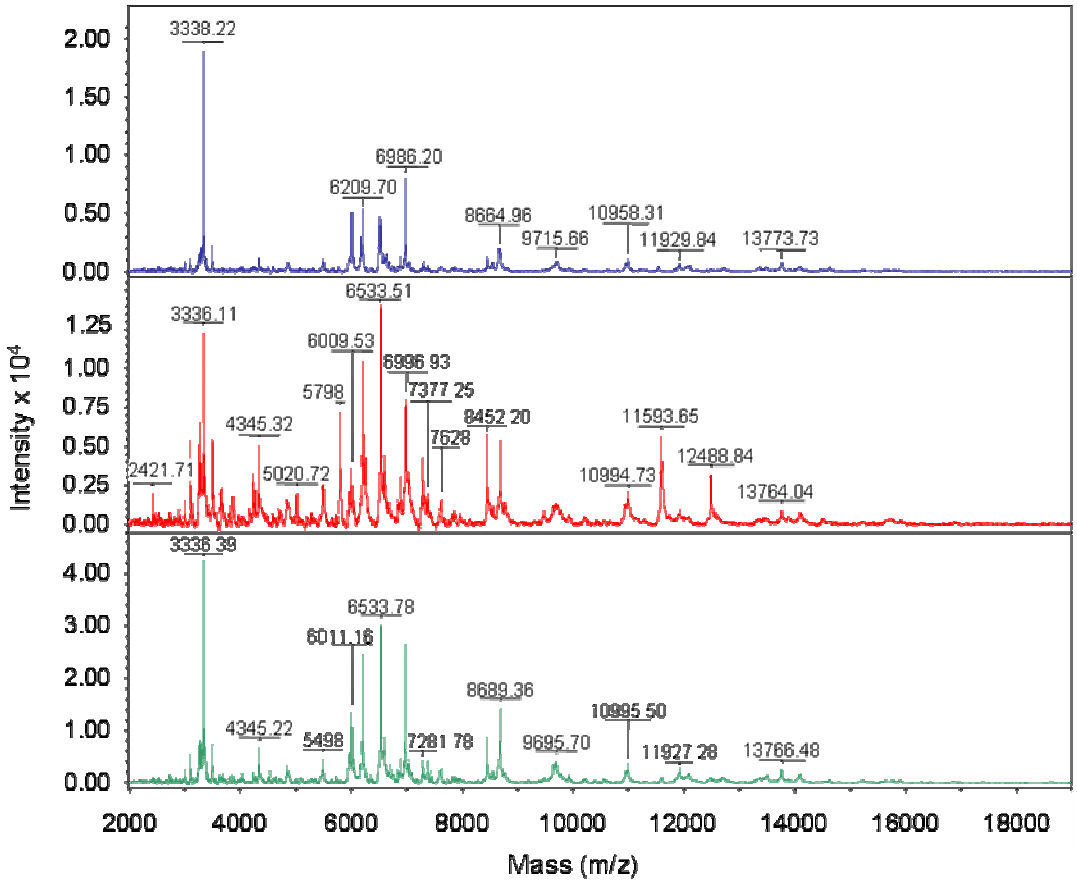


Figure 3.6 MALDI-TOF mass spectral profiles of *S. cerevisiae* strains WS-NEW (blue), BJ2168 (red) and ATCC 9763 (green). Samples were prepared from suspensions containing 1 x 10^8 cells that were aerobically grown on YEPD agar medium at 30°C for 48 hours. Alpha-cyano-4-hydroxycinnamic acid was employed as the matrix solution

Principle component analysis of the mass spectra obtained from the three different *S. cerevisiae* strains, illustrated that the spectra grouped separately according to the specific *S. cerevisiae* strain (Fig. 3.7). Although mass spectra from individual strains clustered into defined groups without outliers, mass spectral differences were observed within the spectra generated

from certain strains. Importantly these spectral differences did not compromise the BioTyper software-based identification of these strains. The data seems to strongly suggest that the optimized ethanol/formic acid procedure is capable of generating significantly different spectral signals that may be employed in the differentiation of these S. cerevisiae strains.

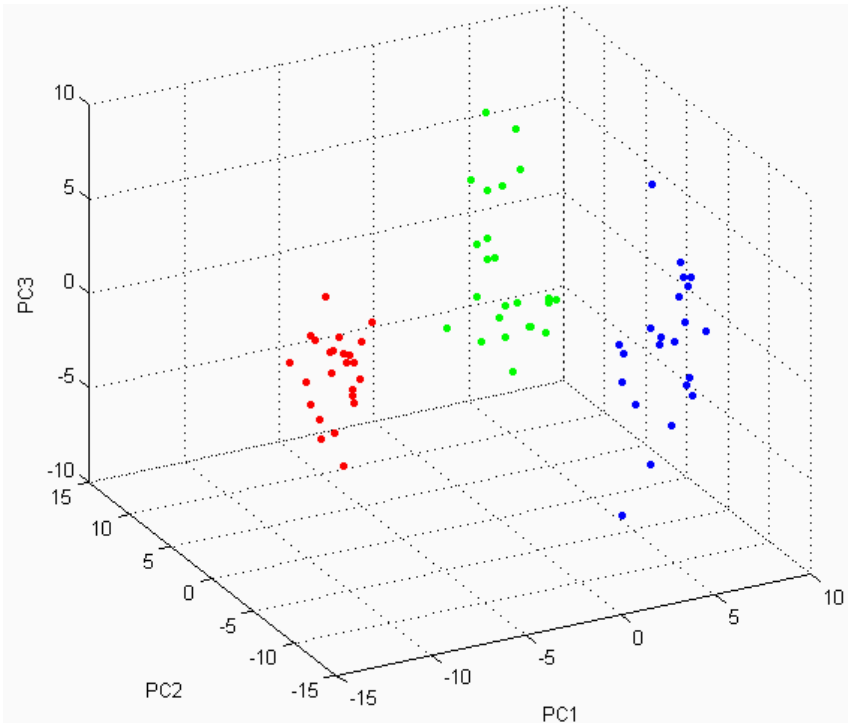


Figure 3.7 Principal component analysis of mass spectral fingerprints derived from *S. cerevisiae* WS-NEW (blue), BJ2168 (red) and ATCC 9763 (green).

3.5 DISCUSSION

To date, this is the first study to report on the identification of laboratory and industrial *S. cerevisiae* strains using an ethanol/formic acid extraction specifically optimized for this purpose. This study shows the potential of MALDI-TOF MS as a rapid tool for *S. cerevisiae* characterization to the strain level. Various parameters including cell number, alcohol fixation, cultivation medium and MALDI matrix were considered during the optimization of an ethanol/formic acid extraction protocol that was previously developed by Bruker Daltonic (2007) for the broad-based MALDI-TOF identification of microorganisms. Sample preparation using the optimized extraction procedure, yielded unique and highly reproducible mass spectral fingerprints for the three *S. cerevisiae* strains employed in this study.

Direct MALDI-TOF MS analysis of whole intact microbial cells coated with matrix was previously demonstrated to yield distinctive mass spectral signatures for bacteria (Claydon*, et al.*, 1996; Ilina*, et al.*, 2009; Krishnamurthy and Ross, 1996; Mandrell*, et al.*, 2005; Mellmann*, et al.*, 2008) and other yeast genera (Amiri-Eliasi and Fenselau, 2001; Qian*, et al.*, 2008). However, our data clearly illustrated that this simpler and faster protocol was not suitable for biotyping of *S. cerevisiae* strains employed in this study. In contrast, Qian and co-workers (2008) reported that the direct technique is suitable for the generation of unique mass spectral profiles for *S. cerevisiae* strains ATCC 9763 and ATCC 24903. Comparison of the mass spectral fingerprints reveals that five diagnostic ions on average were evident for the profiles obtained in this study (using an ethanol/formic acid protein extraction procedure) whilst only three ions were evident in directly obtained mass spectra reported by Qian and co-workers. In addition, comparison of the mass spectral profiles of all three *S. cerevisiae* strains used in our study to the Bruker Daltonics BioTyper mass spectral database were representative of red score values indicating no reliable identification even to the *Saccharomyces* genus level. Moreover, a recent study by van Veen *et al*. (2010) demonstrated that ethanol/formic acid pretreatment procedure is critical in yielding unique MALDI-TOF mass spectral fingerprints for the identification of nineteen clinical yeast isolates to the genus and species level.

The second parameter investigated involving the use of ethanol or methanol in the alcohol/formic acid extraction procedure was found to have no affect on the identification of all three *S. cerevisiae* strains employed in this study. All strains were correctly identified to both the species and strain level. However, comparatively the use of ethanol in the alcohol/formic acid extraction protocol yielded characteristic and more reproducible MALDI-TOF mass spectra than methanol. This finding is supported by other research studies that have also employed ethanol (Giebel*, et al.*, 2008; Ilina*, et al.*, 2009; Marklein*, et al.*, 2009; van Veen*, et al.*, 2010). In contrast another study has reported that lyticase lysis of yeast cells prior to methanol/trifluoroacetic acid extraction is suitable for the generation of characteristic mass spectral fingerprints that allows for differentiation of *Candida* species and S. *cerevisiae* strains (Sherburn and Jenkins, 2003). The sample preparation protocol described in the present study is more rapid and economical than that described by Sherburn and Jenkins (2003).

Improvement of the contact of ethanol/formic acid extracted yeast molecules (analyte) with the MALDI matrix material requires an optimal ratio between the matrix and extracted yeast molecules. This is a crucial step in sample preparation as an uneven matrix/analyte ratio can lead to a reduction or even total suppression of yeast-derived diagnostic ion signals (Qian*, et al.*, 2008). In this study it was found that the matrix HCCA (20 μg μl^{-1}) yielded unique mass spectral fingerprints for the identification of *Saccharomyces cerevisiae* strains to the species and strain level. In addition the spectra obtained were more reproducible than when SA was

employed. However, it must be noted that other studies have shown SA to be the most suitable matrix for the MALDI-TOF MS differentiation of other yeast genera (Amiri-Eliasi and Fenselau, 2001; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003).

