TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Technische Mikrobiologie

Differentiation of yeasts and their metabolic status by MALDI-TOF Mass Spectrometry

Julia C. Usbeck

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender:		Univ.-Prof. Dr. W. Schwab
Prüfer der Dissertation	1.	Univ.-Prof. Dr. R. F. Vogel
	2.	Univ.-Prof. Dr. S. Scherer

Die Dissertation wurde am 14.04.2016 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 15.08.2016 angenommen.

DANKSAGUNG

Die vorliegende Arbeit entstand im Rahmen eines, aus Haushaltsmitteln des BMWi über die AiF-Forschungsvereinigung Wissenschaftsförderung der Deutschen Brauwirtschaft e. V. geförderten Projektes (AiF 16576 N).

Mein besonderer Dank gilt meinem Doktorvater Prof. Dr. Rudi F. Vogel für die Überlassung des Themas und der Möglichkeit die praktischen Arbeiten an seinem Lehrstuhl durchführen zu können. Er hat mich mit vielen Anregungen, Diskussionen und unangenehmen Fragestellungen gelehrt, den roten Faden nie zu verlieren.

Bei Herrn Prof. Dr. Siegfried Scherer möchte ich mich für die Begutachtung dieser Arbeit bedanken und Herrn Prof. Dr. Wilfried Schwab danke ich für die Übernahme des Prüfungsvorsitzes.

Ohne Dr. Jürgen Behr wäre die Arbeit nicht in diesem Umfang entstanden. Danke für viele, neu gestaltete Auswerte-Tools, Rat bei der Experimentplanung, ein immer offenes Ohr und die stets motivierenden Interpretation meiner Ergebnisse.

Bei Prof. Dr. Matthias Ehrmann und Prof. Dr. Ludwig Niessen bedanke ich mich für ihre Diskussionsbereitschaft und die stete Möglichkeit sie mit fachlichen Fragen zu behelligen.

Bei Dr. Mathias Hutzler und seinem Team vom Forschungszentrum Weihenstephan für Brau- und Lebensmittelqualität möchte ich mich für die Bereitstellung der Stämme ganz herzlich bedanken. Seine Bereitschaft mich mit Fachwissen über Hefen zu versorgen hat mir manch tieferen Einblick in die Praxis ermöglicht.

Der Firma Bruker Daltonik, Bremen, danke ich für ihr Entgegenkommen bei Support, Ersatzteilen und diversen Hilfestellungen im Laufe der drei Jahre mit unserem MALDI-TOF MS.

Caroline Wilde von der Firma Lallemand Inc., Montréal, Canada, danke ich für die Kooperation, die das Überlassen der Weinhefe Isolate und deren Charakterisierung sowie das Korrektur lesen der Veröffentlichung und viele wertvolle Diskussionen einschließt.

Dem Zentrallabor für Proteinanalytik der Ludwig-Maximilian Universität in München danke ich für die LC-MS/MS Messungen.

Bei den Festangestellten des Lehrstuhls möchte ich mich für den reibungslosen Ablauf und die Hilfe bei bürokratischen und organisatorischen Abwicklungen bedanken.

Bei meinen „Mitdoktoranden“ bedanke ich mich für die kollegiale Zusammenarbeit, das angenehme Arbeitsklima, den wissenschaftlichen Austausch und diverse praktische Tipps und Anregungen.

Anna, Juli, Patrick, Jasmin, Ben, Katja, Claudia und Frank danke ich für eine sehr schöne Zeit am und außerhalb des Lehrstuhls. Mit sowohl fachlichen als auch privaten Beiträgen habt ihr maßgeblich zu dem Erfolg dieser Arbeit und meinem persönlichen Wohlbefinden in Freising beigetragen.

Im Besonderen möchte ich mich bei Caro bedanken, die als meine Projektpartnerin einige Schrecksekunden mit mir teilte, so manche Laune abbekam und ohne die ich oft aufgeschmissen gewesen wäre.

Meinen Freunden danke ich, dass sie mir trotz der wenigen Zeit und der teilweise großen Entfernung in den letzten Jahren die Treue gehalten haben.

Meinen Eltern und meiner Schwester Theresa möchte ich danken, dass sie stets hinter mir standen, egal bei welcher Entscheidung und egal wie lange das Ergebnis auf sich hat warten lassen.

Ganz besonders möchte ich Chris danken für zahlreiche Ergebnisdiskussionen, seinen kritischen Blick für Details, der sprachlichen Hilfe und vor allem für seine unendliche Geduld mit mir.

ABBREVIATIONS

°C	degree Celsius
µ	micro (10^{-6})
1*24	24 spectra per strain from one extraction procedure
4*3	4 independent triplicates
3MH	3-mercaptohexan-1-ol
3MHA	3-mercaptohexyl acetate
4MMP	4-mercapto-4-methylpentan-2-one
a	aerobic
aa	anaerobic
AATase	alcohol acetyltransferase activity
a. u.	arbitrary units
ACN	acetonitrile
ADH	alcohol dehydrogenase
ADP	adenosine diphosphate
AFLP	amplified fragment length polymorphism
AHAS	acetohydroxy acid synthase
ALDH	aldehyde dehydrogenase
AmBic	ammonium bicarbonate
ANOSIM	analysis of similarities
APCI	atmospheric pressure chemical ionization
ATP	adenosine triphosphate
BF	bottom-fermenting
BIC	Bayesian Information Criterion
BLQ	Forschungszentrum Weihenstephan für Brau- und Lebensmittelqualität
BT	Biotyper 3.0 software (Bruker Daltonics, Bremen)
BTS	bacterial test standard
BU	bitterness units
C.	*Candida*
cf.	lat.: confer, compare
CHCA	= HCCA
CO_2	carbon dioxide
CoA	coenzyme A
Cr.	*Cryptococcus*
D.	*Debaryomyces*
Da	Dalton
DAPC	discriminant analysis of principal components
DB	database
dH_2O	distilled water
DHB	2,5-dihydroxybenzoic acid
Dk.	*Dekkera*

DNA	deoxyribonucleic acid
DTT	1,4-dithio-D,L-threitol
e.g.	lat.: exempli gratia, for example
E_{kin}	kinetic energy
E_p	potential energy
ER	endoplasmic reticulum
et al.	lat.: et alii, and others
EtOH	ethanol
FA	formic acid
FAB	fast atom bombardment
Fig.	figure
FWHM	full width at half maximum
g	gram
GA	genetic algorithm
h	hour
H_2S	hydrogen sulfide
HCCA	α-cyano-4-hydroxycinnamic acid
HPLC	high performance liquid chromatography
λ	wavelength
i.e.	lat.: id est, that is
IMS	inter membrane space
INRA	National Agricultural Research Institute
IS	ion source
k	kilo
L	liter
LAMMA	laser microprobe mass analyzer
LC	liquid chromatography
LD	laser desorption
LIMPIC	linear MALDI-TOF MS peak indication and classification
m	meter, milli (10^{-3})
M	mol
m/z	mass to charge ratio
MALDI	matrix-assisted-laser-desorption/ionization
MAS	in-house software, based on MASCAP implemented in octave
MASCAP	mass spectrometry comparative analysis package
MCFA	medium chain fatty acid
MDS	multidimensional scaling
mgf	mascot generic format
min	minute
MPI	message passing interface
MS	mass spectrometry
MSP	main spectra projection

n	nano (10^{-9})
NAD	nicotinamide adenine dinucleotide
NADP	nicotinamide adenine dinucleotide phosphate
Nano-ESI	nano electrospray ionization
NCBI	National Center for Biotechnology Information
P.	*Pichia*
PAGE	polyacrylamide gel electrophoresis
PC	principal component
PDC	pyruvate decarboxylase
PDH	pyruvate dehydrogenase
pdr	peak detection rate
PFGE	pulsed-field gel electrophoresis
PIE	pulsed ion extraction
POF	phenolic off-flavor
QC	quick classifier
R.	*Rhodotorula*
RAPD	random amplified polymorphic DNA
RFLP	restriction fragment length polymorphism
s	second
S.	*Saccharomyces*
S. c.	*Saccharomyces cerevisiae*
S. p.	*Saccharomyces pastorianus*
SDS	sodium dodecylsulfate
SNN	supervised neural network
SNR	signal-to-noise-ratio
SOP	standard operation protocol
SRS	sulfate reduction sequence
SVM	support vector machine
T.	*Torulaspora*
TCA	tricarboxylic acid cycle
TEMED	N,N,N',N'-tetramethylethylen-diamine
TF	top-fermenting
TFA	trifluoroacetic acid
TOF	time-of-flight
Tricine	N-[tris(hydroxymethyl)methyl]-glycine
Tris	Tris-(hydroxymethyl)-aminomethane
UPGMA	unweighted pair group method with arithmetic mean
VDK	vicinale diketone
W.	*Wickerhamomyces*
ZfP	Zentrallabor für Proteinanalytik

CONTENTS

1 INTRODUCTION

1.1 Yeasts

Fermented foods and beverages such as beer, wine or bread, have been part of the human diet for over 8000 years (Steensels et al., 2012). Historically, the fermentation was a spontaneous process, relying on naturally occurring microbes, basically yeasts and bacteria that were present in the environment and the raw material. This natural microbiota was gradually replaced by well-defined microbial starter cultures consisting of only one or a few selected microbial strains, when scientists like van Leeuwenhoek, Pasteur, Hansen and others unraveled the underlying mechanisms of fermentation (Steensels and Verstrepen, 2014). Therefore, the majority of today's fermentations processes are based on this principle with *Saccharomyces (S.) cerevisiae* as one of the most prominent representatives.

1.1.1 Yeast carbohydrate metabolism

According to the role of oxygen in their metabolism, yeasts can be classified as (i) obligate aerobes, with an exclusively respiratory metabolism, (ii) facultative fermentatives, capable of using respiratory as well as fermentative metabolism pathways, and (iii) obligate fermentatives (Merico et al., 2007). To give an overview of yeast metabolism, this section focuses on *S. cerevisiae*, since it is the most widely studied species and was particularly used in this thesis. However, it has to be mentioned that in many physiological aspects *S. cerevisiae* is a rather exceptional yeast, since it is one of the few yeasts that is able to grow anaerobically (Visser et al., 1990). Additionally, this yeast shows an unusual behavior during aerobic growth: When grown aerobically at a low rate under sugar limitation, cultures tend to spontaneously synchronize their cell cycle (Kaspar von Meyenburg, 1969; van Dijken et al., 1993). Furthermore, *S. cerevisiae* is a Crabtree-positive yeast, where at high glucose concentrations alcoholic fermentation can occur even at aerobic growth conditions (de Deken, 1966), because glucose represses the respiratory chain (Hommes, 1966). In the following paragraphs, the main metabolic pathways and their influence factors are summarized.

1.1.1.1 Sugar uptake

Monohexose sugars are transported through the plasma membrane of *S. cerevisiae* by facilitated diffusion, which is mediated by hexose transporters (*HXT*) (Bisson et al., 1993; Kruckeberg, 1996). The hexose transporter family of *S. cerevisiae* comprises 18 proteins, i.e. Hxt1-17*,* Gal2*.* Two transporter-like glucose sensors, Snf3 and Rgt2*,* with different affinity towards glucose, where Snf3p senses low and Rgt2p high extracellular glucose levels (Özcan et al., 1998), enable the yeast to detect and respond to environmental conditions by expressing the appropriate transporters (Özcan and Johnston, 1999).

1.1.1.2 Glycolysis

The glycolytic pathway is the major process to oxidize glucose to two molecules pyruvate generating ATP and NADH. An overview is given in Fig. 1.1. Firstly, glucose is activated by the hydrolysis of one ATP to generate glucose 6-phosphate, catalyzed by hexokinases, which are encoded by *HXK1, HXK2,* and *GLK1.* Afterwards, a second ATP molecule is employed to phosphorylate the fructose 6-phosphate to fructose 1,6-bisphosphate catalyzed by the phosphofructokinase, where the expression of the encoded genes *PFK1* and *PFK2* are strongly induced by the glucose concentration, whereas high ATP and citrate concentrations inhibit the activity. Fructose 1,6-bisphosphate is divided into two three-carbon products, which can be reversibly interconverted and which were than oxidized in several steps while electrons were removed by the reduction equivalent NAD. Finally, the overall reaction can be summarized as:

Equation 1:

$$\text{Glucose} + 2\,[\text{NAD}]^+ + 2\,\text{ADP} + 2\,P_i \rightarrow 2\,\text{Pyruvate} + 2\,\text{NADH} + 2\,H^+ + 2\,\text{ATP} + 2\,H_2O$$