Lastly it was found that the mass spectral signatures of the three *S. cerevisiae* strains employed in this study produced more than adequate diagnostic ions in a manner seemingly independent of the growth medium employed so as to identify them to the species and strain level. This finding is supported by other research studies that have made similar observations (Marklein*, et al.*, 2009; van Veen*, et al.*, 2010). However, it must be highlighted that only two different types of media were used in this study and interestingly, Qian and co-workers (2008) have reported that MALDI-TOF spectra obtained from the analysis of yeast cells cultivated on different media is strain dependent. In the interests of maintaining a standardized protocol the data of the present study also indicates that it is advisable to use a single cultivation medium for the purpose of biotyping *S. cerevisiae* strains.

The data thus far seems to strongly support the hypothesis that provided an optimized ethanol/formic acid extraction procedure is coupled with HCCA as the MALDI matrix, MALDI-TOF MS can be used to produce highly reproducible and distinctive mass spectral fingerprints of *S. cerevisiae* strains to the strain level. Furthermore the mass spectral fingerprints of all three *S. cerevisiae* strains employed in this study were considered suitable by Bruker Daltonics for the purpose of biotyping. This proof of concept will be further assessed by biotyping a larger cohort of *S. cerevisiae* strains which will be the main topic of the proceeding chapter.

3.6 ACKNOWLEDGEMENTS

This work was made possible through financial support from the National Research Foundation and UKZN.

3.7 REFERENCES

Amiri-Eliasi, B. and Fenselau, C. (2001). Characterization of Protein Biomarkers by MALDI from whole fungal cells. *Anal. Chem* **73,** 5228-5231.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1995). *Current Protocols in Molecular Biology*. John Wiley and Sons.

Beavis, R. C. and Chait, B. T. (1990). Rapid, sensitive analysis of protein mixtures by mass spectrometry. *Proc Natl Acad Sci U S A* **87,** 6873-7.

Bruker (2007). Bruker BioTyper User Manual. **Version 2.0**.

Claydon, M. A., Davey, S. N., Edwards-Jones, V. and Gordon, D. B. (1996). The rapid identification of intact microorganisms using mass spectrometry. *Nat Biotechnol* **14,** 1584-6.

Giebel, R. A., Fredenberg, W. and Sandrin, T. R. (2008). Characterization of environmental isolates of *Enterococcus spp*. by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Water Res* **42,** 931-40.

Ilina, E. N., Borovskaya, A. D., Serebryakova, M. V., Chelysheva, V. V., Momynaliev, K. T., Maier, T., Kostrzewa, M. and Govorun, V. M. (2009). Application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for the study of *Helicobacter pylori*. *Rapid Commun. Mass Spectrom.* **24,** 328-334.

Kettaneh, N., Berglund, A. and Wold, S. (2005). PCA and PCL with very large data sets. *Comp. Stat. Data Anal.* **48,** 69-85.

Krishnamurthy, T. and Ross, P. L. (1996). Rapid identification of bacteria by direct matrix-assisted laser desorption/ionization mass spectrometric analysis of whole cells. *Rapid Commun Mass Spectrom* **10,** 1992-6.

Mandrell, R. E., Harden, L. A., Bates, A., Miller, W. G., Haddon, W. F. and Fagerquist, C. K. (2005). Speciation of *Campylobacter coli, C jejuni, C. helveticus, C. lari, C. sputorum* and *C. upsaliensis* by Matrix-Assisted Laser Desorption Ioniztion-Time of Flight Mass Spectrometry. *Appl. Environ. Microbiol.* **71,** 6292-6307.

Marklein, G., Josten, M., Klanke, U., Muller, E., Horre, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A. and Sahl, H. G. (2009). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* **47,** 2912-7.

Martorell, P., Querol, A. and Fernandez-Espinar, M. T. (2005). Rapid identification and enumeration of *Saccharomyces cerevisiae* cells in wine by real-time PCR. *Appl Environ Microbiol* **71,** 6823-30.

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., Dunn, J., Hall, G., Wilson, D., LaSala, P., Krostrzewa, M. and Harmsen, D. (2008). Evaluation of Matrix-Assisted Laser Desorption Ionization-Time-of-Flight Mass Spectrometry in Comparison to 16S rDNA Gene Sequencing for Species Identification of Nonfermenting Bacteria. *J. Clin. Microbiol.* **46,** 1946-1954.

Qian, J., Cutler, J. E., Cole, R. B. and Cai, Y. (2008). MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers. *Anal Bioanal Chem* **392,** 439-49.

Sherburn, R. E. and Jenkins, R. O. (2003). A novel and rapid approach to yeast differentiation using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Spectroscopy* **17,** 31-38.

van Veen, S. Q., Claas, E. C. and Kuijper, E. J. (2010). High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *J Clin Microbiol* **48,** 900-7.

Chapter 4

RESEARCH RESULTS II

Biotyping *S. cerevisiae* strains

Biotyping *Saccharomyces cerevisiae* strains

Anushka Moothoo-Padayachie [1], Thavendran Govender [2], Hendrik G. Kruger [3], Glenn E. M. Maguire [3] and Patrick Govender [1].

Department of Biochemistry [1], School of Pharmacy [2], School of Chemistry [3], University of KwaZulu-Natal, Private Bag X54001, Durban, 4000, South Africa.

4.1 ABSTRACT

Phenotypic and biochemical methods have traditionally been used for identification of yeast but can be time consuming and often unreliable. Molecular biology-based approaches have opened up new avenues for yeast identification and presently are considered the methods of choice. These methods can also be slow, labour-intensive and require considerable technical expertise. Matrix assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS) is a promising tool for microbe identification and in recent years has been shown to be useful for the unequivocal identification of bacteria and clinical yeast isolates. However, for industrial purposes of patent protection, strain verification and contaminating strain/s identification, a method that allows further discrimination at the strain level (biotyping) is required. In this study, MALDI-TOF MS is investigated as a method to rapidly differentiate and identify *S. cerevisiae* strains to the strain level. Several *S. cerevisiae* strains commonly employed in South African fermentation-based industries were biotyped in an attempt to create a yeast reference database within a local context. A reference database of forty five *S. cerevisiae* strains was established and the data revealed that MALDI-TOF MS can be used for the rapid and accurate identification of laboratory and industrial *S. cerevisiae* strains.

4.2 INTRODUCTION

Since 1883, when Emil Christian Hansen first used a pure culture of brewer's yeast on a large scale, the application of yeast starter cultures has become essential for the production of fermented beverages (Barszczewski and Robak, 2004). Starter cultures are used to ensure product quality and consistency (Lopez*, et al.*, 2003; Manzano*, et al.*, 2006; Wightman*, et al.*, 1996). Commercial *S. cerevisiae* strains play a crucial role in the production of baked and fermented foods, potable alcohol and industrial ethanol (Dombek and Ingram, 1987). Certain fermentation processes can only be carried out by a select strain/s of *S. cerevisiae*. As a result the accurate identification of *S. cerevisiae* strains has become a key issue (Gonzalez-Techera*, et al.*, 2001). Furthermore spoilage yeasts and bacteria if undetected in starter cultures or in a fermentation can result in large financial losses in production, processing, preservation and storage (Hierro*, et al.*, 2004). The accurate identification of *S. cerevisiae* strains is also important in industry for the effective protection of patented strains. Therefore in industry there is a need for a simple, rapid and accurate method that allows for the discrimination of *S. cerevisiae* to the strain level (biotyping).

Any biotyping method used for industrial purposes should provide extensive and reproducible discrimination. The ideal method should allow for testing a large number of isolates, provide rapid results, require easy and inexpensive procedures and permit automation. Conventional phenotypic methods are not suitable as a biotyping tool for yeast identification as these methods are too time consuming, cannot be automated easily and can yield ambiguous results (Buzzini*, et al.*, 2007).

Molecular biology tools have now been developed to discriminate between closely related yeast including *S. cerevisiae* strains (Barszczewski and Robak, 2004; Barszczewski and Robak, 2006; Buchaille*, et al.*, 1998; Giusto*, et al.*, 2006; Hierro*, et al.*, 2004; Logan*, et al.*, 2009). A great number of molecular biology techniques rely on the polymerase chain reaction [PCR, (Barszczewski and Robak, 2006)]. One such method involves the identification of yeast strains using PCR with delta primers. The delta sequences are direct repeat elements of 0.3 kb that flank the TY1 retrotransposon. The position and/or number of the delta elements are strain dependant. Therefore the amplification products will be different according to the strains and will allow for their discrimination. One major short coming of the technique is that it can only be used for the identification of *S. cerevisiae* strains and related species thus is not a very extensive biotyping tool. Molecular biology methods can be used as biotyping tools for the effective discrimination of *S. cerevisiae* strains but some are time-consuming, selective for only a particular species and require considerable expertise (Ness*, et al.*, 1993; Qian*, et al.*, 2008; Sheehan*, et al.*, 1991; Sherburn and Jenkins, 2003).

Since the invention of MALDI-TOF MS in 1988, the technique has been developed for the identification of microbes (Amiri-Eliasi and Fenselau, 2001; Ilina*, et al.*, 2009; Mandrell*, et al.*, 2005; Marklein*, et al.*, 2009; Mellmann*, et al.*, 2008; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003; van Veen*, et al.*, 2010). These microbial biotyping endeavours have thus far primarily focused on the identification of bacteria and clinical yeast isolates such as *Candida* species. Studies based on the identification of clinical yeast isolates using MALDI-TOF MS unlike bacteria, have only shown the application to resolve clinical yeast isolates to the species level and not the strain level. Thus far and to the best of our knowledge a comprehensive study demonstrating the application of MALDI-TOF MS for *Saccharomyces cerevisiae* strain identification has not yet been conducted. The latter is supported by the commercially available Bruker Daltonics mass spectral database of three thousand seven hundred and forty microorganisms of which only seven are *S. cerevisiae* strains. As it currently stands, a major limiting aspect of MALDI-TOF MS for *S. cerevisiae* identification is the need for a comprehensive strain specific mass spectral database. Thus for the method to be an extensive biotyping tool for the identification of *S. cerevisiae* to the strain level, the development of such a comprehensive *S. cerevisiae* database is critical.

In this study a mass spectral database of forty five *S. cerevisiae* strains was initially established. Subsequently it was demonstrated that MALDI-TOF MS can be used for the rapid and accurate identification of laboratory and industrial *S. cerevisiae* species to the strain level.

4.3 MATERIALS AND METHODS

4.3.1 Strains

The yeast strains employed in this study are listed in Table 4.1.

TABLE 4.1 *S. cerevisiae* strains employed in this study

Strain	Use	Source
285	Wine making	Lallemand Inc., Canada
ALCOHOL D3	Ethanol	Natal Cane Products
ATCC 2601	Ethanol	American Type Culture Collection
ATCC 7754	Ethanol	American Type Culture Collections
ATCC 9763	Ethanol	American Type Culture Collection
BAKERS D4/2	Ethanol	Natal cane Products
BJ2168	Laboratory	American Type Culture Collection
BM45	Wine making	Lallemand Inc., Canada
BP725	Wine making	Maurivin Yeast, Australia
BY4742	Laboratory	Euroscarf, Frankfurt
Cross evolution	Wine making	Lallemand Inc., Canada
DV10	Wine making	Lallemand Inc., Canada
DY802	Baking	Anchor Yeast, South Africa
EC1118	Wine making	Lallemand Inc., Canada
Enoferm BDX	Wine making	Lallemand Inc., Canada
F5	Wine making	Laffort, France
F15	Wine making	Laffort, France
Fermicru 4F9	Wine making	DSM Food specialities, France
Fermol Chardonnay	Wine making	AEB Africa (PTY) LTD
Fermol Sauvignon	Wine making	AEB Africa (PTY) LTD
FY23	Laboratory	(Winston*, et al.*, 1995)
Lalvin PdM	Wine making	Lallemand Inc., Canada
NCYC79	Baking	National Collection of Yeast Cultures
NT50	Wine making	Anchor Yeast, South Africa
N96	Wine making	Anchor Yeast, South Africa
NT112	Wine making	Anchor Bio-Technologies
NT116	Wine making	Anchor Bio-Technologies
NT202	Wine making	Anchor Bio-Technologies
NT45	Wine making	Anchor Bio-Technologies
NWS Chardonnay	Wine making	Intec International Technology
NWS Pinot	Wine making	Intec International Technology
PM	Wine making	Laffort, France
RMS2	Wine making	Laffort, France
ROSE´	Wine making	Laffort, France
SAB-DY	Brewing	South African Breweries Ltd.
SABMC-CO	Brewing	South African Breweries Ltd.
SABMC-UF	Brewing	South African Breweries Ltd.
VIN13	Wine making	Anchor Yeast, South Africa
VIN2000	Wine making	Anchor Bio-Technologies
VIN7	Wine making	Anchor Bio-Technologies
W15	Wine making	Lalleman Inc., Canada
WE372	Wine making	Anchor Yeast, South Africa
WS-NEW	Brewing	South African Breweries Ltd.
WS-CAR	Brewing	South African Breweries Ltd.
X5	Wine	Laffort, France

4.3.2 Media and cultivation conditions

Strains of *S. cerevisiae* were routinely cultivated at 30ºC on nutrient-rich YEPD medium containing 1% (wt/vol) yeast extract, 2% (wt/vol) peptone, 2% (wt/vol) glucose and 2% (wt/vol) agar (Difco) for 48 hours. Replicate cell samples were harvested from cultures grown under identical conditions on three different days. For long-term storage, yeast strains were cryopreserved at -80ºC in YEPD supplemented with 15% (vol/vol) glycerol (Ausubel*, et al.*, 1995).

4.3.3 Optimized ethanol/formic acid extraction procedure

Yeast cells were collected from the agar plate and suspended in 1 ml distilled water. A cell population of 1×10^8 cells was enumerated and the required volume was transferred to a 2 ml microcentrifuge tube. Yeast cells were harvested by centrifugation at 16000 rpm for 2 minutes (Eppendorf, centrifuge 5417R, Germany). The yeast cell pellet was resuspended in 300 µl of distilled water by vigorous vortexing and 900 µl of absolute ethanol was added to the cell suspension. The suspension was homogenized and centrifuged at 16000 rpm for 2 minutes and the supernatant was discarded. Centrifugation was repeated to completely remove all residual liquid. The ethanol-exposed cell pellet was air-dried in a laminar-flow hood at room temperature for 5 minutes. For yeast protein extraction, 50 µl of 70% formic acid (v/v) was added to the pellet and the contents were thoroughly mixed by vortexing. Thereafter 50 µl of 100% acetonitrile was added to the solution and mixed thoroughly by vortexing. The suspension was centrifuged at 16000 rpm for 2 minutes. Thereafter 1 µl of supernatant containing extracted yeast proteins was applied to a specified spot on a ground-steel MALDI target plate and air-dried in a laminar-flow hood at room temperature for 5 minutes. The applied sample was then overlaid with 1 µl of HCCA solution (20 $\mu g\ \mu l^{-1}$) and air dried. Once the sample was dry the target plate was analyzed by MALDI-TOF MS. All samples were prepared in triplicate. Replicate samples comprised of cultures grown under identical conditions on three different days.

4.3.4 Acquisition of MALDI-TOF MS data

The MALDI-TOF mass spectra were acquired on an AutoFlex III Smartbeam MALDI-TOF MS instrument (Bruker, Germany) using the instrument's pre-programmed FlexControl 3.0 method MBT_FC.par (Bruker, Germany). The flex-control method stipulated the following parameter settings. The intensity of the nitrogen laser (at 337 nm) was set significantly above the threshold for desorption/ionization. The voltage of ion source one and two were set at 20 kV and 25 kV respectively. The pulsed-ion extraction time was 100 ns. For each sample spot (approximately 5 mm in diameter on the ground-steel MALDI target plate), an average of 600 shots was delivered at one point, and the final spectrum was an average accumulation of all the spectra gathered from at least six different points on a given sample spot. The protein molecular weight detection limit was set in the medium range between 2000 and 20000 Daltons.

4.3.5 Creation of mass spectral database entries

Each strain was spotted at 8 positions on the MALDI target plate and measured three times per spot. Thus twenty four spectra were generated. Using replicate cultures, this was repeated on three separate occasions thereby generating 72 spectra per *S. cerevisiae* strain. A mass spectral profile (MSP) consisting of an average of ± 24 most reproducible spectra were created using the Biotyper 3.0 software individually for all forty five *S. cerevisiae* strains to generate a reference mass spectral database. An MSP is an average "spectrum" or "a reference peak list" which is created by a defined creation method from a number of preprocessed spectra acquired from a known *S. cerevisiae* strain. The MSP represents the individual peak pattern including peak intensity distribution and peak frequency of the respective *S. cerevisiae* strain by extracting the typical peak information. For quality assurance, all raw spectra used for MSP creation were processed according to a standardized protocol as advised by Bruker Daltonics.

To validate the authenticity of mass spectral data generated by the MALDI-TOF MS instrument, each batch of samples contained the Bruker bacterial standard [BTS, (Bruker, Germany)]. The mass spectral profile of the BTS sample was acquired using the instrument's pre-programmed FlexControl 3.0 method, MBT_autoX.axe (Bruker, Germany) and it was routinely employed to calibrate the instrument. Creation of a MSP dendrogram to determine the appropriate relatedness of the *S. cerevisiae* strains was conducted using the Bruker Daltonics BioTyper 3.0 software.

4.3.6 Yeast chromosomal DNA isolation

Chromosomal DNA was extracted by the method described by Ausubel *et al*. (1995). Yeast strains were grown at 30°C with shaking (160 rpm) for 18 hours in 10 ml YEPD medium. The cells were collected by centrifugation at 3000 rpm at 4°C for 2 minutes. The medium was then discarded and the cells were resuspended in 500 µl of distilled water. The cell suspension was then transferred to a 2 ml microcentrifuge tube and centrifuged at 12000 rpm at 4°C for 30 seconds (Eppendorf Centrifuge 5417R, Germany). The supernatant was discarded and the pellet was resuspended by vortexing briefly in the residual liquid. Thereafter 200 µl of breaking buffer, 0.3 g of glass beads and 200 µl of a phenol/chloroform/isoamyl alcohol [25:24:1 (v/v/v)] mixture were added to the cell suspension. This was then followed by vortexing (Vortex Genie 2) at the highest speed for 5 minutes. Subsequently 200 µl of a buffer containing 10 mM Tris-HCl and 1 mM EDTA pH 8.0 (TE) was added to the mixtures and the tubes were briefly vortexed. Reaction mixtures were centrifuged at 14000 rpm at 4°C for 10 minutes. The aqueous layer (~400 µl) was transferred into a new 1.5 ml microcentrifuge tube and 1 ml of 100% ethanol was added to the aqueous fraction. To facilitate chromosomal DNA precipitation, the contents were mix by inverting the microcentrifuge tube and mixtures were stored at -20°C for 20

minutes. The tubes were then microcentrifuged at 14000 rpm at 4°C for 5 minutes. The supernatant was discarded and the pellet was resuspended in 400 µl TE. Contaminating RNA was degraded by the addition of 3 µl of RNase (10 $mg.ml^{-1}$) and reaction mixtures were incubated at 37°C for 5 minutes. Thereafter chromosomal DNA was precipitated by the addition 10 µl of 4M ammonium acetate and 1 ml of 100% ethanol. The contents were mixed by inversion and stored at -80°C for 20 minutes. Tubes were then centrifuged at 14000 rpm at 4°C for 3 minutes and supernatants discarded. The pellet was dried in a vacuum concentrator (Eppendorf Concentrator 5301, Germany) and isolated DNA was resuspended in 100 µl of TE. A 1/10 dilution of isolated chromosomal DNA was routinely employed as template in amplification reactions.