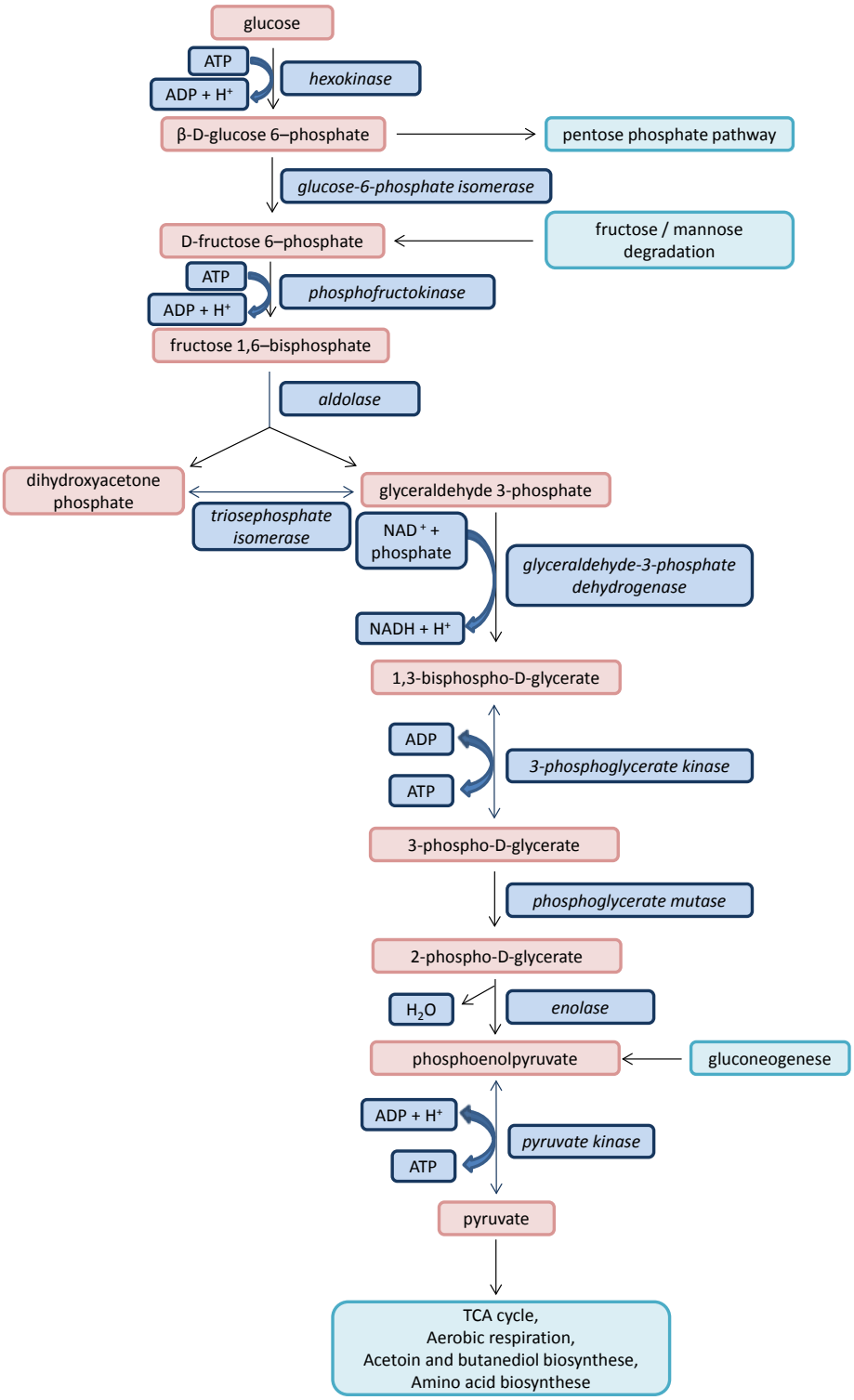

Fig. 1.1: Degradation of glucose to pyruvate via the glycolytic pathway. The intermediates are presented in red, enzymes and reduction equivalents are displayed in blue. Teal boxes indicate pathways that are not depicted in detail (modified according to KEGG pathway sce00010).

1.1.1.3 Pyruvate metabolisms

In Fig. 1.2, key reactions of pyruvate metabolism in *S. cerevisiae* are displayed depending on environmental conditions (Pronk et al., 1996): (i) the transport of pyruvate into mitochondria (**) and the subsequent oxidative decarboxylation to acetyl-CoA catalyzed by the pyruvate dehydrogenase complex (PDH). Acetyl-CoA then enters the TCA cycle (→) (Pronk et al., 1994). (ii) the decarboxylation of pyruvate by cytoplasmic pyruvate decarboxylase (PDC) via the pyruvate bypass to acetaldehyde, which leads to the production of either ethanol (EtOH) by alcohol dehydrogenase (ADH) (→) or acetyl-CoA by acetaldehyde dehydrogenase (ALDH) and acetyl-CoA synthetase (→) (Holzer and Goedde, 1957; Pronk et al., 1994). (iii) the carboxylation of pyruvate to oxaloacetate as a TCA cycle intermediate catalyzed by pyruvate carboxylase (→). Major factors determining the actual fate of pyruvate include the regulation of the enzymes involved and their kinetic properties (Pronk et al., 1996).

The respiratory dissimilation of carbohydrates (i) is predominantly responsible for the conversion of pyruvate to acetyl-CoA during glucose-limited, aerobic growth (Käppeli, 1986; Kresze and Ronft, 1981a; Kresze and Ronft, 1981b; Postma et al., 1989; Pronk et al., 1994) and results in the highest energetic yield (→).

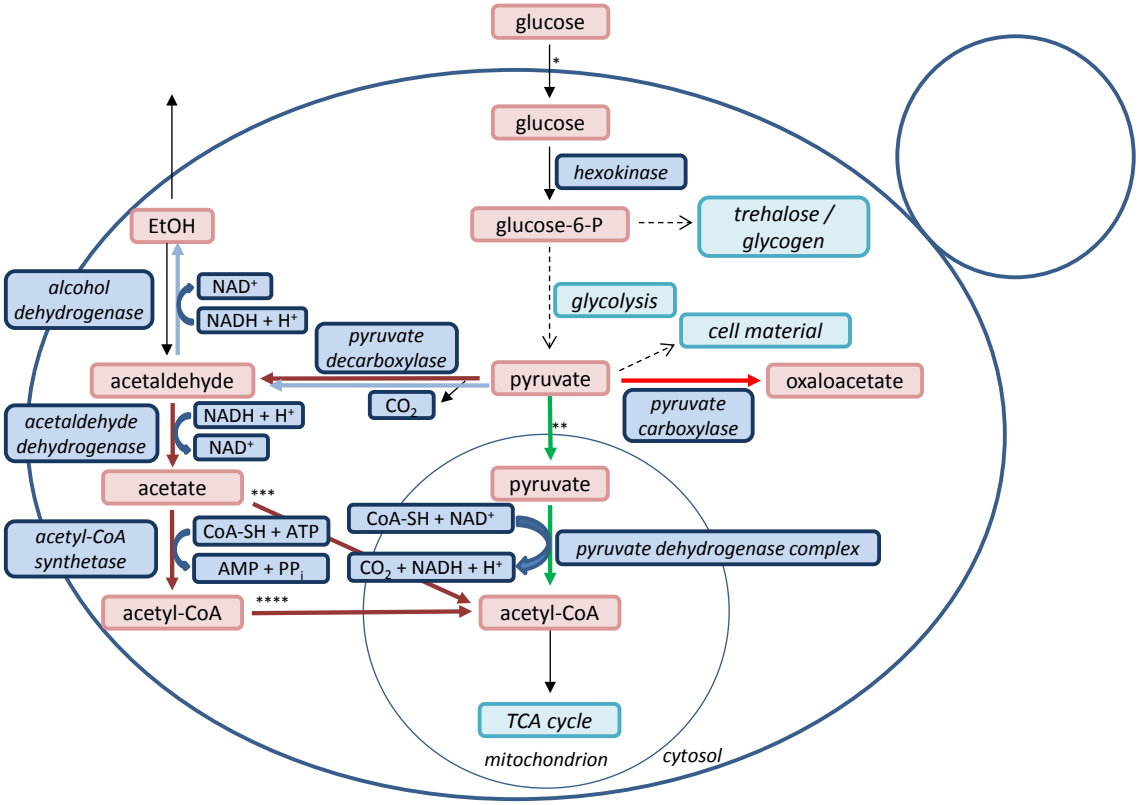

Fig. 1.2: Key enzyme reactions of pyruvate in *S. cerevisiae.* Asterisks indicate transport types, i.e. (*) glucose transport: facilitated diffusion or active transport, (**) mitochondrial pyruvate transport, (***) transport of acetate into mitochondria and formation of acetyl-CoA via mitochondrial acetyl-CoA synthetase, and (****) transport of acetyl-CoA into mitochondria via the carnitine acetyl-transferase shuttle. The intermediates are presented in red. Enzymes and reduction equivalents are displayed in blue. Metabolic pathways that are not depicted in detail are indicated by teal boxes. Colored arrows indicate (→) oxidative decarboxylation (referred to as (i) in the paragraph above), (→) "pyruvate dehydrogenase bypass" (ii), (→) alcoholic fermentation, and (→) pyruvate carboxylase (iii) (modified according to KEGG pathway sce00620).

The K_m for oxidation of pyruvate by isolated mitochondria is approximately tenfold lower than the K_m of PDC (van Urk et al., 1989). Due to the fact that when a substrate concentration exceeds the K_m of an enzyme, the substrate is converted by the enzyme with a higher K_m (Schönenberg, 2010), pyruvate is preferentially channeled into the TCA cycle (cf. 1.1.1.4) only when the intracellular pyruvate concentration is low. The respiratory pathway can be replaced by the pyruvate bypass (→), which involves ATP hydrolysis and, therefore, requires a considerable amount of energy. However, this bypass is of physiological significance, because it provides cytosolic acetyl-CoA, which is necessary for lipid synthesis and not repressed by high glucose amounts, as it was demonstrated by Flikweert et al. (1996) and Pronk et al. (1996). Indeed, the bypass route is favored during high glucose availability (0.1 g/L as the limit for Crabtree-positive *S. c.*) or when the intracellular pyruvate concentration is high. In the first step, pyruvate is decarboxylated to acetaldehyde by PDC, which shows a 3-4 times increased activity

(Rodrigues et al., 2006). Afterwards acetaldehyde can be either reduced to EtOH (→) or oxidized to acetate (→). However, since the amount of active ALDH is decreased under such conditions (high glucose or high pyruvate levels), alcoholic fermentation is favored.

Under anaerobic conditions, fermentation takes place, whereas under aerobic conditions the reduction of acetaldehyde to EtOH is performed by Crabtree-positive yeasts (Piškur et al., 2006). This is mainly due to the glucose repression of the respiratory pathway (de Deken, 1966; Johnston, 1999) and serves as a recycling reaction for NADH with acetaldehyde as the electron acceptor. Accordingly, *S. cerevisiae*, as a Crabtree-positive yeast, regulates respiration and fermentation in response to sugar availability (de Deken, 1966; Gancedo, 1998). A glucose pulse induces alcoholic fermentation in the presence of oxygen and activates PDC and Adh1, which is the isoenzyme involved in alcoholic fermentation. Whereas transcription of *ADH1* is up-regulated during growth on glucose (Denis et al., 1983; van Urk et al., 1990), *ADH2* is glucose-repressed (Dickinson and Kruckeberg, 2006). This is different in the majority of other known yeasts responding to oxygen availability rather than glucose concentration (van Dijken and Scheffers, 1986; Visser et al., 1990). Finally, when glucose is depleted and oxygen is still available the so called “diauxic shift” takes place, where accumulated EtOH can be used as substrate and converted back to acetaldehyde by Adh2 (Denis and Young, 1983; Pronk et al., 1996). In Crabtree-negative yeasts the PDC activity is on average 6-fold lower than in Crabtree-positive ones, but ALDH and acetyl-CoA synthetase show relatively high activity. Accordingly, the pyruvate bypass can effectively oxidize acetaldehyde in Crabtree-negative yeasts (van Urk et al., 1990).

The anaplerotic synthesis of oxaloacetate from pyruvate catalyzed by pyruvate carboxylase (→) is a magnesium- and ATP-dependent carboxylation of pyruvate. This pathway where biotin in the cytosol acts as a co-enzyme (Rohde et al., 1991) is stimulated by acetyl-CoA (Gailiusis et al., 1964; Ruiz-Amil et al., 1965). The withdrawal of TCA-cycle intermediates for biosynthesis is compensated by this reaction (Pronk et al., 1996), which, therefore, presents a part of the gluconeogenesis.

1.1.1.4 TCA cycle

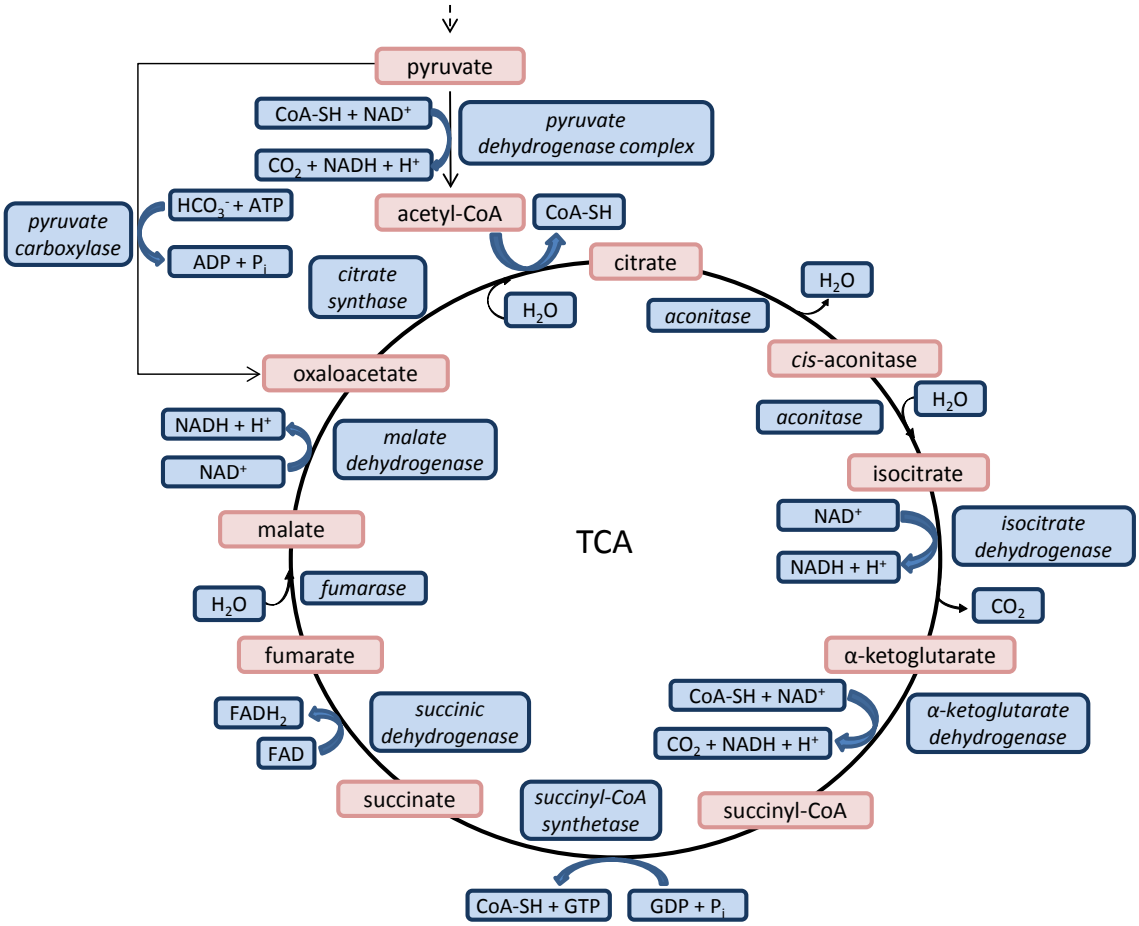

Fig. 1.3: Tricarboxylic acid (TCA) cycle located in the mitochondrion. The intermediates are presented in red. Enzymes and reduction equivalents are displayed in blue (modified according to KEGG pathway sce00020).