4.3.7 PCR-based yeast strain identification

All PCR reactions were performed using the Takara Ex Taq™ PCR system (Takara Bio Inc., Otsu, Japan) as per manufactures instructions. The delta primers employed in this study were designed according to Ness et al. (1993) and are listed in Table 4.2. All amplification reactions were performed in an Eppendorf Mastercycler Gradient (Germany). The PCR conditions were one denaturation cycle at 94°C for 2 minutes, followed by 30 cycles of 94°C for 30 seconds, 45°C for 30 seconds, 72°C for 2 minutes. Aliquots (10 µl) of the amplification reaction were loaded onto a 1% agarose gel and subjected to electrophoresis at 60 V for 5 hours. A 1 kb molecular weight marker (Fermentas, Germany) was employed.

Table 4.2 Primers employed for PCR reactions

Primer Name [a]	Primer sequence (5'→3')
δ-F	CAAAATTCACCTAT$^{A}/_{T}$TCTCA
δ-R	GTGGATTTTTATTCCAACA

[a] *F, forward primer and R, reverse primer.*

4.3.8 Contour-clamped homogenous electric field gel electrophoresis

The cultivation of yeast strains, preparation of chromosomal DNA plugs and contour clamped homogenous electric field gel electrophoresis (CHEF) in a Biorad DR III CHEF apparatus was performed according to the method described by Van der Westhuizen and co-workers (1999).

4.3.9 MALDI-TOF-based biotyping of blind-coded *S. cerevisiae* strains

For biotyping purposes, blind-coded *S. cerevisiae* strains of laboratory and industrial origins were employed to evaluate the robustness of the newly created *S. cerevisiae* MALDI-TOF MS reference database. Twenty blind-coded *S. cerevisiae* strains were obtained from the Agricultural Research Council (ARC), Infruitec-Nietvoorbij (Institute for Deciduous Fruit, Vines and Wine), Stellenbosch, South Africa. In addition seven blind-coded strains were obtained from NCP Alcohols, Durban, South Africa. All blind-coded samples were subjected to the optimized ethanol/formic acid extraction procedure.

Following extraction and acquisition of MALDI-TOF mass spectra, the Bruker Daltonics BioTyper 3.0 software was used to compare the peak pattern of the unknown spectra generated from the blind-coded *S. cerevisiae* strains to the pattern of the MSP's entered into the *S. cerevisiae* mass spectral database. The identification results were based on the log (score) values calculated by the Bruker Daltonics BioTyper 3.0 Software as shown in Table 4.3. These BioTyper thresholds were empirically determined on the basis of information in an in-house database with data from approximately three thousand seven hundred and forty microorganisms, seven of which are *S. cerevisiae* strains. MALDI-TOF identification to the strain level was only accepted if *S. cerevisiae* were correctly identified to the genus, species and strain level as a "best hit". The results of the identification were verified by the ARC, Infruitec-Nietvoorbij and NCP Alcohols.

Table 4.3 Interpretation of score values (Bruker, 2007).

Range	Description	Colour
2.300 to 3.000	Highly probable species identification	Green
2.000 to 2.299	Secure genus and probable species identification	Green
1.700 to 1.999	Probable species identification	Yellow
0.000 to 1.699	No reliable Identification	Red

4.3.10 Principle component analysis (PCA)

Principle component analysis (PCA) was employed to statistically evaluate MALDI-TOF mass spectra. PCA is a broadly used mathematical technique for reducing the dimensionality of multivariate data whilst preserving most of the variance (Kettaneh*, et al.*, 2005). PCA was performed using ClinProTools software 3.0 (Bruker, Germany). The PCA is managed by an external MATLAB software tool, which is integrated in ClinProTools. All spectra were normalised before PCA analysis. Complex relationships between samples were explored and illustrated with the use of PCA plots. These models allow one to determine similarities or dissimilarities between samples. For example samples that cluster closer together are more similar than samples that are further apart.

4.4 RESULTS

4.4.1 *S. cerevisiae* database creation

Employing the previously described (Section 3.3.5) optimised ethanol/formic acid extraction procedure; the mass spectral fingerprints of forty five laboratory and commercial *S. cerevisiae* strains were employed in the creation a *S. cerevisiae* mass spectral database. It was found that all of these strains even those that are closely related could be differentiated as depicted in Figure 4.1. Importantly, none of the *S. cerevisiae* strains were observed to be identical. This indicated that the *S. cerevisiae* database contains unique mass spectral profiles for all forty five *S. cerevisiae* strains. Close similarity between the *S. cerevisiae* strains used in this study, was not unexpected as there exists only a limited number of characteristic diagnostic ion peaks that differentiate these strains. From the MSP dendrogram it was shown that these strains could be clustered into five major groups.

It seems that wine yeast strain *S. cerevisiae* VIN13 is the most dissimilar from the other *S. cerevisiae* strains within the database. Also with the exception of VIN7 and NT45, the remaining wine yeast strains appear in the largest cluster (blue, Fig. 4.1). This implies that these strains may share common ancestral origins.

With the exception of strain SAB DY the other brewing strains including *S. cerevisiae* WS CAR, SAB MC UF, SAB MC CO and WS NEW clustered close to one another (red, Fig. 4.1). This suggests that their surface chemistry may be more similar than to *S. cerevisiae* SAB DY. A possible explanation may reside in the fact that South African Breweries have identified strain SAB DY as a problematic contaminant in the beer production process. Interestingly, the industrial ethanol strains labelled ALCOHOL D3 and BAKERS D42 obtained from NCP Alcohols were grouped within this cluster but could clearly be differentiated. The brewing strains (WS NEW and SAB MC CO) also reside within this cluster.

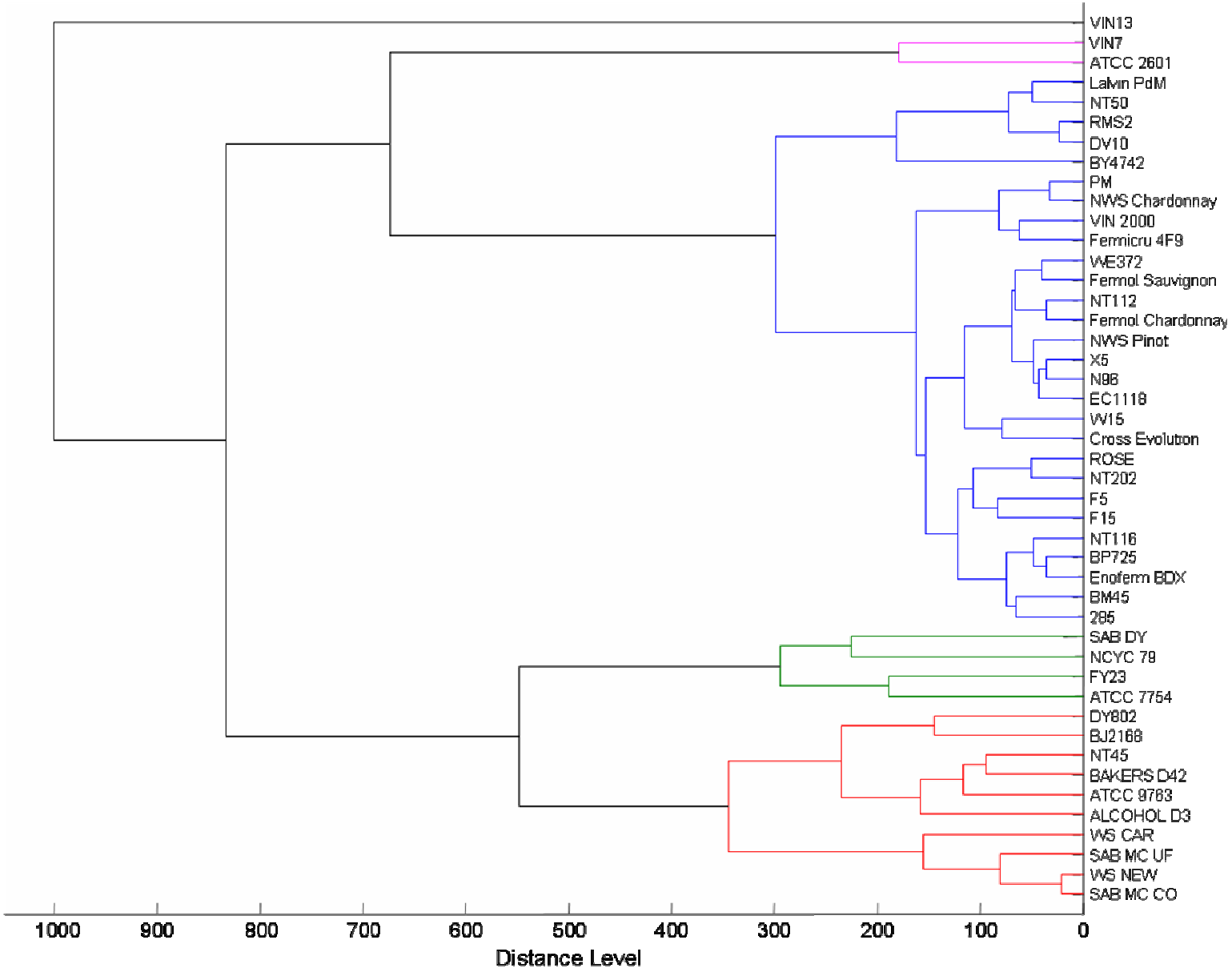


Figure 4.1 Dendrogram created from the mass spectral profile of forty five *S. cerevisiae* strains using the BioTyper 3.0 software

4.4.2 PCR-based yeast strain identification

A PCR method employing delta specific primers as described by Ness *et al*. (1993) was used to discriminate *S. cerevisiae* strains by targeting the δ elements which are repeat sequences that flank the TY1 retrotransposon. Only twenty of the *S. cerevisiae* strains assessed by this method were easily differentiated by the specific polymorphic banding patterns of their amplicons. From the data it is evident that this PCR-based method failed to discriminate certain strains from each other namely, Group A (Fig 4.2), Groups B, C, D and E (Fig 4.3), and Groups F and G (Fig. 4.4). The latter strains represent a total of twenty five *S. cerevisiae* yeast strains that displayed similar monomorphic banding patterns.

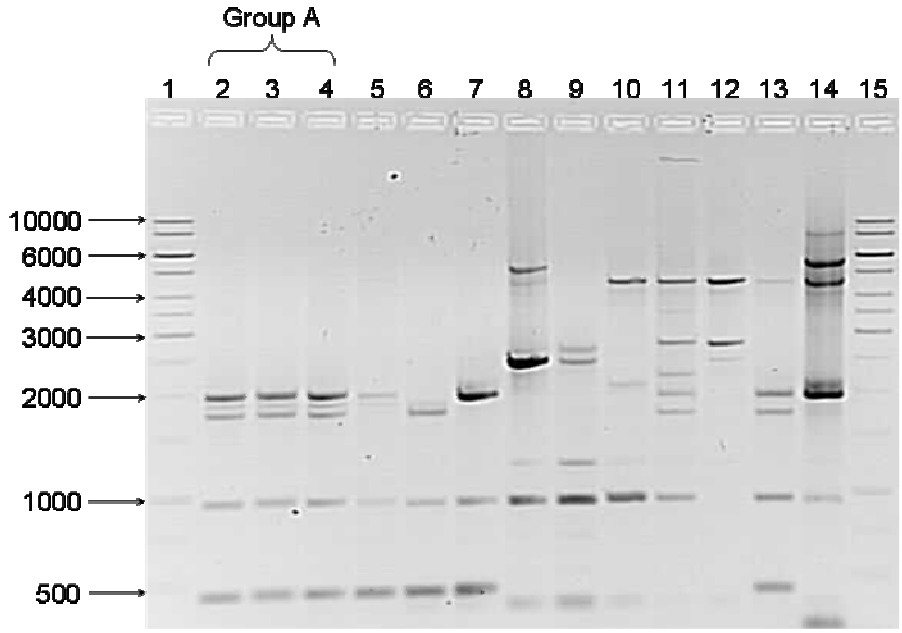


Figure 4.2 PCR profiles of different *S. cerevisiae* wine yeast strains: lane 2: N96, lane 3: DV10, lane 4: EC1118, lane 5: Fermol Sauvignon, lane 6: NT50, lane 7: 285, lane 8: WE372, lane 9: W15, lane10: F15, lane11: X5, lane 12: F5, lane 13: NT116 and lane: 14 BM45. Lane 1 and lane 15: contained the 1 kb DNA.

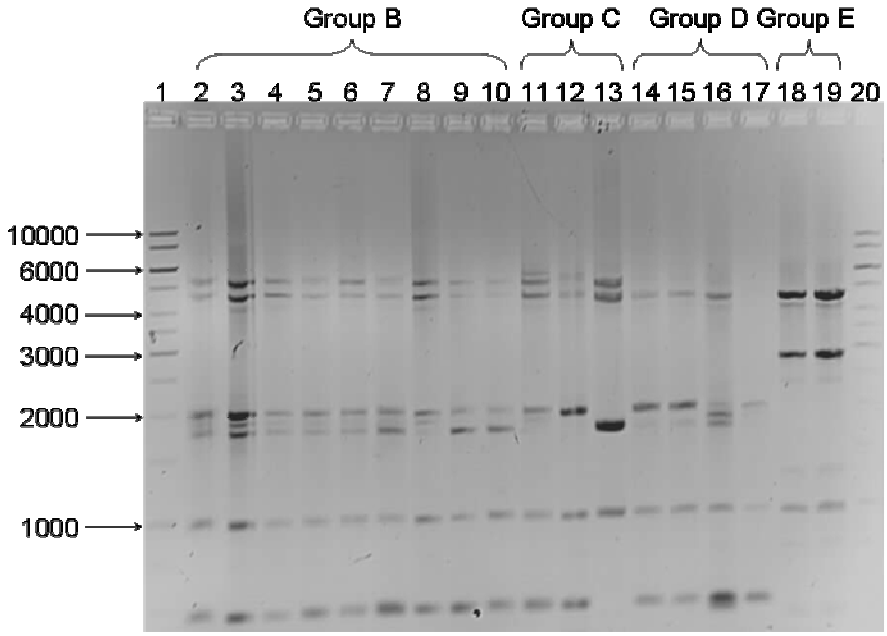


Figure 4.3 PCR profiles of different *S. cerevisiae* wine yeasts strains: lane 2: Lalvin PdM, lane 3: NWS Chardonnay, lane 4: VIN 2000, lane 5: NT202, lane 6: Fermicru 4 F9, lane 7: NT112, lane 8: NWS Pinot, lane 9: RMS2, lane 10: PM, lane 11: NT45, lane 12: Cross evolution, lane 13: ROSE´, lane 14: VIN13, lane 15: Fermol Chardonnay, lane 16: NT116, lane 17: VIN7, lane 18: BP725 and lane 19: Enoferm BDX. Lane 1 and lane 20: contained the 1 kb DNA.

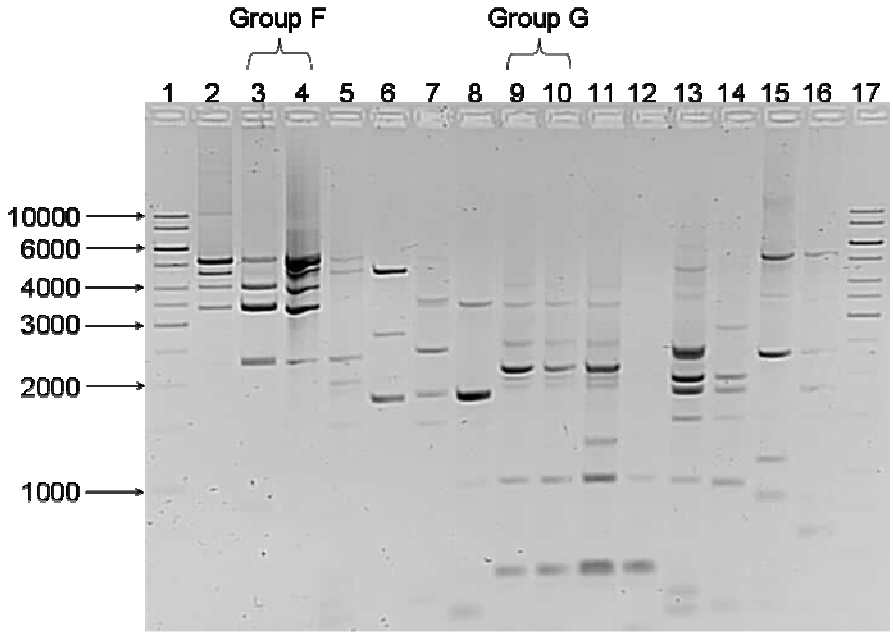


Figure 4.4 PCR profiles of different *S. cerevisiae*: Laboratory strains lane 2: FY23, lane 3: BY4742, lane 4: BJ2168, lane 5: ATCC 7754, lane 6: ATCC 2601, lane 7: NCYC 79, lane 8: ATCC 9763, brewing strains lane 9: SAB MC CO, lane 10: SAB MC UF, lane11: WS NEW, lane 12: WS CAR, lane 13: SAB DY, industrial alcohol strains lane 14: BAKERS D4/2, lane 15: ALCOHOL D3 and lane 16: DY802. Lane 1 and lane 17: contained the 1 kb DNA

4.4.3 CHEF-based karyotype strain differentiation

Contour-clamped homogenous electric field gel electrophoresis (CHEF) was employed to assess the ability of karyotyping to differentiate between a selected subset of the *S cerevisiae* strains employed in this study. From Figure 4.5 it is evident that the karyotypes as resolved by CHEF electrophoresis is capable of distinguishing between six (N96, NT50, WE372, VIN13, BM45, W15) of the strains. However, the karyotype banding pattern of the remaining two strains SAB MC CO and SAB MC UF appear to be identical.