The main catalytic function of the TCA cycle is to provide reducing equivalents to the respiratory chain through the oxidative decarboxylation of acetyl-CoA. As detailed above, acetyl-CoA can be derived from pyruvate via PDH, from EtOH or acetate via acetyl-CoA synthase, or from fatty acid oxidation. It enters the TCA cycle reacting with oxaloacetate to form citrate, which is catalyzed by citrate synthase. The TCA cycle provides several intermediates that are crucial to the de novo biosynthesis of several amino acids and cellular compounds. Similar to the other metabolic pathways mentioned above, the TCA cycle is regulated in response to environmental conditions. An example can be found in high glucose and glutamate concentrations repressing aconitase, which is responsible for the isomerization of citrate to isocitrate.

1.1.2 Brewing strains

Traditionally, two types of *Saccharomyces* yeasts are involved in beer fermentation, originally classified along their flocculation behavior, i.e. top-fermenting (TF) and bottom-fermenting (BF) yeasts. These groups can also be distinguished taxonomically by the ability of BF strains to ferment the disaccharide melibiose enabled by the production of the extracellular enzyme α-galactosidase (melibiase) (Stewart et al., 1984). Genetically, the TF group consists of *S. cerevisiae* (*S. c.*) strains, whereas the BF yeast is known under a variety of names such as *S. carlsbergensis* or *S. uvarum* and nowadays generally designated as *S. pastorianus* (*S. p.*) (Vaughan-Martini and Martini, 2011). It presents a natural hybrid organism in which chromosome sets from two parental species are present. However, the exact parentage was a matter of dispute for long time and suggested to involve *S. c.* and most likely a kind of *S. bayanus* (Dunn and Sherlock, 2008; Nakao et al., 2009; Ogata et al., 2008; Rainieri et al., 2006; Yamagishi and Ogata, 1999). In 2011, Libkind et al. discovered the supposedly true origin of the second parent in Patagonia as *S. eubayanus,* which delivered an almost exact genetic match with the non-*S. c.* part in the yeast.

In addition to their genetic background, the brewing yeast strains vary in their technological and biochemical properties. TF strains are employed for ale fermentation, which is conducted at temperatures between 18 and 25 °C, whereas BF strains are applied for lager fermentation, carried out at temperatures between 6 and 15 °C (Gibson et al., 2007). Furthermore, the ability to form flocs that will either rise to the surface of the bulk medium due to entrapped gas bubbles (ale yeasts) or sediment to the bottom of the fermentation tanks (lager yeasts) (Dengis and Rouxhet, 1997b) is of considerable importance for the brewing industry, as it provides an effective, simple, environmental-friendly, and cost-free way to remove yeast cells from green beer at the end of fermentation (Verstrepen et al., 2003b). By contrast, weak flocculation forces the brewer to employ filters or centrifugation systems, which is expensive, time-consuming, and complicates repitching of yeasts (van Mulders et al., 2010).

The comparative analysis of the flocculation of top- and bottom-fermenting yeasts by Dengis et al. (1995) and Dengis and Rouxhet (1997a) showed that the underlying mechanisms appear to be completely different. For BF strains the onset of flocculation occurs during the stationary growth phase, which is controlled by the appearance of flocculins at the cell surface, which interacts with surface located mannose chains of other cells (Miki et al., 1982) and by the decrease in sugar concentration in the medium, since free sugars occupy the flocculin binding sites so that they can no longer bind to the

mannose residues at the cell surface of other cells (Verstrepen et al., 2003b). Additionally, flocculation specifically requires Ca^{2+} so that flocculins achieve the active conformation (Miki et al., 1982; Stratford, 1992). It occurs in a narrow pH range different from the isoelectric point, and is not influenced by EtOH (Dengis et al., 1995). In contrast, the flocculation process of TF yeasts takes place without Ca^{2+} and is induced by EtOH or other organic solvents (Dengis and Rouxhet, 1997a), which is possibly due to the adsorption of EtOH to the yeast surface reducing the local dielectric constant and, thus, decreasing electrostatic repulsion between single cells (Dengis et al., 1995). Additionally, flocculation of TF yeasts occurs at the cell isoelectric point.

Flocculins are encoded by specific genes, so-called *FLO* genes, whose transcriptional activity is influenced by the nutritional status and other stress factors. In particular, the expression of *FLO1* and the highly homologues genes *FLO5* and *FLO9*, only present in *S. c.*, encode a mannose-specific, lectin-like protein, which causes flocculation of the "Flo1 phenotype". *Lg-FLO1,* is responsible for the flocculation phenotype of BF strains coding for a mannose/glucose-specific, lectin-like protein called "NewFlo1 phenotype". The most important characteristic of this phenotype is that flocculation is inhibited during fermentation when the concentration of typical wort carbohydrates is sufficiently high (Kobayashi et al., 1998; Ogata et al., 2008). Accordingly, competitive binding only occurs, when all fermentable sugars are depleted. Surprisingly, the *Lg-FLO1* gene was also detected in three ale yeast strains by van Mulders et al. (2010).

In addition to these general differences, a great variety of different top- and bottom-fermenting strains is available for the production of beer, influencing the characteristic taste of the final product with their significantly varying biochemical and technological peculiarities. The primary metabolites are EtOH and CO_2, which both have inhibitory effects on cell growth. Secondary metabolites include esters, fusel alcohols, aldehydes, organic acids, fatty acids, vicinal diketones, and some sulphur components (Romano et al., 2006). Esters, which are the largest group of flavor compounds, can be divided into two main groups of flavor active esters, i.e. (i) *acetate esters* and (ii) *ethyl esters*. (i) *Acetate esters* are formed intracellularly by a reaction of acetate and EtOH or higher alcohols with the most important representatives being ethyl acetate (solvent-like), phenylethyl acetate (roses, honey) (Saerens et al., 2008a), and isoamyl acetat, which is the main component in the ester fraction and has a typical banana flavor (Nykänen and Nykänen, 1977). Being lipid-soluble, acetate esters diffuse completely through the membrane into the medium and are therefore the quantitatively dominant portion. (ii) *Ethyl esters* are formed with EtOH as the alcohol group and medium-chain fatty acids

(MCFA), with their most prominent representatives ethyl hexanoate (anise seed, applelike aroma), ethyl octanoate (sour apple aroma), and ethyl decanoate (floral odor). Depending on the chain length of the fatty acid, the ethyl esters are increasingly transferred to the medium with decreasing chain length, e.g. 100 % of ethyl hexanoate (C_6) are transferred to the medium compared with 83-92 % of ethyl decanoate (C_{10}) remain in the cell (Nykänen and Nykänen, 1977; Verstrepen et al., 2003a). Basically, the concentration of the substrates and the total activity of the enzymes involved in the formation are important for the rate of ester formation, where the enzyme activity varies depending on the used strain type (Peddie, 1990). This is possibly due to differences in the alcohol acetyltransferase activity (AATase) (Hiralal et al., 2014) (well reviewed by Mason and Dufour (2000)), which, in turn, is influenced by the temperature: larger ester amounts are transferred to the medium at higher temperatures (Suomalainen, 1981). Furthermore, lager yeasts retain more esters in the cell compared with ale yeasts (Lodolo et al., 2008). Additional factors that influence the aroma profile include high gravity brewing, which results in an overproduction of acetate esters, and the capacity of fermentation vessels, which generally became larger and larger due to economic reasons finally leading to the inhibition of yeast metabolism caused by CO_2 as reported by Meilgaard (2001).

So finally, the choice of the employed strain has a deep impact on the final product. An overview is provided in Tab. 1.1. Many small breweries owe their success the uniqueness of their product caused by special yeast strains and individual conditions. Therefore it is important to control and to differentiate brewing yeasts at strain level. Generally it can be assumed that *S. c.* yeasts are genetically more diverse, whereas lager yeasts are more conserved (Lodolo et al., 2008). Wheat beer yeast, for instance, produces beer with spicy, clove, vanilla, and nutmeg flavor notes due to the presence of the POF (phenolic off-flavor) gene (*POF1, Pad1*) (Meaden and Taylor, 1991). The compound 4-vinylguaiacol, which is formed by thermal (Fiddler et al., 1967) or enzymatic decarboxylation of the highly abundant hydroxycinnamate ferulic acid (Zupfer et al., 1998) is responsible for this phenolic flavor effect. The use of wheat and wheat malt contributes to a higher level of ferulic acid, but there are some other factors influencing the final amount of this unwanted compound. Such factors include the initial mash temperature and pH, techniques and duration of wort filtration, the application of beer filter aids removing phenolic compounds (Papp et al., 2001). Most importantly, the volatile profile is highly yeast strain depended (Coghe et al., 2004), which makes the selection of suitable strains a crucial factor for the production of good-tasting beer. In comparison with wheat beer, the TF “Altbier” is old style ale, still popular in Düsseldorf.

It is fermented at relatively warm temperatures and is subjected to a lagering or cold post-fermentation conditioning stage. This contributes to smoothness and increases full blended flavor, it is estery and copper-brown. To improve foam quality, wheat malt is sometimes added (Hardwick, 1995).

Tab. 1.1: Classical beer types (modified according to Hardwick (1995); (Burberg and Zarnkow, 2009; Jackson, 1999))

		Type	Character	Origin	Alc. [% w/v]	Flavor Appearance
Top-fermented (TF)	Barley	Ale	Ale	UK, US, Canada	2.5-5.0	Hoppy, estery, bitter
		Altbier		Düsseldorf	4.0	Estery, bitter
		Kölsch		Cologne	4.4	Light, estery, hoppy
		Porter	Stout	London, US, Canada	5-5.7	Very malty, rich
		Stout		Ireland	4-7	Dry, bitter
		Strong Ale		UK	6-8.4	Estery, heavy, hoppy
	Wheat	Lambic	Acidic ale	Brussels	5+	Acidic, estery
		Berliner Weisse	Lager	Berlin	2.5-3	Lightly flavored, mild
		Wheat beer	Lager/Ale	Bavaria	5-6	Full bodied, low hops
Bottom-fermented (BF)		Bock	Lager	Bavaria, US, Canada	6	Full bodied
		Doppelbock	Lager/Ale	Bavaria	7-13	Full bodied, estery
		Dortmunder	Lager	Dortmund	5+	Light hops, dry, estery
		Light beers	Lager	US	4-5	Light hops, light body
		Münchner	Lager/Ale	Munich	4-4.5	Malty, dry, moderate bitter
		Pilsner	Lager	Pilsen	4.5-5	Full bodied, hoppy
		Vienna (Märzen)	Lager	Vienna	5.5	Full bodied, hoppy

1.1.3 Wine strains

Nowadays, wine production by selected *Saccharomyces* strains, which are commercially available as active dried yeasts, is a widely spread oenological practice to improve the reproducibility in fermentation and final quality. On the other hand, specific strains are used to achieve an individual flavor, because, in addition to the production method and fermentation conditions, the choice of the employed strain can largely influence the

resulting product and makes it unique. In addition to the amount of EtOH and CO_2 produced during fermentation, the final concentration of a large number of metabolic by-products, e.g. acetic, succinic, or lactic acid as well as higher alcohols, glycerol, and volatile compounds, can vary significantly (Antonelli et al., 1999; Lilly et al., 2000; Mateo et al., 1992; Richter et al., 2013; Swiegers et al., 2009).

The most important strains for alcoholic fermentation of grape wine belong to the species *S. c.* and *S. uvarum* (*S. bayanus* var. *uvarum*) (Vaughan-Martini and Martini, 2011), which are highly adapted to grow on substrates characterized by a low pH (2.9 - 3.8) and high osmolarity (sugars of 200 – 300 g/L) (Pizarro et al., 2007). They belong to the genus *Saccharomyces*, formerly known as *Saccharomyces sensu stricto* complex (Rainieri et al., 2003; Vaughan-Martini and Martini, 2011). This complex contains further species including *S. cariocanus, S. kudriavzevii, S. mikatae* (Naumov et al., 2000a), and *S. paradoxus*. Other species of this complex such as the recently found species from China, *S. arboricola* (Naumov et al., 2010; Wang and Bai, 2008), and the hybrid *S. p. (S. c. x S. bayanus* var. *bayanus)* with its wild parent *S. eubayanus* from Patagonia (Libkind et al., 2011) are not likely to play an important role in wine fermentation (Replansky et al., 2008; Sipiczki, 2008). Furthermore, hybrid strains are often used in the wine industry to introduce new/unique sensory attributes (Bellon et al., 2011) or to influence the fermentation performance under different conditions. For example, strain Anchor VIN7, used for cold white wine such as Riesling and Sauvignon Blanc, contains a complete set of chromosomes from *S. kudriavzevii* and *S. c.* lineages (Borneman et al., 2012). The role of the *S. kudriavzevii* genome in the hybrid was described by Peris et al. (2012), claiming its importance in the maintenance of good fermentative performance at low temperatures, which is combined with the glucose and EtOH tolerance of *S. c.* (González et al., 2006). Sparkling wine is preferentially fermented by hybrids of *S. c. x S. uvarum* (*S. bayanus* var. *uvarum*), because they ferment well at a wide temperature range, sediment rapidly (a plus for bottle fermentation), and form minor products of fermentation (glycerol, succinic, acetic, and malic acid) compared with their parental strains (Coloretti et al., 2006). However, different demands depending on varying grape types impede a free strain selection. Chardonnay is typically characterized by ethyl esters (ethyl butanoate, hexanoate, octanoate and decanoate) and the acetate ester, hexyl acetate. These compounds are associated with a specific fruity aroma (Richter et al., 2013), whereas Sauvignon Blanc is divided in a “green” class, in which mainly methoxypyrazines are responsible for the grassy, green pepper character, and a “tropical” (grapefruit, passion fruit) class (Coetzee and du Toit, 2012), owing their characteristics to volatile thiol compounds.

It is a tightrope walk for winemakers to enhance the production of beneficial volatile thiols and simultaneously erase the production of undesirable H_2S and mercaptans.

There are three main grape-derived thiol compounds, which are desired in the product and released from cysteine and glutathione conjugates by yeast carbon-sulfurylase enzymes during fermentation. Namely these beneficial volatile thiols are 4-mercapto-4-methylpentan-2-one (4MMP), 3-mercaptohexan-1-ol (3MH), and its esterified form 3-mercaptohexyl acetate (3MHA), which is converted by alcohol acetyltransferase (Subileau et al., 2008; Swiegers et al., 2009; Swiegers et al., 2006). These are highly potent aroma substances due to their perception thresholds in the ng/L range (King et al., 2011) and of particular importance in Sauvignon Blanc wines, where they impart “box-tree” (4MMP), “passion fruit”, “grapefruit”, “gooseberry” and “guava” aromas (3MH, 3MHA). Depending on the β-lyase encoding gene *IRC7*, the amount of 4MMP can be influenced (Thibon et al., 2008), whereas the *ATF1* gene encodes the alcohol acetyltransferase to transform 3MH to 3MHA (Swiegers et al., 2006).

On the other hand, H_2S is formed in cases of organic sulfur deficiency, which triggers the yeast to synthesize the required organic sulfur amount from inorganic sources through the sulfate reduction sequence (SRS) pathway (Rauhut, 1994; Swiegers and Pretorius, 2007). It has been demonstrated by overexpression studies that the genes *MET14*, encoding an adenosylphosphosulfate kinase and *SSU1*, encoding a sulfite transporter, regulate the formation (Donalies and Stahl, 2002).

Large variations in the mentioned abilities of different commercially available yeast strains make the strain selection an extremely important tool to modulate wine aroma.

1.1.4 Spoilage strains in beverages

In the beverage industry, culture yeasts are used as starters for fermented drinks such as beer, wine, sake, or cider. However, “wild yeasts” unequal to the production strain can contaminate beverages or starter cultures, which does not necessarily lead to the spoilage of a product. By definition, spoilage strains mandatorily cause direct or indirect spoilage of the end product (Priest, 1981), which is often characterized by the production of off-flavors, e.g. phenolic, acidic, fatty acid, and estery notes as well as hazes, turbidity, and EtOH loss (Lawrence, 1988). This can occur at infection levels as low as one yeast per 10^7 culture yeasts (Thurston, 1986) and leads to considerable economic losses for the beverage industry (Fleet, 2006).

In many cases, spoilage cannot be easily defined and largely depends on the desired characteristics of the end product. For example, the brewing yeast can be regarded as a contaminant if it is detected in beer after filtration (Eidtmann et al., 1998). Similarly, the production of 4-ethyl-phenol in red wines by *Brettanomyces/Dekkera* spp. contributes to desirable aromatic notes of spices, leather, smoke, or game, when an amount of up to 400 µg/L are present, whereas higher levels from 620 µg/L upwards are regarded as spoilage (Chatonnet et al., 1992). These examples clearly show that spoilage is a question of definition. A distinction can be made between primary and secondary contamination. Using the example of a brewery, sources for primary contaminations comprise raw material such as malt, hops, water, or the pitching yeast in combination with the equipment of complex pipelines, brewhouse vessels, and, finally, the packaging material, which can be heavily contaminated when returned from trade and, hence, should be properly cleaned (Vaughan et al., 2005). Secondary contaminations can occur during bottling, canning, or kegging, e.g. when unsealed bottles run from the cleaner to the filler and from there to the point of sealing, where airborne microorganisms have time enough to contaminate the product (Storgårds, 2000). Furthermore, biofilm formation enables protection of a complex consortium of microorganisms during cleaning and disinfection by slime production (Storgårds, 2000). Biofilms can persist even when the microorganisms are exposed to high hydrodynamic strengths and the removal success generally decreases with increasing biofilm age (Brugnoni et al., 2012).

1.1.4.1 Spoilage of alcoholic beverages – on the example of beer

Two major challenges need to be overcome to successfully spoil beer. Firstly, beer displays an adverse environment with restrictions for microbiological growth due to the presence of EtOH, CO_2, hop compounds (iso-α-acids), a low level of nutrients and oxygen, and a low pH in the final product. Thus, relatively few species are able to grow in beer (depending on the exact beer type). Secondly, the availability of nutrients markedly changes during the brewing process, which requires an adjustment of the microorganisms to different environments. For example, there is an increased supply of fermentable carbohydrates and nitrogen sources during malting and mashing followed by their depletion when fermentation proceeds. Oxygen availability is limited to the raw material, the initial period of fermentation, and partly during packaging. Additionally, temperature ranges from those suitable for thermophiles in mash and sweet wort down to refrigerated temperatures during storage (Lawrence, 1988). Therefore, a plurality of different species appears during this process.

Depending on meteorological conditions, molds can infect and colonize barley grains and produce potent carcinogenic mycotoxins (van Rensburg et al., 1985) such as deoxynivalenol (DON) and zearalenone (ZEA) (Niessen et al., 1993) produced by *Fusarium sp.*, of which some species are known to additionally induce gushing (Sarlin et al., 2005). Furthermore, aflatoxins can be produced as a metabolic byproduct of *Aspergillus* and *Penicillium* species. A major problem associated with this type of spoilage is that the toxic or gushing inducing substances cannot be easily removed and might largely persist during the brewing process (Schwarz et al., 1995).

Among bacteria, the most detrimental beer spoilers can be found in *Lactobacillus spp. (L.),* with nearly half of the reported incidents caused by *L. brevis* due to their hop resistance (Kern et al., 2014a; Preissler, 2011; Priest and Campbell, 2003; Sakamoto and Konings, 2003; Suzuki et al., 2006). Another 20 % are associated with *L. lindneri* (Back, 1994) and *Pediococcus* spp. (Sakamoto and Konings, 2003). As anaerobic representatives, *Pectinatus* spp. (Haikara et al., 1981; Kern et al., 2014b; Lee et al., 1980; Tholozan et al., 1997) and *Megasphaera* spp. (Haikara and Helander, 2006; Paradh et al., 2011) are examples for bacteria that were reported to cause spoilage. An extensive overview is given by Sakamoto and Konings (2003), Suzuki et al. (2006) and Vaughan et al. (2005).

Last but not least: wild yeast! To provide greater clarity, wild yeasts are traditionally divided into *Saccharomyces* and non-*Saccharomyces* species (Back, 1987) including the genera *Brettanomyces, Candida, Dekkera, Debaryomyces, Hanseniaspora, Pichia* (Campbell and Msongo, 1991), *Torulaspora, Zygosaccharomyces*, and many more (van der Aa Kühle and Jespersen, 1998). They occur in all production steps and attack the brewing process by metabolic products causing turbidity and undesirable flavor (Lawrence, 1988) on the one hand and, on the other hand, they compete against the brewing yeast for nutrients so that fermentation gets stuck. Thus, ensuring fermentation under ideal conditions for growth of the production strain, hygienic packaging processes (depending on the choice of the material), equipment design avoiding rough surfaces and dead space (Kumar and Anand, 1998; Shi and Zhu, 2009), and the filler hygiene are the most important critical control points for yeast spoilage.

In the last decade, ready-packed new low-alcohol and value-added products, such as fusion drinks, which consist of alcoholic and alcohol-free beverages in various ratios, increasingly gain popularity. Hereby, the spoilage potential largely depends on the recipe and the process control. For example, it can be assumed that lemonade and its raw materials initially have a low risk potential, which significantly increases with the mixing

process and in the final product with higher sugar and lower EtOH contents (Hutzler et al., 2008).

1.1.4.2 Spoilage of alcohol-free beverages

Soft drinks are available carbonated or non-carbonated, with or without fruit juice, and often supplemented with organic acid preservatives. Such generally nutrient-poor media can be spoiled by relatively few organisms, i.e. usually yeasts and a few acid tolerant bacteria and fungi (Jay and Anderson, 2001), which are able to grow at a low pH, high sugar content, and refrigeration temperatures. Carbonation further shifts the spoilage flora to yeasts, which are more tolerant to CO_2 than the very sensitive molds and bacteria (Wareing and Davenport, 2005). Fruit juices, concentrates, or nectars can be consumed fresh, pasteurized, or unpasteurized. In the last years, high hydrostatic pressure is increasingly used for processing of fruit juices with the advantage of a better color and aroma stability, sustainment of the original fresh mouth feeling, and a significantly expanded shelf life (Oey et al., 2008). Yeast growth can result in the production of CO_2 and EtOH, cause flocculation, and affect the product stability due to pectinolytic enzyme activity (Fleet, 2007; Jay and Anderson, 2001). The genera most frequently associated with juice spoilage incidents include *Candida, Hanseniaspora, Pichia, Rhodotorula, Torulopsis, Saccharomyces, Zygosaccharomyces, Hansenula*, and *Trichosporon* (Arias et al., 2002; Renard et al., 2008). Spoilage can be often traced back to highly contaminated raw materials, and deficient hygiene (Juvonen et al., 2011). Additionally, the occurrence of highly resistant spoilage strains can facilitate spoilage. For example, although the decimal reduction time D_T is low, some strains can show high z-values as demonstrated by Tchango Tchango et al. (1997). Furthermore, spoilage strains are frequently resistant to preservatives as it was observed within the species *Zygosaccharomyces bailii* (Thomas and Davenport, 1985), *Zygosaccharomyces rouxii* (Lenovich et al., 1988), *Candida krusei*, *Saccharomyces bisporus*, *S. c.*, *Schizosaccharomyces pombe*, and *Pichia membranifaciens* (Lima Tribst et al., 2009). Examples for resistance mechanism include *S. c.*, which actively extrudes acid via a specific ATP binding cassette (ABC) transporter (Pdr12) in the plasma membrane, and *Z. bailii*, which mainly limits the initial diffusion of the acid into the cell and can furthermore catalyze the oxidative degradation of sorbate and benzoate (Piper et al., 2001).

1.2 MALDI-TOF MS

The above mentioned examples clearly show the importance of the choice of strains to obtain the desired fermentation product with regard to process consistency, efficiency and overall quality. Furthermore, the identification of spoilage strains reveals helpful hints for process and product safety, which makes the strain identification a valuable and necessary tool in food industry. In the following the respective limits of conventional techniques are summarized and compared to the potential of MALDI-TOF MS (matrix-assisted-laser-desorption/ionization time-of-flight mass spectrometry).

1.2.1 Identification of yeasts by conventional techniques

Several, mostly DNA based, methods for identification, characterization, or typing of closely related yeasts have been developed in the past, which allow a discrimination at strain level (Schuller et al., 2004). Some prominent examples, especially for *Saccharomyces* (*S.*) strains, include the karyotyping of chromosomes by pulsed-field gel electrophoresis (PFGE) (Demuyter et al., 2004; Esteve-Zarzoso et al., 2001; Naumov et al., 2000b; Naumov et al., 2001) and the restriction fragment length polymorphism (RFLP) analysis of nuclear (Wightman et al., 1996) or mitochondrial DNA (Beltran et al., 2002; Esteve-Zarzoso et al., 2004; Torija et al., 2003). Additionally, PCR-based techniques such as random amplified polymorphic DNA (RAPD) (Tornai-Lehoczki and Dlauchy, 2000), micro- and minisatellites (Legras et al., 2005; Marinangeli et al., 2004; Masneuf-Pomarède et al., 2007), delta-PCR (Legras and Karst, 2003), amplified fragment length polymorphism (AFLP) (Esteve-Zarzoso et al., 2010), and PCR-DHPLC have been developed (Hutzler et al., 2010), the letter especially to differentiate brewing yeast strains. However, in food industry, their application for routine analysis of yeasts is limited because of their frequently high costs, the requirement for highly skilled personnel, and the time consuming effort.

1.2.2 History of MALDI-TOF MS

The technical breakthrough in direct mass spectrometric analysis of high molecular weight compounds, especially of substances of biological and biomedical importance, succeeded in the 1980th with the development of “soft ionizing” MS-methods. Fast Atom Bombardment (FAB) is an ion source to determine thermally labile, non-volatile, polar compounds without pre-derivatization steps using the sample solved in low volatile solvent molecules, mostly glycerol (Hsu et al., 1983). Samples were subjected to a beam

of fast neutral atoms, mostly noble gases, so that non-fragmented macromolecules (Biemann and Martin, 1987) up to 20 kDa get desorbed and can be detected in the MS (Barber et al., 1981; Barber et al., 1987). In the routine analysis nowadays, this technique is replaced by ESI (Electrospray Ionization), APCI (Atmospheric Pressure Chemical Ionization), or MALDI, because of its much lower susceptibility, the poorer ion yield and the higher degree of pollution of the ion source.