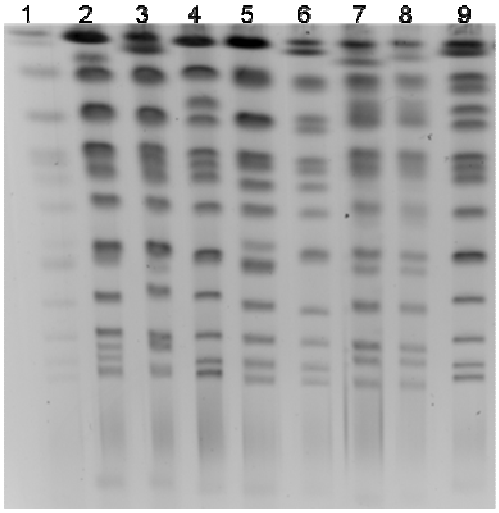


Figure 4.5 CHEF profiles of commercial yeast strains. Lane 1: VIN13, lane 2: N96, lane 3: NT50, lane 4: WE372, lane 5: VIN13, lane 6: BM45, lane 7: SAB MC CO, lane 8: SAB MC UF, lane 9: W15.

Although both the PCR and karyotype-based identification methods failed to discriminate between *S. cerevisiae* strain SAB MC CO and SAB MC UF, MALDI-TOF MS-based differentiation using the multivariate statistical method, principle component analysis unequivocally distinguished between these strains (Fig. 4.6).

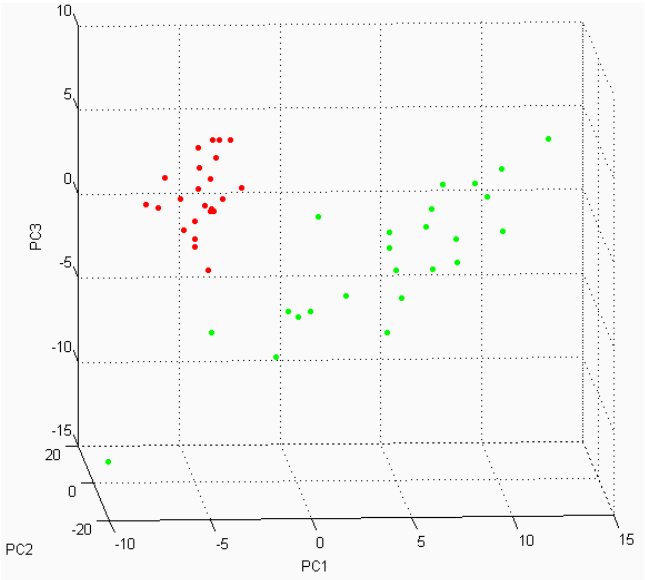


Figure 4.6 PCA of mass spectral signatures of *S. cerevisiae* strains SAB MC UF (red) and SAB MC CO (green).

4.4.4 MALDI-TOF-based biotyping of blind-coded *S. cerevisiae* strains

To evaluate the accuracy of identifying *S. cerevisiae* strains by MALDI-TOF MS, twenty blind-coded *S. cerevisiae* strains were sourced from an independent research facility namely the ARC, Infruitec-Nietvoorbij. The blind-coded samples were analysed by MALDI-TOF MS and the raw spectral data of these samples were compared against the Bruker Daltonics mass spectral database containing the forty five *S. cerevisiae* strain mass spectral profiles (MSP's) recorded in this study (Table 4.4). All twenty *S. cerevisiae* blind-coded strains submitted were identified correctly to the genus and species level. Significantly eighteen of these were also identified to the strain level as well. Therefore MALDI-TOF MS could identify the *S. cerevisiae* to the genus and species level with 100% accuracy and more importantly to the strain level with an accuracy of 90%. However, on visual inspection of blind-coded yeast cultures it must be stressed that the two samples that were incorrectly identified to the strain level appeared to have been contaminated.

To investigate the capacity of MALDI-TOF MS to biotype industrial *S. cerevisiae* strains, seven blind-coded *S. cerevisiae* strains were obtained from a local industrial ethanol producing plant, NCP Alcohols. The blind-coded samples were analysed by MALDI-TOF MS and the raw spectral data of these samples were also compared against the Bruker Daltonics mass spectral database as described above using the Bruker Daltonics BioTyper software programme. Of these seven strains; only four strains were identified correctly (Table 4.5). According to the NCP Alcohols industrial ethanol plant the four *S. cerevisiae* strains that were correctly identified corresponded to pure cream yeast cultures that were obtained from NCP yeast. In addition NCP Alcohols identified the three incorrectly biotyped strains as being recovered from problematic three fermentation tanks that had produced a very low alcohol yield. As such it may be suggested that the low ethanol yield may be attributed to contaminating *S. cerevisiae* strains in these fermentation processes. This view is also supported by PCA analysis of the MALDI-TOF MS spectra of these strains. From the PCA plot it was observed that the three supposedly *S. cerevisiae* ALCOHOL D3 strains obtained from the three fermentation tanks grouped separately from the databased mass spectra of *S. cerevisiae* ALCOHOL D3 (Fig 4.7).

Table.4.4 **MALDI-TOF-based identification of blind-coded *S cerevisiae* strains obtained from ARC, Infruitec-Nietvoorbij using Bruker Daltonics Biotyper 3.0 software.**

Sample	*S. cerevisiae* strain best match	Score value	Identification
1	NT50	2.612	Correct
2	WE372	2.644	Correct
3	BM45	2.551	Correct
4	EC1118	2.565	Correct
5	ROSE	2.593	Correct
6	F5	2.626	Correct
7	F15	2.644	Correct
8	X5	2.652	Correct
9	EC118	2.634	*Incorrect*
10	285	2.605	Correct
11	RMS2	2.687	Correct
12	ATCC 2601	2.650	Correct
13	ATCC 7754	2.720	Correct
14	NCYC 79	2.404	Correct
15	PM	2.445	Correct
16	FY23	2.577	Correct
17	WS CAR	2.424	Correct
18	SAB DY	2.667	Correct
19	FY23	2.663	*Incorrect*
20	DY802	2.558	Correct

Italics represent incorrect identifications.

Table.4.5 **MALDI-TOF-based identification of blind-coded *S cerevisiae* strains obtained from NCP Alcohols using Bruker Daltonics Biotyper 3.0 software.**

Sample	*S. cerevisiae* strain best match	Score value	Identification
Cream yeast A	BAKERS D42	2.524	Correct
Cream yeast B	BAKERS D42	2.672	Correct
Cream yeast C	BAKERS D42	2.809	Correct
Cream yeast D	ALCOHOL D3	2.608	Correct
Cream yeast A, new tank	BAKERS D42	1.773	*Incorrect*
Cream yeast B, new tank	BAKERS D42	2.468	*Incorrect*
Cream yeast C, old tank	BAKERS D42	2.316	*Incorrect*

Italics represent incorrect identifications.

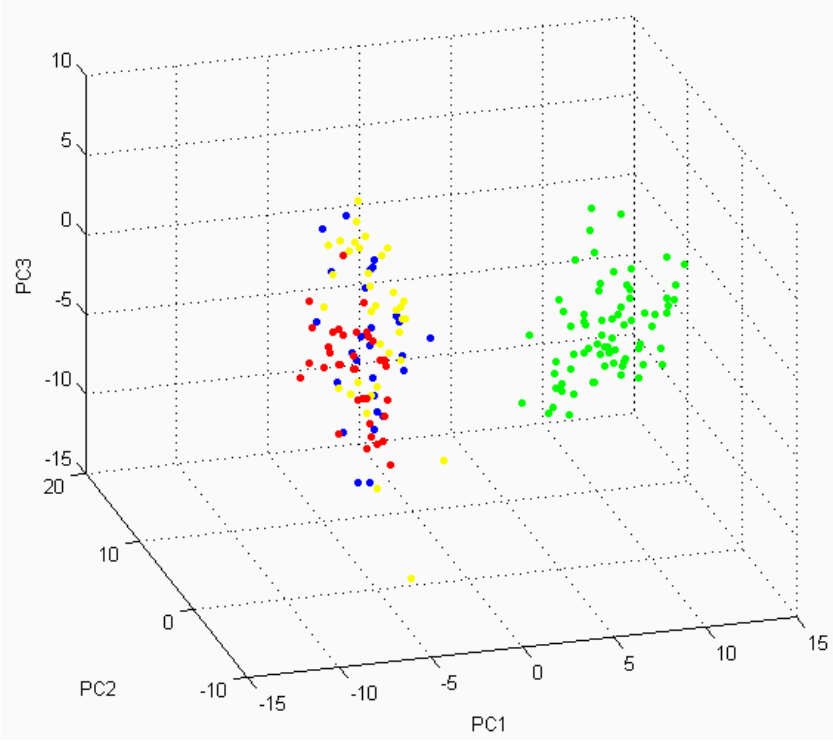


Figure 4.7 3D PCA scores plot of Cream yeast old tank C (yellow), Cream yeast B new tank (blue), Cream yeast A new tank (red) and Alcohol D3 (green).

4.5 DISCUSSION

This is the first report that attempts to create an *S. cerevisiae* MALDI-TOF mass spectral database for the explicit purpose of biotyping these industrially important microorganisms. Previous research studies have primarily focused on the identification of bacteria (Giebel*, et al.*, 2008; Ilina*, et al.*, 2009; Mandrell*, et al.*, 2005) and clinical yeast isolates such as *Candida* (Marklein*, et al.*, 2009; Qian*, et al.*, 2008; van Veen*, et al.*, 2010) using MALDI-TOF MS. These studies relied heavily on the commercially available Bruker Daltonics mass spectral database (Marklein*, et al.*, 2009; van Veen*, et al.*, 2010). The Bruker Daltonics taxonomy tree has a collection of three thousand seven hundred and forty MSP's administered within its database. However, only seven MSP's have been constructed for *S. cerevisiae* strains.