In addition to the option to ionize the analyte indirectly via carriers as bombarding particles, samples were coated on a metal surface and exposed to sub-microsecond laser pulses to produce intact ionized molecules and a rather low level of fragmentation, so called Laser Desorption (LD)-MS (Posthumus et al., 1978; Stoll and Röllgen, 1979). In the 1980th Hillenkamp and co-workers began to expand the Laser Microprobe Mass Analyzer (LAMMA) application to detect small organic compounds, triggered by the observation of chemical noise in spectra, which apparently represented fragments of the organic matrix in the samples with moderate success (Feigl et al., 1983). In their retrospection of the development of MALDI, Hillenkamp and Karas (2000) described the main idea as a discovery by chance: signals of tryptophane and alanine could be detected in a mixture even though the laser fluence was adjusted to 266 nm, at which only tryptophane should have been desorbed and ionized, and - with the knowledge of previous experiments that the desorption threshold fluence at 266 nm wavelength of alanine was more than 10 times higher than tryptophane (Karas et al., 1985) - the consequence was that “Alanine came riding along piggyback with the Tryptophane” (Hillenkamp and Karas, 2000) and the idea of matrix assisted desorption was born (Karas et al., 1987). In 1988, at the International Mass Spectrometry Conference in Bordeaux, Karas and co-workers presented the first MALDI mass spectrum of a β-D-galactosidase in nicotinic acid (>100 kDa) (Karas et al., 1989). At this point, the awareness of the beneficial effect of matrix for desorption was already known and widely discussed, as mentioned above for FAB-MS (Barber et al., 1981; Hsu et al., 1983), plasma desorption-MS (Jonsson et al., 1986) or even LD-MS (Zakett et al., 1981). Hillenkamp and Karas postulated in 2000 that the concept of controllable deposition of the energy per volume of sample makes the difference and distinguishes the MALDI technique with its necessity of the use of an assisting matrix from other desorption techniques where this is essentially not possible. Nowadays, the use of an assisting matrix is applied exclusively in the MALDI technique, but until now the mechanistic background is not clarified in detail (cf. 1.2.3).

Within a few months after the publication of the new achievements, Beavis and Chait (1989) reported a substantially higher resolution. They demonstrated a capability of the instrument to detect separate signals from ions of similar masses defined as m/Δm (mass of an ion divided by full width of the signal at half maximum intensity (FWHM)) when a linear time-of-flight (TOF) analyzer instrument was coupled. The combination was ideally suited because of their theoretically unlimited mass range, high ion transmission, the small size, low costs, and the ability to obtain a complete mass spectrum in a few seconds (Opsal et al., 1985). The major limitation was still the resolution decreasing with increasing molecular sizes (Colby et al., 1994). Among the most common approaches to improve the spectra were the use of reflector analyzers, originally described by Mamyrin et al. (1973), deflecting the ions with an electric field and thereby doubling the ion flight path. Additionally, the delayed pulsed ion extraction (PIE) with high ion source bias voltage was employed to compensate initial velocity, temporal and spatial distribution of the produced ions with first experiments published by Wiley and McLaren (1955), improved and diversified over the years (Brown and Lennon, 1995; Colby et al., 1994; Juhasz et al., 1997; Opsal et al., 1985; Vestal and Juhasz, 1998; Vestal et al., 1995; Whittal and Li, 1995). Both solutions are still popular in modern constructions (cf. 1.2.3).

1.2.3 Technical overview of MALDI-TOF MS

Depending on the task, the three most commonly used matrix substances for proteins are sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid) (Beavis et al., 1992; Beavis and Bridson, 1993), CHCA or HCCA (α-cyano-4-hydroxycinnamic acid) (Beavis et al., 1992; Beavis et al., 1989), and DHB (2,5-dihydroxybenzoic acid) (Strupat et al., 1991). Juhasz et al. (1993) introduced HABA (2-(4-hydroxyphenylazo) benzoic acid) as a valuable matrix for peptides, proteins and glycoproteins up to 250 kDa. The determination of suitable matrices was, to some extent, by trial and error. Factors considered were basic chemical characteristics such as a fairly low molecular weight to guarantee easy vaporization. On the other hand, the vapor pressure has to be large enough to avoid the evaporation during sample preparation or in the spectrometer (Westman et al., 1994). Furthermore, a high absorption capacity of the laser irradiation in the individual wavelength range is essential (Ehring, 1992) and mostly achieved by several conjugated double bonds. Delocalized π-electron systems such as naphthalene compounds can act as an electron acceptor and can be used for analytes that also possess conjugated π-systems such as chlorophyll (Nazim Boutaghou and Cole, 2012). Generally, polar groups of matrix compounds enable water solubility and their acidic

character acts as a proton donator to support the ionization of analytes such as peptides or proteins (Beavis and Chait, 1989), but also alkaline matrices were reported (Fitzgerald et al., 1993).

A generally postulated main principle is the co-crystallization of analyte and matrix in a high molar excess as a prerequisite for a successful analysis (Horneffer et al., 1999). Both substances are spotted onto a MALDI metal plate and solvents vaporize so that the analyte incorporation into the matrix crystal lattice can take place. The plate is placed in a vacuum inside of the MS and a UV laser (N_2 laser in our case, 337 nm) is employed to ablate/desorb the sample. The ionization mechanism remained a debatable issue throughout all the years with various hypothesis including photoexitation and pooling (Ehring, 1992), proton transfer from excited states of matrix molecules (Gimon et al., 1992), or cluster models (Karas and Krüger, 2003), reviewed in recurring regularity by various experts (Dreisewerd, 2003; Knochenmuss, 2006; Knochenmuss, 2013; McEwen and Larsen, 2014; Zenobi and Knochenmuss, 1998). There seems to be a common agreement that the idea of one uniform MALDI mechanism may have to be given up (Glückmann et al., 2001). Karas et al. (2000), for instance, described matrix as a mediator for energy absorption followed by an explosive phase transition like a “gas jet” of the matrix, which entrains the analyte. Generally it is assumed that primarily the matrix absorbs the laser irradiation (Knochenmuss, 2006) so that the lattice structure is destroyed and matrix molecules can get desorbed. During this phase transition the embedded analyte molecules are ionized indirectly to prevent fragmentation (Westman et al., 1994; Zenobi and Knochenmuss, 1998).

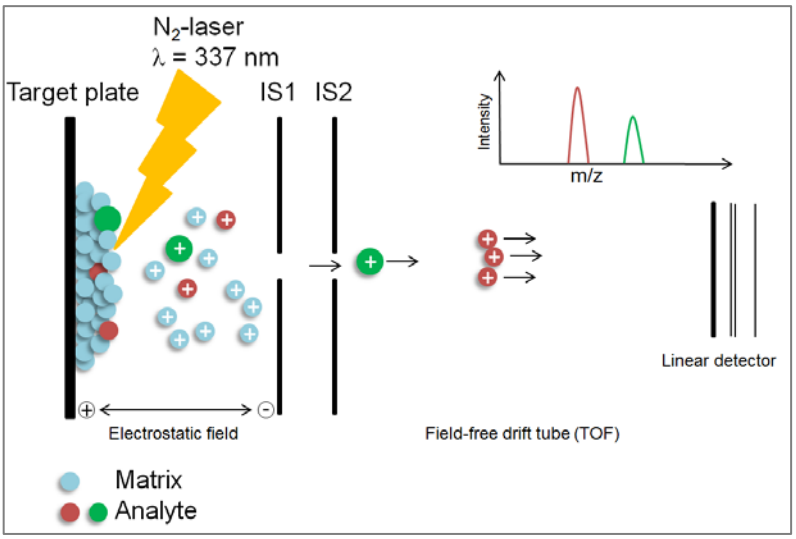

Fig. 1.4: Schematic illustration of the ionization process

MALDI-TOF mass spectrometers are preferably equipped with UV lasers such as nitrogen lasers (337 nm) and frequency-tripled and quadrupled Nd:YAG lasers (355 nm and 266 nm respectively). However, Horneffer et al. (1999) observed matrix adduct formation as a general problem at shorter wavelengths of 266 and 308 nm. Not as widespread but still in use are infrared lasers provoking a softer mode of ionization via absorption by vibration modes of matrix (Dreisewerd, 2003). This mode, a low degree of metastabile ion fragmentation (Dreisewerd et al., 2003), less low-mass interferences, and compatibility with other matrix-free laser desorption mass spectrometry methods (Berkenkamp et al., 1997) makes infrared lasers a valuable alternative for large molecules (> 500 kDa). Unfortunately, this approach has a low sensitivity, is expensive, relatively complicated, and does not work well with commercially available MALDI-TOF MS systems (Yoshihashi-Suzuki et al., 2008).

Using TOF MS the mass-to-charge ratio is determined via time measurement. Ions are accelerated by an electric field, where the potential energy E_p is related to the charge of an ion (z) and the potential difference of the field (U). The ion enters the field-free flight tube, where the potential energy is converted to kinetic energy E_{kin}. This is dependent on mass and velocity (v) of the ion, with the velocity as a quotient of the way in the tube (d) and the required time (t) for this flight.

Equation 2:

$$E_p = z * U \qquad E_{kin} = \frac{1}{2}mv^2 \qquad v = \frac{d}{t} \quad \gg \quad z * U = \frac{1}{2}m\left(\frac{d}{t}\right)^2 \quad \gg \quad t = \frac{d}{\sqrt{2U}}\sqrt{\frac{m}{z}}$$

Finally, the mass to charge ratio of the ion can be calculated from the measured time of flight. Measurement inaccuracies are likely to be caused during ion formation, which is basically due to differences in initial ion velocity, temporal, and spatial distribution. Consequently, ions of the same mass do not have the same kinetic energy after passing the acceleration field and, therefore, the time of flight is slightly different and influences the arrival at the detector. This results in broadening of peaks, which decreases resolution. To improve the resolution, an ion mirror, first proposed by Mamyrin et al. (1973), is used, which employs an retarding electric field to reverse the flight path of the entering ions. This leads to a diminished spread of flight times of identical ions, since ions with higher kinetic energy penetrate further into the electric field and follow a slightly longer path-length when their trajectories are reversed. Thus, such ions emerge from the reflector timely lagged behind lower-energy ions but catch up at the focal point, where the detector is located (Cornish and Cotter, 1993). Today, commercially available MALDI-TOF MS systems provide a resolving power of > 26,000 and mass accuracy ≤ 2 ppm in reflectron mode (e.g. Autoflex, Bruker Daltonik, Bremen).

Furthermore, a still popular possibility to improve resolution and sensitivity is a pulsed ion extraction (PIE). There, the acceleration voltage of the ion source is divided into two stages, i.e. the first one between the target plate and the first electrode IS1 (cf. Fig. 1.4) and the second one between IS1 and IS2. During the laser shot, target and IS1 are set to equal full potentials so that the generated ions are not yet exposed to any forces except the kinetic energy caused by the MALDI process. Therefore, ions can expand according to their different velocities and spread through the area between target and IS1. After a delay of a few nanoseconds, IS1 is pulsed down to the IS2 potential, so that the ions are accelerated towards IS1. However, due to their different positions the formerly slower ions obtain higher kinetic energy than those already closer to IS1, which results in focusing of identical ions (Brown and Lennon, 1995; Vestal et al., 1995).

Since the invention of MALDI-TOF MS, there are multiple discovered, solved, or simply discussed factors/problems that can influence its performance. Important influence factors include (i) sample-, solvent-, and pH-dependent incorporation of the analyte into the matrix (Cohen and Chait, 1996; Yao et al., 1998), which is considered to be the reason for sample suppression effects, as well as strong spot-to-spot ion-intensity fluctuations (Glückmann et al., 2001), (ii) the quality of the matrix-analyte layer to guarantee homogenous microcrystalls (Vorm et al., 1994), (iii) laser spot size, irradiance or angle of incidence (Ingendoh et al., 1994).

1.2.4 Potential and limits of yeast identification by MALDI-TOF MS

MALDI-TOF MS has been demonstrated to be a versatile method to analyze macromolecules from biological origin (Fenselau and Demirev, 2001), is already commonly employed in clinical diagnosis (Cassagne et al., 2013; Dhiman et al., 2011; Goyer et al., 2012; Marklein et al., 2009; Marvin et al., 2003; Stevenson et al., 2010; van Veen et al., 2010), and represents a promising tool for the rapid and reliable identification of yeasts along their mass fingerprints, mainly at species level but with a certain potential for strain identification (Marklein et al., 2009; Qian et al., 2008; van Veen et al., 2010).

For a MALDI-TOF MS analysis, cells are prepared and co-crystallized with matrix in a way that yields a sufficient number of medium-sized ions (2 – 20 kDa) in the mass spectra. The identification of microbiological samples by this method relies on the acquisition of mass fingerprints and the subsequent comparison of the obtained data with a database (DB). Dozens of sample preparation protocols for bacteria profiling are available in literature. Such protocols differ in particular steps of the procedure including different combinations of extraction solvents (Domin et al., 1999), physical methods to improve the cell wall disruption employing microbeads (Fagerquist et al., 2005), thermolysis (Vargha et al., 2006) or sonication (Easterling et al., 1998), and enzymatic treatments up to 20 h (Giebel et al., 2008). However, in comparison with bacteria, yeast cells exhibit a higher cell wall stability, which complicates protein extraction and solubilization. For bacterial cells it has previously been reported that the mass spectra can be affected by the growth medium, period of incubation, pH, and the cell preparation method used (Cohen and Chait, 1996; Valentine et al., 2005; Wunschel et al., 2005). Additional literature mainly focuses on the limitations of MALDI-TOF MS technique to identify microorganisms correctly (Marklein et al., 2009; Putignani et al., 2011; van Veen et al., 2010). However, effects of differences in the physiological state of yeast cells on the reliability of the identification are less explored.

A major problem is that the acquired spectra always exhibit complex features, since protein signals can be contaminated due to several chemical and/or physical processes during sample preparation or the measurement itself (Gras et al., 1999; Satten et al., 2004). Interfering signals, resulting in a baseline drift or background noise, can be produced by electronic disturbances and fragments of material with rapid fluctuations randomly varying over small mass ranges (Mantini et al., 2007). Other reasons for interfering signals include chemical noise, which is influenced by sample preparation and contaminations, the temperature in the flight tube, and software signal read errors (Tong et al., 2011). Consequently, a very sensitive and accurate peak detection method, able

to separate protein peaks correctly from the noise, is required (Diamandis, 2004). Numerous data preprocessing techniques were proposed including baseline correction, smoothing/denoising, data binning, peak alignment, peak detection, and sample normalization (Jeffries, 2005; Tong et al., 2011; Yang et al., 2009).

1.3 Aim

Since there exists an essential lack of knowledge regarding the reliable identification of yeasts in food and beverage microbiology using MALDI-TOF MS, the aim of this study was to solve following questions:

- Can beverage spoiling yeasts be identified at species level?
- Can sensitivity, reproducibility, and mass accuracy be enhanced by optimizing sample preparation protocols?
- Is the impact of different cultivation conditions and metabolic states observable in the mass spectrum?
- Are differences in mass spectra reproducible to use them as biomarkers for growth conditions?
- Is it possible to assign biomarker mass peaks to proteins?
- What is the possible level of differentiation achievable with MALDI-TOF MS?
- How large is the impact of the applied software on the level of differentiation?
- Is the achievable level of differentiation sufficient to select yeasts for specific applications in brewing or wine making?

For this purpose three species, commonly associated with spoilage incidents were used to determine the impact of sample preparation methods and measurement parameterization on the quality of generated mass spectra as evaluated by LIMPIC software (linear MALDI-TOF MS peak indication and classification) (Mantini et al., 2007).

The influence of varying culturing conditions on protein mass signatures was analyzed in order to evaluate their effect on the identification of spoilage yeasts. Data obtained this way was also used to search for biomarkers that enable a classification according to conditions of a specific yeast was grown at. To assign such biomarker peaks to possible proteins, a genome sequenced *S. cerevisiae* strain was used and proteins were isolated and identified via SDS-PAGE and LC-MS/MS sequencing.

Strain level differentiation experiments were conducted on the basis of a set of brewing strains for top- and bottom- fermented beer types and a set of wine strains with differences in their typical commercial application. In addition to commercially available software, bioinformatic tools were tested to optimize strain level identification results.

2 MATERIAL AND METHODS

2.1 Material

Tab. 2.1: Devices

Device	Model	Manufacturer
1-D gel electrophoresis chamber	Mini protean® Tetra System	BioRad, München, Germany
Analytical balance	SI-234	Denver Instrument, Bohemia, NY, USA
Balance	SBA 52 / SPO 61	Scaltec Instruments, Heiligenstadt, Germany
Centrifuge	Hermle Z383 K	Hermle Labortechnik, Wehningen, Germany
Fast Prep 24		MP Biomedicals, Eschwege, Germany
HPLC-RI	UltiMate 3000	Dionex, Idstein, Germany
Incubator	Memmert INB series	Memmert, Schwabach, Germany
Magnetic stirrer	AREC heating	VELP, Scientifica, Usmate, Italy
MALDI target	Anchor Chip96, polished steel	Bruker Daltonik, Bremen, Germany
MALDI-TOF MS	Microflex LT	Bruker Daltonik, Bremen, Germany
Microwave	Intellowave	LG Electronics, Ratingen, Germany
Mini Centrifuge	MCF-1350	LMS, Tokyo, Japan
NanoDrop spectrophotometer	NanoDrop 1000	Thermo Scientific, Wilmington, DE, USA
pH electrode	InLab 412, pH 0-14	Mettler Toledo, Gießen Germany
pH measuring device	Knick pH 761 Calimatic	Knick elektronische Geräte, Berlin, Germany
Photometer	LKB Biochrom 4060	Pharmacia Biotech, Uppsala, Sweden
Plate reader	Tecan SpectraFluor / Tecan Sunrise	TECAN Deutschland GmbH, Crailsheim, Germany
Pipettes	Pipetman (10 – 1000 µL)	Gilson International B.V., Limburg-Offheim, Germany
Pipettes	1 -10 µL	Eppendorf, Hamburg, Germany
Rotary evaporator		Heidolph Instruments, Schwabach, Germany

Device	Model	Manufacturer
Rotary Shaker		Witeg Labortechnik, Wertheim, Germany
Ultrasonic bath	Sonorex Super RK103H	Bandelin electronic, Berlin, Germany
Ultrasonic stick	Sonopuls	Bandelin electronic, Berlin, Germany
Vortex mixer	Vortex Genie 2	Scientific Instruments, NY, USA

Tab. 2.2: Consumables

Material	Type	Manufacturer
Bottle with screw cap	100-1000 mL	Schott Duran, Wertheim, Germany
Conical flasks	50-500 mL	Schott Duran, Wertheim, Germany
Cotton plugs	No. 18	Zefa, Harthausen, Germany
Cuvette	1 mL	Sarstedt, Nümbrecht, Germany
Glass beads	0.5 µm*1 mm	Scientific Industries, New York, USA
HPLC vials	1.5 mL	Techlab GmbH, Erkerode, Germany
HPLC crimp caps	PTFE	Techlab GmbH, Erkerode, Germany
Inoculation loops	1 µL	VWR, Darmstadt, Germany
Petri dishes	Ø 94 mm	Greiner Bio-One GmbH, Frickenhausen, Germany
Pipette tips	1 µL	Eppendorf AG, Hamburg, Germany
Pipette tips	20 µL-1 mL	Gilson Inc., Middleton, USA
Syringe filters, RC	15 mm, 0.2 µm	Phenomenex, Aschaffenburg, Germany
Sterile tubes	1.5-50 mL	Sarstedt, Nümbrecht, Germany

Tab. 2.3: Chemicals

Chemicals	Purity	Manufacturer
Acetic acid	100 %, glacial	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Acetone	> 99.5 %, for synthesis	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Acetonitrile	HPLC grade	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Acrylamide / Bisacrylamide	29:1, 30 % (w/v)	SERVA, Heidelberg, Germany
Agar-agar	European agar	Difco, BD science, Heidelberg, Germany
Ammonium persulfate	electrophoresis grade	SERVA, Heidelberg, Germany
Bacterial Test Standard	for mass spectrometry	Bruker Daltonik, Bremen
Bromophenol blue	for electrophoresis	Sigma-Aldrich, Steinheim, Germany
Casein peptone	for microbiology	Merck, Darmstadt, Germany
Dextrin		Sigma-Aldrich, Steinheim, Germany
DTT	high purity	Gerbu Biotechnik GmbH, Heidelberg, Germany
Ethanol absolute	> 99.8 %	VWR, Prolabo, Foutenay-sous-Bois, France
Formic acid		Fluka, Steinheim
Fructose monohydrate	HPLC grade	Merck, Darmstadt, Germany
Glucose monohydrate	for microbiology	Merck, Darmstadt, Germany
Glycerol	99.5 %	Gerbu Biotechnik GmbH, Heidelberg, Germany
HCl 37 %	p.a.	Merck, Darmstadt, Germany
H_2SO_4	95 – 97 %	Merck, Darmstadt, Germany
Iodoacetamide	≥ 99 %	AppliChem, Darmstadt, Germany
Isopropanol	≥ 99.9 %	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
K_2HPO_4	p.a.	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Malt extract	for microbiology	AppliChem, Darmstadt, Germany
Maltose	for microbiology	Merck, Darmstadt, Germany
Meat extract	for microbiology	Merck, Darmstadt, Germany

Chemicals	Purity	Manufacturer
Methanol	HPLC grade	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Molecular Weight Marker for peptides (range 2.5 – 17 kDa)	research grade	Sigma-Aldrich, Steinheim, Germany
NH_4Cl	≥ 99.5 %, p.a.	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Peptone from casein	for microbiology	Merck, Darmstadt, Germany
Peptone from soybeans	for microbiology	Oxoid, Hampshire, England
Perchloric acid	70 %	Merck, Darmstadt, Germany
Roti®-blue solution	Colloidal Coomassie staining of proteins (5x conc.)	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Na_2CO_3	≥ 99.8 %, p.a.	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
NaOH	≥ 99 %, p.a.	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Sucrose		Merck, Darmstadt, Germany
SDS	research grade	SERVA, Heidelberg, Germany
TEMED	p.a.	Merck, Darmstadt, Germany
Tricine	≥ 99 %	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Trifluoroacetic acid (TFA)		Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Tris(hydroxymethyl)-aminomethan-Base / HCl	99.5 %	Gerbu Biotechnik GmbH, Heidelberg, Germany
Yeast extract	for microbiology	Carl Roth GmbH & Co. KG, Karlsruhe, Germany
Yeast nitrogen based	for microbiology	Difco, BD science, Heidelberg, Germany
α-cyano-4-hydroxy-cinnamic acid		Bruker Daltonik, Bremen, Germany

Tab. 2.4: Software

Software	Manufacturer
Biotyper 3.0	Bruker Daltonik, Bremen, Germany
Biotyper Automation Control 2.0	Bruker Daltonik, Bremen, Germany
Chromeleon V6.6	Dionex, Idstein, Germany
ClinProTools 3.0	Bruker Daltonik, Bremen, Germany
FlexAnalysis 3.0	Bruker Daltonik, Bremen, Germany
FlexControl 3.3	Bruker Daltonik, Bremen, Germany

Tab. 2.5: Bacterial Test Standard (BTS) (Bruker Daltonik) composition (Bruker Daltonik, 2008)

Protein	Reference mass [Da]	± 300 ppm range [Da]
RL29 $[M+2H]^{2+}$	3637.8	3636.7 – 3638.8
RS32 $[M+H]^{+}$	5096.8	5095.3 – 5098.3
RS34 $[M+H]^{+}$	5381.4	5379.8 – 5383.0
RS33met $[M+H]^{+}$	6255.4	6253.5 – 6257.3
RL29 $[M+H]^{+}$	7274.5	7272.3 – 7276.7
RS19 $[M+H]^{+}$	10300.1	10297.0 – 10303.2
RNAse A $[M+H]^{+}$	13683.2	13679.1 – 13687.3
Myoglobin $[M+H]^{+}$	16952.3	16947.2 – 16957.4

Tab. 2.6: Gels, buffers and solutions for SDS-PAGE

Used for	Ingredients	Concentration	Amount
sample buffer	Tris Base	0.5 M, pH 6.8	1 mL
	Glycerol		920 µL
	SDS	10 %	1.75 mL
	DTT	1 %	
			Ad 8 mL
Tris-HCl / SDS-buffer	Tris-HCl		182 g
	SDS		1.5 g
			Ad 500 mL
cathode buffer	Tris Base	0.1 M	12.11 g
	Tricine	0.1 M	17.92 g
	SDS	0.1 %	1 g
			Ad 1 L
anode buffer	Tris Base	0.2 M	121.1 g
			Ad 0.5 L
			pH 8.9
separation gel	Acrylamide / Bisacrylamide	29:1	18.61 mL
	Tris HCl/SDS Buffer		11.1 mL
	Glycerol		4.46 g
	Ammonium persulfate	10 %, freshly prepared	200 µL
	Temed		20 µL
stacking gel	Acrylamide / Bisacrylamide	29:1	1.35 mL
	Tris HCl/SDS Buffer		2.58 mL
	dH_2O		6.42 mL
	Ammonium persulfate	10 %, freshly prepared	70 µL
	Temed		70 µL

Used for	Ingredients	Concentration	Amount
Coomassie staining solution	Rotiblue solution		200 mL
	Acetic acid		92 mL
	Methanol		354 mL
			Ad 1 L
fixing solution	Trichloroacetic acid (TCA)		120 mL
			Ad 1 L
destaining solution I	Methanol		500 mL
	Acetic acid		100 mL
			Ad 1 L
destaining solution II	Methanol		100 mL
	Acetic acid		100 mL
			Ad 1 L
AmBic buffer I	ammonium bicarbonate		0.1 M
	DTT		10 mM
AmBic buffer II	ammonium bicarbonate		0.1 M
	Iodoacetamide		55 mM

2.2 Strains

Tab. 2.7: Strains used for the optimization of experimental design and modeling parameters (software optimization) for the differentiation of beverage spoiling yeasts

Species	TMW	Reference
Saccharomyces cerevisiae var. *diastaticus*	3.236	(Jespersen and Jakobsen, 1996; van der Aa Kühle and Jespersen, 1998)
Wickerhamomyces anomalus	3.237	(Kurtzman, 2011; Timke et al., 2008)
Debaryomyces hansenii	3.238	(Deak and Beuchat, 1996; Suzuki et al., 2011)

Tab. 2.8: Used yeast species for DB enlargement and method validation provided by the BLQ

Species	Frequency of identification [abs.]	[%]
Candida boidinii	6	3.2
Candida colliculosa	1	0.5
Candida glabrata	2	1.1
Candida inconspicua	2	1.1
Candida intermedia	5	2.6
Candida krusei	9	4.7
Candida mesenterica	3	1.6
Candida palmioleophila	1	0.5
Candida parapsilosis	6	3.2
Species	**Frequency of identification [abs.]**	**[%]**
Candida pararugosa	1	0.5
Candida sake	1	0.5
Clavispora lusitaniae	6	3.2
Cryptococcus luteolus	1	0.5
Cryptococcus liquefaciens	1	0.5
Debaryomyces hansenii	2	1.1
Dekkera anomala	4	2.1
Pichia guilliermondii	4	2.1
Pichia manshurica	4	2.1
Pichia membranifaciens	23	12.1
Pichia norvegensis	4	2.1
Rhodotorula mucilaginosa	12	6.3
Saccharomyces cerevisiae	25	13.2
Saccharomyces pastorianus	17	8.9
Torulaspora delbrueckii	7	3.7
Trichosporon asahii	5	2.6
Wickerhamomyces anomalus	39	20.5
Yarrowia lipolytica	3	1.6
Zygosaccharomyces bailii	1	0.5

Tab. 2.9: Background information about *S. c.* brewing yeast strains provided by the BLQ