In this study an *S. cerevisiae* database comprising of forty five laboratory and industrial *S. cerevisiae* strains was created. The MSP dendrogram revealed that all forty five *S. cerevisiae* strains grouped separately which suggests that unique mass spectral fingerprints were achievable for all *S. cerevisiae* strains making up the newly created mass spectral database. This taxonomic classification also allows one to determine the relationship between the data. As mentioned previously the largest cluster consisted of the wine yeast strains, implying that they

share common ancestral origins. This is not surprising especially since most modern commercially available wine yeast strains have been derived from wine yeast selection using genetic techniques such as hybridization (mating, spore cell-mating, rare mating, cytoduction and spheroplast fusion), clonal selection of variants and mutagenesis (Pretorius and Bauer, 2002). It was observed that four closely related *S. cerevisiae* grouped closer together such as *S. cerevisiae* SAB MC UF, SABMC CO, WS NEW and WS CAR which as shown in the dendrogram clustered together and had a high degree of similarity. Interestingly the fifth brewing strain *S. cerevisiae* which according to the supplier is a contaminant strain was grouped in a separate cluster.

In recent years molecular biology techniques have developed as the method/s of choice for the identification of *S. cerevisiae* strains to the species and strain level (Barszczewski and Robak, 2006; Buzzini*, et al.*, 2007; Hierro*, et al.*, 2004). In the present study it was found that both PCR using delta primers according to Ness *et al*. (1993) and CHEF as described by Van der Westhuizen and co-workers (1999) were not as discriminative as MALDI-TOF MS biotyping. Brewing strains *S. cerevisiae* SAB MC CO and *S. cerevisiae* SAB MC UF could not be distinguished using either of these molecular biology methods described and were clearly differentiated using MALDI-TOF MS. It should be noted that further optimization of both molecular methods could have lead to greater discrimination of the *S. cerevisiae* strains employed in this study. MALDI-TOF MS was found to be a far superior biotyping tool than PCR and CHEF in terms of duration, expertise and automation (Table 4.6).

Table 4.6 Comparison of biotyping tools

Biotyping tool	Duration (h)	Technical expertise	Automation
PCR	± 14	extensive	no
CHEF	± 156	extensive	no
MALDI-TOF MS	± 2	minimal	yes

Note duration is not inclusive of yeast cultivation time.

Recent studies have shown that identification of yeast using MALDI-TOF MS and the commercially available Bruker Daltonics mass spectral database, identification can only be achieved accurately to the species level (Marklein*, et al.*, 2009; van Veen*, et al.*, 2010). In the present study the created *S. cerevisiae* database was assessed to determine if MALDI-TOF MS could be used to identify *S. cerevisiae* to the strain level and determine the accuracy of this identification. It was found that 90% of blind-coded *S. cerevisiae* strains compared against the *S. cerevisiae* spectral database were identified correctly. This confirms our proof of concept that MALDI-TOF MS can be employed for the identification of *S. cerevisiae* strains provided a comprehensive *S. cerevisiae* strain mass spectral database is available.

For the first time in this study the potential of MALDI-TOF MS as a rapid biotyping tool for industrial *S. cerevisiae* strain identification was established. Seven blind-coded *S. cerevisiae* strains were supplied for MALDI-TOF MS analysis from NCP Alcohols. Positive identifications were obtained for four of the blind-coded *S. cerevisiae* strains. However, three *S. cerevisiae* ALCOHOL D3 strains were incorrectly identified as *S. cerevisiae* BAKERS D4/2. It was found that these three strains unlike the previous four strains were not pure cultures instead were obtained from fermentation tanks that contained a low ethanol yield. From the MALDI-TOF MS data obtained it was found that the low ethanol yield may have been due to contamination. Thus MALDI-TOF MS has great potential as a tool for the rapid identification of industrial *S. cerevisiae* strains and may be applied to ensure quality control of starter cultures and to detect contamination which if left undetected could lead to large scale losses.

4.6 ACKNOWLEDGEMENTS

This work was made possible through financial support from the National Research Foundation and UKZN.

4.7 REFERENCES

Amiri-Eliasi, B. and Fenselau, C. (2001). Characterization of Protein Biomarkers by MALDI from whole fungal cells. *Anal. Chem* **73,** 5228-5231.

Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (1995). *Current Protocols in Molecular Biology*. John Wiley and Sons.

Barszczewski, W. and Robak, M. (2004). Differentiation of contaminating yeasts in brewery by PCR-based techniques. *Food Microbiology* **21,** 227-231.

Barszczewski, W. and Robak, M. (2006). PCR-Based Differentiation and Homology of Brewing and Type strains of Genus *Saccharomyces*. *J. Inst. Brew* **112,** 165-172.

Bruker (2007). Bruker BioTyper User Manual. **Version 2.0**.

Buchaille, L., Freydiere, A. M., Guinet, R. and Gille, Y. (1998). Evaluation of six commercial systems for identification of medically important yeasts. *Eur J Clin Microbiol Infect Dis* **17,** 479-88.

Buzzini, P., Turchetti, B. and Vaughan-Martini, A. E. (2007). The use of killer sensitivity patterns for biotyping yeast strains: the state of the art, potentialities and limitations. *FEMS Yeast Research* **7,** 749-760.

Dombek, K. M. and Ingram, L. O. (1987). Ethanol production during batch fermentation with *Saccharomyces cerevisiae*: changes in glycolytic enzymes and internal pH. *Appl Environ Microbiol* **53,** 1286-91.

Giebel, R. A., Fredenberg, W. and Sandrin, T. R. (2008). Characterization of environmental isolates of *Enterococcus spp*. by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Water Res* **42,** 931-40.

Giusto, C., Iacumin, L., Comi, G., Buiatti, S. and Manzano, M. (2006). PCR-TTGE and RAPD-PCR Techniques to analyze *Saccharomyces cerevisiae* and *Saccharomyces carlsbergensis* isolated from craft beers. *J. Inst. Brew* **112,** 340-345.

Gonzalez-Techera, A., Jubany, S., Carrau, F. M. and Gaggero, C. (2001). Differentiation of industrial wine yeast strains using microsatellite markers. *Letters in Applied Microbiology* **33,** 71-75.

Hierro, N., Gonzalez, A., Mas, A. and Guillamon, J. M. (2004). New PCR-based methods for yeast identification. *J Appl Microbiol* **97,** 792-801.

Ilina, E. N., Borovskaya, A. D., Serebryakova, M. V., Chelysheva, V. V., Momynaliev, K. T., Maier, T., Kostrzewa, M. and Govorun, V. M. (2009). Application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry for the study of *Helicobacter pylori*. *Rapid Commun. Mass Spectrom.* **24,** 328-334.

Kettaneh, N., Berglund, A. and Wold, S. (2005). PCA and PCL with very large data sets. *Comp. Stat. Data Anal.* **48,** 69-85.

Logan, J., Edwards, K. and Saunders, N. (2009). Real-time PCR Current Technology and applications, Caister Academic press.

Lopez, V., Fernandez-Espinar, M. T., Barrio, E., Ramon, D. and Querol, A. (2003). A new PCR-based method for monitoring inoculated wine fermentations. *Int J Food Microbiol* **81,** 63-71.

Mandrell, R. E., Harden, L. A., Bates, A., Miller, W. G., Haddon, W. F. and Fagerquist, C. K. (2005). Speciation of *Campylobacter coli, C. jejuni, C. helveticus, C. lari, C. sputorum* and *C. upsaliensis* by Matrix-Assisted Laser Desorption Ioniztion-Time of Flight Mass Spectrometry. *Appl. Environ. Microbiol.* **71,** 6292-6307.

Manzano, M., Medrala, D., Giusto, C., Bartolomeoli, I., Urso, R. and Comi, G. (2006). Classical and molecular analyses to characterize commercial dry yeasts used in wine fermentation. *Food Microbiol*.

Marklein, G., Josten, M., Klanke, U., Muller, E., Horre, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A. and Sahl, H. G. (2009). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* **47,** 2912-7.

Mellmann, A., Cloud, J., Maier, T., Keckevoet, U., Ramminger, I., Iwen, P., Dunn, J., Hall, G., Wilson, D., LaSala, P., Krostrzewa, M. and Harmsen, D. (2008). Evaluation of Matrix-Assisted Laser Desorption Ionization-Time-of-Flight Mass Spectrometry in Comparison to 16S rDNA Gene Sequencing for Species Identification of Nonfermenting Bacteria. *J. Clin. Microbiol.* **46,** 1946-1954.

Ness, F., Lavellee, F., Dubourdieu, D., Aigle, M. and Dulau, L. (1993). Identification of Yeast strains using the polymerase chain reaction. *J Sci Food Agric* **62,** 89-94.

Pretorius, I. S. and Bauer, F. F. (2002). Meeting the consumer challenge through genetically customized wine-yeast strains. *Trends Biotechnol.* **20,** 426-32.

Qian, J., Cutler, J. E., Cole, R. B. and Cai, Y. (2008). MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers. *Anal Bioanal Chem* **392,** 439-49.

Sheehan, C. A., Weiss, A. S., Newsom, I. A., Flint, V. and O'Donnell, D. C. (1991). Brewing yeast identification and chromosome analysis using high resolution CHEF gel electrophoresis. *J. Inst. Brew* **97,** 163-167.

Sherburn, R. E. and Jenkins, R. O. (2003). A novel and rapid approach to yeast differentiation using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Spectroscopy* **17,** 31-38.

Van der Westhuizen, T. J., Augustyn, O. P. H. and Pretorius, I. S. (1999). The value of long-chain fatty acid analysis, randomly amplified polymorphic DNA and electrophoretic karyotyping for the characterization of wine yeast strains. *S. Afr. J. Enol. Vitic.* **20,** 3-10.

van Veen, S. Q., Claas, E. C. and Kuijper, E. J. (2010). High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *J Clin Microbiol* **48,** 900-7.