Species	TMW	Commercial strain code	Application	Origin
S. cerevisiae	3.250	TUM 68	Wheat beer	Weihenstephan
	3.251	TUM 127		Weihenstephan
	3.253	TUM 149		Munich
	3.255	TUM 175		Weihenstephan
	3.258	TUM 205		Kelheim
	3.343	TUM 505		Weihenstephan
	3.252	TUM 148	Altbier	Düsseldorf
	3.257	TUM 184		Düsseldorf
	3.259	TUM 308		Rhein-Main-area

	3.336	TUM 192		Düsseldorf
	3.337	TUM 338		Düsseldorf
	3.256	TUM 177	Kölsch	Krefeld
	3.332	TUM 998		Cologne
	3.260	TUM 210	Ale	Nutfield, England
	3.261	TUM 211		Nutfield, England
	3.262	TUM 213		Nutfield, England
	3.338	TUM 503		Odell, USA
	3.339	TUM 506		Montréal, Canada

Tab. 2.10: Background information about *S. p.* brewing yeast strains provided by the BLQ

Species	**TMW**	**Commercial strain code**	**Property**	**Application**	**Origin**
S. pastorianus	3.275	TUM 34/70	Flocculation	Pilsner beer, Lager, Export	Nuremberg
	3.276	TUM 34/78			Nuremberg
	3.277	TUM 59			Nuremberg
	3.278	TUM 69			Nuremberg
	3.279	TUM 120			Fürth
	3.280	TUM 128			Vienna
	3.281	TUM 168			Gießen
	3.282	TUM 8-I-4			
	3.283	TUM 8-J-4			
	3.284	TUM 8-J-5			
	3.285	TUM 66/70	Non-Flocculation	Pilsner beer, Lager, Export	Dortmund
	3.286	TUM 204			Munich
	3.350	TUM 53			
	3.351	TUM 92			
	3.352	TUM 106			
	3.353	TUM 139			
	3.354	TUM 145			
	3.355	TUM 162			
	3.356	TUM 167			
	3.357	TUM 170			
	3.358	TUM 182			
	3.359	TUM 183			

Tab. 2.11: Wine yeast strains provided by Lallemand Inc., Montréal, Canada

Strain code	Genetic information	Commercial strain code	Application	Reference method
Lalld 1	*S. c. x S. k*	Lalvin W46	cold white wine	*1, 2
Lalld 2	*S. c.*	Lalld 2	Chardonnay	*1, 2
Lalld 3	*S. c. x S. u.*	Lalvin DV10	Champagne	*3
Lalld 4	*S. c. x S. u.*	Levuline CHP	Champagne	*3
Lalld 5	*S. c.*	Lalvin W15	cold white wine	*1
Lalld 6	*S. c. x S. u.*	Lalld 6	Champagne	*3
Lalld 7	*S. c.*	Lalvin ICV D47	Chardonnay, white wine	*1, 2
Lalld 8	*S. c.*	Lalvin V1116	Chardonnay, white wine	*1, 2
Lalld 9	*S. c.*	Lalvin 71B	Beaujolais	*1
Lalld 10	*S. c. x S. u.*	Lalvin EC1118	Champagne	*3
Lalld 11	*S. boulardii*	Levucell SB20	Probiotics	*1
Strain code	**Genetic information**	**Commercial strain code**	**Application**	**Reference method**
Lalld 12	*S. b.*	Lalvin QA23	Champagne, Chardonnay, Sauvignon blanc	*3
Lalld 13	*S. c.*	Lalvin ICV OKAY	Chardonnay	*1, 2
Lalld 14	*S. c. x S. u.*	Affinity	Champagne	*3
Lalld 15	*S. c. x S. k.*	Lalvin W27	Cold white wine	*1, 2
Lalld 16	*S. c. x S. u.*	Lalvin CH14	Champagne	*3
Lalld 18	*S. c.*	Lalvin CY3079	Chardonnay	*1
Lalld 19	*S. c.*	Lalld19	red wine	*1, 2
Lalld 20	*S. c.*	Uvaferm HPS	red wine	*1, 2
Lalld 21	*S. c.*	Lalvin ICV D254	red wine	*1, 2
Lalld 23	*S. c.*	ICV K1M	Chardonnay, white wine	*1
Lalld 24	*S. c.*	Uvaferm WAM	Sauvignon blanc, Chardonnay, white wine	*1, 2
Lalld 25	*S. c. x S. u.*	Lalvin S6U	White, rosé wine	*1
Lalld 26	*S. c.*	Siha levactif 6	Distillery	*1
Lalld 27	*S. c.*	Zymaflore F15	Chardonnay, red wine	*1
Lalld 28	*S. c.*	Vitilevure C	Sauvignon blanc, starter for all types	*1
Lalld 29	*S. c. x S. u.*	Vitilevure Quartz	Champagne	*3
Lalld 31	*S. c. x S. k.*	Anchor VIN7	Cold white wine, Sauvignon blanc	*1
Lalld 32	*S. c. x. S. k.*	Levuline ALS	Cold white wine, Sauvignon blanc	*1
Lalld 33	*S. c. x S. b.*	CrossEvolution	Sauvignon blanc	*1
Lalld 34	*S. c. x S. b.*	Anchor VIN13	Sauvignon blanc	*1
Lalld 35	*S. c. x S. b.*	Vitilevure Elixir	Sauvignon blanc	*1
Lalld 36	*S. c. x S. b.*	Zymaflore X5	Sauvignon blanc	*1

S. c.: Saccharomyces cerevisiae; S. b.: Saccharomyces bayanus var. *bayanus; S. u.: Saccharomyces bayanus* var. *uvarum; S. k.: Saccharomyces kudriavzevii*

*1: RAPD-PCR, *2: RFLP, *3: Microsatellite (by Lallemand Inc.)

Tab. 2.12: Other strains used in this work

Species	TMW	Information / Reference
S. pastorianus type strains (non IGS2 rDNA hybrid strains) Isolation source: beer, Denmark (Fernández-Espinar et al., 2003)	3.287	CBS 1503, formerly *S. monacensis* (Nguyen and Gaillardin, 2005)
	3.288	CBS 1538*, neotypestrain, *S. pastorianus* (Rainieri et al., 2006)
	3.289	DSM 6580 NT*, *S. pastorianus*
S. cerevisiae S288C	3.308	Genome sequenced

* Discrepancies have been reported that type strains maintained in different culture collections are not identical, namely CBS 1538 and DSM 6580 NT (Joubert et al., 2000; Rainieri et al., 2003)

2.3 Media

To adjust the pH of the media, HCl and NaOH (1 mol/L or higher dilutions) were used. All media were sterilized for 20 min at 121 °C. Sugars were sterilized separately and added aseptically to the media after cooling to approx. 50 °C.

Tab. 2.13: YPG

Ingredients	Amount
Casein peptone	10 g/L
Yeast extract	5 g/L
Glucose	20 g/L
Agar-agar	15 g/L
	Ad 1 L
pH	6.5 ± 0.2

Tab. 2.14: Universal medium (YM)

Ingredients	Amount
Yeast extract	3 g/L
Malt extract	3 g/L
Peptone from soybeans	5 g/L
Glucose	10 g/L
Agar-agar	15 g/L
	Ad 1 L

Tab. 2.15: Malt extract (ME)

Ingredients		Amount
Malt extract		20 g/L
Peptone from Soybeans		2 g/L
Agar-agar		15 g/L
		Ad 1 L
	pH	5.6 ± 0.2

Tab. 2.16: Sabouraud-Glucose (SAB)

Ingredients		Amount
Meat peptone		5 g/L
Casein peptone		5 g/L
Glucose		40 g/L
Agar-agar		15 g/L
		Ad 1 L
	pH	5.6 ± 0.2

Tab. 2.17: Wort medium (W)

Ingredients		Amount
Malt extract		15.0 g/L
Maltose		12.75 g/L
Dextrin		2.75 g/L
K_2HPO_4		1.0 g/L
NH_4Cl		1.0 g/L
Peptone		0.75 g/L
Agar-agar		15 g/L
		Ad 1 L
	pH	4.8 ± 0.2

Tab. 2.18: Yeast nitrogen based medium (YNB)

Ingredients	Concentration	Amount
Yeast nitrogen based		6.7 g/L
		Ad 800 mL
sugar stock solution	250 g/L	
		Ad 1 L

YNB solved medium was membrane filtered to achieve sterility according to the instructions of the manufacturer. Sugar stock solutions were autoclaved as described above and cooled to room temperature. The desired sugar concentration was added aseptically to the YNB stock medium and filled with sterile deionized water ad 1 liter.

2.4 Methods

2.4.1 Optimization of the sample preparation method

To optimize the sample preparation, yeasts were grown from -80 °C, 80 % glycerol stocks on YPG and YM agar plates at 30 °C. Single colonies were used to inoculate 15 mL of both growth media followed by an aerobic incubation overnight at 30 °C in 50 mL flasks sealed with cotton plugs on a rotary shaker at 180 rpm. Subsequently, another 15 mL of the respective growth medium were inoculated with 1 % of the overnight culture and incubated as described above.

At least three biological replicates were used. All of such independent samples were spotted in triplicate onto the MALDI stainless polished steel target for technical replication. Different strategies for rapid cell wall lysis were employed as follows:

2.4.1.1 Direct transfer (D)

Cell material of a freshly grown single colony was smeared directly onto a spot as a thin film using a sterile toothpick.

2.4.1.2 On target extraction (FA)

The smeared spot (cf. 2.4.1.1) was overlaid with 1 µL 70 % formic acid (FA) and dried at room temperature.

2.4.1.3 Extraction (E)

900 µL liquid culture were transferred into a 1.5 mL tube and centrifuged at 11.800 × g for 2 min. The supernatant was discarded and the pellet was mixed thoroughly with 300 µL dH_2O. Afterwards, 900 µL absolute EtOH were added and mixed. The tube was centrifuged at 11.800 × g for 2 min, the supernatant was discarded and the pellet was air dried for minimum 30 min until the solvent evaporated completely. Dependent on the pellet size, 10 up to 50 µL 70 % FA, i.e. approx. 10 µL per 100 µg dried pellet, were added and mixed thoroughly until the extract was resuspended. This was followed by the addition of an equivalent volume of acetonitrile (ACN) and mixed again. The mixture was centrifuged at 11.800 × g for 2 min and subsequently 1 µL of the supernatant was spotted onto a MALDI target and dried at room temperature.

2.4.1.4 Extraction without centrifugation (EZ)

Cells were prepared as described above (cf. 2.4.1.3) except of the last centrifugation step, i.e. the suspension was spotted directly onto the MALDI target.

2.4.1.5 Glass beads (GB)

Cells were prepared as described above (cf. 2.4.1.3), but after discarding the growth medium, 50 mg 0.5 mm glass beads were added and mixed for 2 min.

2.4.1.6 Fast prep (FP)

Cells were prepared as described above (cf. 2.4.1.5), but instead of using a shaker for all mixing steps, a fast prep 24 was used with 4 m/s for 40 s.

2.4.1.7 Ultrasonication bath (UB)

Cells were prepared as described above (cf. 2.4.1.3), extended to an additional ultrasonication treatment (ultrasonic bath) for 5 min after mixing the cell suspension with absolute EtOH.

2.4.1.8 Ultrasonication stick (U)

Cells were prepared as described above (cf. 2.4.1.3). A sonication treatment after mixing the cell suspension with absolute EtOH was performed using a Sonopuls device to provide continuous application of acoustic waves for one cycle of 30 s increasing the power continuously from 0 to 90 % with the suspended cells being cooled in an ice bath.

After each of the described methods the target spots were covered with 1 µL matrix solution of HCCA, which was freshly prepared in 50 % ACN and 2.5 % trifluoroacetic acid (TFA) with a final concentration of 10 mg HCCA/mL.

MALDI-TOF MS analysis was performed on a Microflex LT spectrometer using FlexControl (Version 3.3) and Biotyper Automation Control (Version 2.0) to define measurement parameters and target layout. For the DB comparison Biotyper 3.0 was employed. The spectra were recorded in the linear positive ion mode with a nitrogen laser (337 nm) at a laser frequency of 60 Hz within a mass range from 2 to 20 kDa at a voltage of 20, 16.8, and 7 kV for the first ion extraction stage, the second ion extraction stage, and lense, respectively. Pulsed ion extraction of 200 ns was used and the laser power was modulated between 30 % and 65 %. For each spectrum, 240 single spectra,

acquired in 40-shot steps from different positions of the target spot, were summarized. An external calibration was performed at regular intervals by using Bacterial Test Standard (BTS).

2.4.1.9 Data analysis

To optimize the sample preparation each raw spectrum was converted to a text file by FlexAnalysis (Version 3.3) software listing intensities versus m/z data points spaced 0.25 Da from each other. Text files were imported in LIMPIC (ported to octave, see below). After a preprocessing step where the single mass spectra were smoothed based on a Kaiser filter (Kaiser and Schafer, 1980) and baseline subtracted, the noise level for each m/z value was used to calculate the signal-to-noise-ratio (SNR). Subsequently, peak picking was performed by finding all local maxima and eliminating those with intensities lower than a non-uniform threshold proportional to the noise level (Currie, 1999; Yasui et al., 2003). Since the mass spectra could be inaccurately aligned after the calibration procedure, a maximum tolerance distance equal to 600 ppm of the m/z value was accepted for the comparison (Fushiki et al., 2006; Wang et al., 2006). Finally, a classification of peaks based on the peak detection rate (pdr) was performed, which was expressed by the ratio between the number of spectra containing the considered peak and the total number of analyzed spectra (Mantini et al., 2007). This was done to compare the preparation procedures to each other with respect to their efficiency and receive the maximum possible number of reproducible peaks in the mass spectrum. ClinProTools (Version 2.2) (Ketterlinus et al., 2005) was used to generate classification models for biomarker detection and evaluation. This software was generally used with default settings except when using Support Vector Machine (SVM), where the peak number was changed to 200. In-house software, used for mass spectra clustering and spectra comparison to a DB, was based on octave (Hornik et al., 2003), a high-level language for numerical computations. It was run in parallel on an open sharedroot computer cluster (ATIX; http://opensharedroot.org) using a message passing interface (MPI) for job control (Gabriel et al., 2004). BASH (http://www.gnu.org/software/bash) scripts were used to generate software pipelines.