Wightman, P., Quain, D. E. and Meaden, P. G. (1996). Analysis of production brewing strains of yeast by DNA fingerprinting. *Lett Appl Microbiol* **22,** 90-4.

Winston, F., Dollard, C. and Ricupero-Hovasse, S. L. (1995). Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. *Yeast* **11,** 53-5.

Chapter 5

GENERAL DISCUSSION AND CONCLUSIONS

5 GENERAL DISCUSSION AND CONCLUSION

5.1 GENERAL DISCUSSION AND CONCLUSION

The discrimination (biotyping) of yeast into specific subgroups (biotypes) or the distinction of individual strains is much needed in industry for purposes such as the detection of contaminants, verification and for the protection of patented strains. Traditionally phenotypic and biochemical methods were used for the identification of *S. cerevisiae* strains however these methods have proved to be time consuming and inaccurate (Andrighetto*, et al.*, 2000; Gonzalez-Techera*, et al.*, 2001). In recent years there has been an emergence of molecular biology techniques developed for the unequivocal identification of *S. cerevisiae* strains such methods include radio-labelled DNA probes (Wightman*, et al.*, 1996), electrophoretic karyotypes by pulsed-field gel electrophoresis (Gomes*, et al.*, 2000; Sheehan and Weiss, 1990), polymerase chain reaction amplification with delta elements (Ness*, et al.*, 1993), intron splice site primers (De Barros Lopes*, et al.*, 1996), micro-satellites or simple sequence repeats (Gonzalez-Techera*, et al.*, 2001; Howell*, et al.*, 2004), randomly amplified polymorphic DNA-polymerase chain reaction (Barszczewski and Robak, 2004; Barszczewski and Robak, 2006), temporal temperature gradient gel electrophoresis (TTGE) (Giusto*, et al.*, 2006), PCR-temperature gradient gel electrophoresis (PCR-TGGE), restriction enzyme analysis (RE) (Manzano*, et al.*, 2006) and real-time PCR (Logan*, et al.*, 2009). However these methods can be labour intensive, time consuming and in some instances not show sufficient polymorphisms to enable differentiation of industrial *S. cerevisiae* strains (Howell*, et al.*, 2004; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003).

Spectroscopic approaches such as pyrolysis-mass spectrometry and fourier transform-infrared spectroscopy can be used for the identification of *S. cerevisiae* strains. The general limitations of these methods are: cumbersome data analysis and the need for skilled interpretation of spectra (Timmins*, et al.*, 1998b). Matrix assisted laser desorption/ionization time of flight mass spectrometry is an emerging identification method for the discrimination of microbes. Studies have shown that yeast do produce characteristic MALDI-TOF MS signals however due to the structure of the yeast cell wall, these signals may be more difficult to obtain than from bacteria and may require cell lysis (Amiri-Eliasi and Fenselau, 2001; Qian*, et al.*, 2008; van Veen*, et al.*, 2010). For MALDI-TOF MS to be used as a biotyping tool, the spectra obtained have to be both diagnostic and reproducible. Studies on the identification of yeast using MALDI-TOF MS have primarily focused on clinical yeast such as *Candida*. Furthermore these studies have only achieved identification of yeast to the species level. A major limiting

aspect of MALDI-TOF MS for microbe identification is the need for a comprehensive database (Marklein*, et al.*, 2009; Qian*, et al.*, 2008; Sherburn and Jenkins, 2003). Studies on yeast to date have relied heavily on the commercial Bruker Daltonics mass spectral database for identification to the species level. Currently the Bruker Daltonics spectral database stands at three thousand seven hundred and forty microorganisms of which only a mere seven are *S. cerevisiae* strains. Thus a more comprehensive *S. cerevisiae* database is critical in an endeavour to efficiently biotype these industrially important microbes. Our data show that forty-five *S. cerevisiae* strains can be identified further to the strain level as long as a comprehensive database is available.

To date, this is the first study to report on the identification of laboratory and industrial *S. cerevisiae* strains using an ethanol/formic acid extraction specifically optimized for this purpose. Sample preparation using the optimized extraction procedure, yielded unique and highly reproducible mass spectral fingerprints for the three *S. cerevisiae* strains employed in this study. The spectra were used to create a reference database which can serve to identify *S. cerevisiae* strains to the strain level. This outcome has not been reported previously. MALDI-TOF MS was found to be a far more sensitive and rapid technique than both PCR and CHEF employed as reference methods in this study. Furthermore MALDI-TOF MS was also shown to have great potential for the identification of industrial *S. cerevisiae* strains and the detection of *S. cerevisiae* contaminant strains. The method was rapid, relatively inexpensive, simple and an identification was achievable in two hours. In this study it was shown that statistical tools such as principle component analysis along side this technique further enable the identification and differentiation of strains.

Future prospects arising from this study that warrant further investigation can be summarized as follows:

- The performance of MALDI-TOF MS as a powerful biotyping tool for *S. cerevisiae* identification in industry can only improve as more spectra of appropriate reference strains are added to the database. In this study only thirty *S. cerevisiae* wine yeast were employed which is only 10% of the commercial *S. cerevisiae* wine yeast strains available on the market (Pretorius, 2000).
- In this study it was observed that by using an optimized ethanol/formic acid extraction procedure specifically designed for *S. cerevisiae* strains, identification could be achieved up to the strain level, if a similar strategy was employed for clinical yeast isolates and strain identification was achievable this would significantly reduce the amount of time required to reach a prognosis.

5.2 ACKNOWLEDGEMENTS

This work was made possible through financial support from the National Research Foundation and UKZN.

5.3 REFERENCES

.

Amiri-Eliasi, B. and Fenselau, C. (2001). Characterization of Protein Biomarkers by MALDI from whole fungal cells. *Anal. Chem* **73,** 5228-5231.

Andrighetto, C., Psomas, E., Tzanetakis, N., Suzzi, G. and Lombardi, A. (2000). Randomly amplified polymorphic DNA (RAPD) PCR for the identification of yeasts isolated from dairy products. *Lett Appl Microbiol* **30,** 5-9.

Barszczewski, W. and Robak, M. (2004). Differentiation of contaminating yeasts in brewery by PCR-based techniques. *Food Microbiology* **21,** 227-231.

Barszczewski, W. and Robak, M. (2006). PCR-Based Differentiation and Homology of Brewing and Type strains of Genus *Saccharomyces*. *J. Inst. Brew* **112,** 165-172.

De Barros Lopes, M., Soden, A., Henschke, P. A. and Lang, P. (1996). PCR Differentiation of Commercial Yeast Strains Using Intron Splice Site Primers. *Appl. Environ. Microbiol.* **62,** 4514-4520.

Giusto, C., Iacumin, L., Comi, G., Buiatti, S. and Manzano, M. (2006). PCR-TTGE and RAPD-PCR Techniques to analyze *Saccharomyces cerevisiae* and *Saccharomyces carlsbergensis* isolated from craft beers. *J. Inst. Brew* **112,** 340-345.

Gomes, L. H., Duarte, K. M. R., Argueso, J. L., Echeverrigaray, S. and Taveres, F. C. A. (2000). Methods of yeast characterization from industrial products. *Food Microbiology* **17,** 217-223.

Gonzalez-Techera, A., Jubany, S., Carrau, F. M. and Gaggero, C. (2001). Differentiation of industrial wine yeast strains using microsatellite markers. *Letters in Applied Microbiology* **33,** 71-75.

Howell, K. S., Bartowsky, E. J., Fleet, G. H. and Henschke, P. A. (2004). Microsatellite PCR profiling of *Saccharomyces cerevisiae* strains during wine fermentation. *Lett Appl Microbiol* **38,** 315-20.

Logan, J., Edwards, K. and Saunders, N. (2009). Real-time PCR Current Technology and applications, Caister Academic press.

Manzano, M., Medrala, D., Giusto, C., Bartolomeoli, I., Urso, R. and Comi, G. (2006). Classical and molecular analyses to characterize commercial dry yeasts used in wine fermentation. *Food Microbiol*.

Marklein, G., Josten, M., Klanke, U., Muller, E., Horre, R., Maier, T., Wenzel, T., Kostrzewa, M., Bierbaum, G., Hoerauf, A. and Sahl, H. G. (2009). Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* **47,** 2912-7.

Ness, F., Lavellee, F., Dubourdieu, D., Aigle, M. and Dulau, L. (1993). Identification of Yeast strains using the polymerase chain reaction. *J Sci Food Agric* **62,** 89-94.

Pretorius, I. S. (2000). Tailoring wine yeast for the new millennium: novel approaches to the ancient art of winemaking. *Yeast* **16,** 675-729.

Qian, J., Cutler, J. E., Cole, R. B. and Cai, Y. (2008). MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers. *Anal Bioanal Chem* **392,** 439-49.

Sheehan, C. and Weiss, A. S. (1990). Yeast artificial chromosomes: rapid extraction for high resolution analysis. *Nucleic Acids Res* **18,** 2193.

Sherburn, R. E. and Jenkins, R. O. (2003). A novel and rapid approach to yeast differentiation using matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Spectroscopy* **17,** 31-38.

Timmins, E. M., Quain, D. E. and Goodacre, R. (1998). Differentiation of brewing yeast strains by pyrolysis mass specrometry and Fourier transform infrared spectroscopy. *Yeast* **14,** 885-893.

van Veen, S. Q., Claas, E. C. and Kuijper, E. J. (2010). High-throughput identification of bacteria and yeast by matrix-assisted laser desorption ionization-time of flight mass spectrometry in conventional medical microbiology laboratories. *J Clin Microbiol* **48,** 900-7.

Wightman, P., Quain, D. E. and Meaden, P. G. (1996). Analysis of production brewing strains of yeast by DNA fingerprinting. *Lett Appl Microbiol* **22,** 90-4.