All protein sequence data were obtained from the UniProtKB database (http://www.uniprot.org/uniprot/?query=taxonomy:559292).

2.4.2 Effect of different culturing conditions

To investigate the influence of different culturing conditions, 30 spectra per variation were collected from 10 different biological replicates, cultured and measured on at least seven different days to take technical and biological variability into account. The extraction protocol described in 2.4.1.3 was used for all preparation procedures.

Yeasts were cultivated aerobically and anaerobically. This was achieved by filling 15 mL sterile tubes to the very top. Five different media were used: YPG, YM, malt extract (ME), Sabouraud-glucose (Sab), and wort (W) with a composition as detailed above (cf. 2.3). On overview of the influence was generated by a principal component analysis (PCA) created by the BT software.

Growth phase experiments were conducted three times analyzing three replicate samples per growth medium. For this purpose 500 mL conical flasks sealed with cotton plugs were filled with 150 mL of either YM or W broth and incubated aerobically. Cultures were inoculated to an optical density (OD_{590nm}) of 0.05 (approx. 1 %) and samples for MALDI-TOF MS analysis and the determination of biomass via OD_{590nm} measurements in 1 mL cuvettes were taken at various time points (2.5, 5, 7, 10, 12.5, 15, 18, 24, 36 and 48 h).

To determine the optimal cell concentration for MALDI-TOF MS analysis the cultures were grown in YPG and YM broth, diluted using the respective broth to an OD_{590nm} of 1.0, 0.75, 0.5, down to 0.25. Data analysis was performed as described in 2.4.1.9.

2.4.3 Database enlargement and validation

From different beverages including beer, shandy, icetea, juice, and lemonade and from varying habitats in the production process, e.g. from coolers, gaskets, filters, fillers, rinsing water, and the final end-product, 190 liquid blind-coded yeast samples were isolated, enriched and identified at the BLQ by sequencing the amplified D1/D2 domain of 26S rDNA. PCR was performed by adding the primer system NL1 (5`-gcatatcaataagcggaggaaaag-3`) and NL4 (5`-ggtccgtgtttcaagacgg-3`). The reaction mix contained 400 nM of each of both primers, 1.5 mM $MgCl_2$, 200 µM dNTPs, and 1.25 U Taq polymerase in a total volume of 25 µl per PCR tube. Amplification was carried out applying the following program: 95 °C for 5 min, 35 cycles of 95 °C for 30 s, 52 °C for 2 min, and 72 °C for 2 min followed by a final elongation step at 72 °C for 10 min. The sequence analyses were performed at GATC Biotech (Germany). Resulting sequences of different lengths were compared to appropriate databases. In parallel, samples were

streaked on YPG agar plates (cf. Tab. 2.13) and examined by MALDI-TOF MS using the optimized extraction method (cf. 3.1).

2.4.4 Brewing yeast typing

To differentiate at strain level, all strains were cultured on YPG medium for 18 h in conical flasks at ten different days to obtain biological replicates. All samples were harvested in the "fourth passage". This means that cryo-conserved stock cultures were streaked on a YPG agar plate, incubated at 30 °C for 72 h until single colonies were visible. One single colony from such plates was again streaked on a YPG agar plate. After 2 days of incubation at 30 °C one colony was picked to inoculate 20 mL YPG broth in a conical flask sealed with a cotton plug, which was incubated for 18 h at 30 °C on a rotary shaker. From this culture, a 1 % inoculation of another 20 mL YPG broth was performed and again incubated for 18 h at 30 °C on a rotary shaker. Finally, 900 µL of the culture were used to extract cell material according to the optimized sample preparation method (cf. 3.1). The resulting extractions were spotted as triplicates on a target for technical replication so that finally 30 spectra per strain were acquired.

The strains were provided by the BLQ Weihenstephan. The commercially used strain collection of this institute guaranteed culture purity and an accurate assignment of the strains provided by employing PCR-DHPLC (Hutzler et al., 2010) and ITS sequencing. The used TF strains are listed in Tab. 2.9, the bottom fermenters in Tab. 2.10.

2.4.4.1 Data analysis

Different approaches were used to analyze the generated spectra. One approach consisted of the Biotyper 3.0 software by Bruker Daltonik (Bremen) (BT), which was provided with the hardware equipment and created for routine usage of MALDI-TOF MS. According to the standard operation protocol of the manufacturer, the DB entries for the BT were generated by spotting the extracted strain material on 8 positions of the target and measuring three times per spot. Thus, 24 spectra were compiled (BT1*24). In parallel, one spot was prepared with bacterial test standard (BTS), which is a mixture of proteins in all mass regions to calibrate the system (cf. Tab. 2.5). From the 24 sample measurements, those spectra were deleted, which shifted > 500 ppm related to the mass-to-charge ratio, considering each section of thousands of the spectra. A minimum of 20 spectra per strain had to remain for the creation of a main spectra projection (MSP). For quality assurance, all raw spectra were processed by the "Biotyper Preprocessing Standard Method", including mass adjustment, smoothing, baseline subtraction,

normalization, peak picking with respect to peak intensity distribution and peak frequency (Bruker Daltonik, 2008).

Arbitrary triplicate measurements as test dataset were then reconciled against the DB entries. The identification was performed automatically by the software without any action required from the user. The software generates a list of peaks up to 100, in which the threshold for peak detection is a SNR of 10. The alignment is performed with a tolerance of 250 ppm in the mass-to-charge ratio. The resulting peak list was matched against the reference library using the integrated standard pattern matching algorithm with its default settings, expressing the results in a score range from 0 to 3 as a logarithmic expression of a score summary from 1 to 1000. The score increases for each peak matching between measured and library spectrum and decreases for each peak that does not match. Additionally, an overall match of the intensity pattern is added (Bruker Daltonik, 2008). Details of the algorithm are protected by IP rights of the company and not accessible to the public. The best matches were put on one level with the identification and were summarized to illustrate the results in hit rate percentages for comparability.

In comparison to the standard procedure of 24 measurements per MSP, a main spectrum was generated by BT software with triplicate measurements from four different approaches to allow for technical and biological variability by aberrations during cultivation, extraction, equipment performance, surrounding influences and many more (BT4*3). Again, the library entries were reconciled against the search entries.

Furthermore, BT software provides the opportunity to generate subtyping MSPs for highly related spectra by analyzing those peak matches with the lowest number of conformity within the MSP set getting the highest weighting. The weightings are multiplied with their corresponding frequency and intensity to avoid very small peaks getting a high weighting.

The first in-house approach was based on MASCAP implemented in octave (MAS), where the MSP measurements were also employed to setup a DB and the triplicate measurements were summarized to one spectrum. As performed by BT, the MSP measurements were replaced by 12 spectra per strain cultured and measured at four different days to account for technical and biological variability and were summarized to a DB entry (MAS4*3), with which an arbitrary test dataset could be compared.

The similarity calculation worked with a maximum mass tolerance of peaks of 600 ppm for matching and only peaks above a cut-off mass of 2500 Da were taken into account.

Additionally, peak intensities were scaled logarithmically and were normalized. Always the first hit, independently of its score value, was used as the best match. This comparison offers the possibility of a yes/no answer at strain level and provides additional information on achieved mismatches and on mass spectra that are very close to each other.

In a next step a similarity calculation was performed, which was graphically displayed as an MDS (Multidimensional Scaling). The aim was to place all objects of interest in an *N*-dimensional metric space so that the distances between the single points exhibit their similarity or non-similarity to each other. This delivers a graphical representation of similarities, which can be correlated to properties all strains of a specific group have in common.

Additionally, DAPC (discriminant analysis of principal components) analysis was applied using the *adegenet* package of R software (Jombart et al., 2014; Team, 2011), described more detailed in chapter 2.4.5.1. Spectra from top- and bottom-fermenting strains were analyzed separately, and only the test spectra were used to run the analysis, taking the inhomogeneity of the dataset into account. Two different approaches are possible: firstly, the analytic approach can be based on the assumption that there is no more information other than the observed values. Accordingly, the whole dataset was tested with the function “find.clusters”, which runs successive *K*-means with an increasing number of clusters after the transformation of data by PCA (Fraley and Raftery, 1998) to predict the best supported model and therefore the number of clusters finally computes the Bayesian Information Criterion (BIC) (Jombart et al., 2014; Jombart et al., 2010; Lee et al., 2009). As a rule of thumb BIC decreases until it reaches the optimal *k*-value*,* hence, the “elbow” of the graph marks the optimal number of clusters. With this determined number of clusters, a DAPC analysis can be performed and the appropriate loading plot represents those values above the threshold of 0.02. Finally, a hierarchical cluster analysis was calculated, where the groups originally detected by “find.clusters” were colored for more clarity. Furthermore, the cluster was colorized according to the known grouping of the beer types (e.g. wheat beer, ale) and was compared with the first one.

In a second approach, a known grouping can be defined before the analysis – in this study the group of different beer type strains – to categorize the dataset along its brewing properties. For this purpose, the spectra were randomly split in one test and one validation dataset. Afterwards, the function “xvalDapc” performed a stratified cross-validation, using a varying number of principal components while the number of discriminant functions was kept constant to reduce the plenty of data. Different numbers

of principal components (PC) were run through to find the best DAPC, which is than performed with the training dataset. The following prediction using the validation dataset gives an idea if the system is valid (Jombart et al., 2014).

2.4.5 Wine yeast typing

All strains were cultured as described above (cf. 2.4.4). The 33 wine yeast strains used were blind-coded, i.e. except arbitrary numbers (Lalld1 – 16, 18 – 21, 23 – 29, 31 – 36), no information about the strains was considered to avoid any influence on the obtained results (clustering of strains according to MALDI-TOF MS analysis). As the standard reference method for strain differentiation, delta-PCR analysis using delta12/delta21 primers according to Legras and Karst (2003) was employed, which was performed by Lallemand Inc. (Montréal, Canada).

2.4.5.1 Data analysis

The delta-PCR cluster analysis was done by the GelCompar II V 5.0 commercial software (Applied Maths, Austin, USA). First, the software created fingerprint types from each strain banding pattern, i.e. densitometric records seen as a profile of bands on the *.tiff files. Bands were detected and normalized to create a DB, which was then clustered using Dice's coefficient and Unweighted Pair Group Method using Arithmetic averages (UPGMA) to construct a dendrogram based on the similarity matrix obtained.

Two different approaches were conducted to analyze MALDI-TOF MS spectra. A DB comparison to prearrange the blind-coded strains in groups, which were then visualized by similarity calculations and MDS graphics. Additionally, a mathematically based approach was performed employing DAPC (discriminant analysis of principal components). The first approach was tested with 30 MALDI-TOF MS measurements per strain of the first 15 blind-coded strains (Lalld1 – 15) to create a training dataset. The set was split into a first DB group with 21 spectra to calculate a DB entry using the commercially available BT software (Sogawa et al., 2011) and a second group with nine spectra as a test dataset for comparison. In addition to this, 24-fold measurements of the strains from one extraction were used to calculate a DB entry according to the standard operation protocol (SOP) of the manufacturer. This DB was then reconciled against the same nine spectra of the test dataset mentioned above. False correlation indicated similarities between different strains so that it was possible to organize the 15 strains (Lalld1 – 15) in six groups. In the next step, MASCAP was used to summarize all spectra in one entry to be able to process the acquired spectra (Mantini et al., 2010).

Subsequently, spectra similarity computations combined with ANOSIM (analysis of similarities) was employed, a non-parametric statistical procedure, which assumes that the distances within groups are smaller than those between groups. R-values present a measure for the separation of groups, where a good separation is indicated by R = 1 and R = 0 indicates a situation where no separation occurs. R-values above 0.75 are commonly interpreted as well separated, R > 0.5 as separated, R < 0.25 as barely separated (Ramette, 2007). An abbreviated pair wise test (ten iterations) between each group (e.g. each strain (1 vs. 2-15, 2 vs. 3-15 and so on) or each of the prearranged six groups) was calculated with 54 different metrices implemented in a FATHOM Toolbox for MATLAB (Jones, 2012). Thereafter, each metrix was checked by an ANOSIM analysis to calculate R-values. Afterwards, the metrix with the highest R-value was used to run an additional enlarged ANOSIM analysis with 1000 iterations. Resulting groupings were used to establish a Voronoi diagram. Additionally, a boxplot graphic was generated to visualize the homogeneity of the groups, where each spectrum of a group was compared with each of the other spectra within the same group. A box corresponds to 50 % of all data and the median is indicated by a red line within the box. Extending vertical lines (whiskers) indicate variability outside the upper and lower quartiles, and outliers are plotted as individual points (cf. p. XLII, Fig. S4).

As a validation dataset, 18 strains, blind-coded as well (Lalld16 – 36), were measured in the same manner as described above and sorted according to the existing optimized concept using DB comparison followed by an MDS correlation to check whether the results are transferable.

Since 33 strains as triplicates mapped on one diagram yields to a confusing mixture of points (n=99), single spectra were summarized to one sum spectrum and calculated as an MDS to reduce the complexity of the diagram. Afterwards, strains that formed distinct clusters separated from other strains (thus could easily be distinguished from the others) were removed to facilitate a closer look on the remaining strains that appeared other regions of the MDS.

In parallel, manual grouping applying DAPC using the *adegenet* package for the R software (Team, 2011) was tested. Originally developed for the analysis of genetic data to harness the large amount of information generated, e.g. by next-generation sequencing technologies, this approach was implemented here to classify MALDI-TOF mass spectra. In contrast to usual Principal Component Analysis (PCA), DAPC optimizes the variance between groups neglecting variations within a group and seeking synthetic variables, the discriminant functions, which show differences between groups as best as

possible (Jombart, 2013). In practice, the raw data were transformed using PCA, followed by *k*-means algorithms with increasing values of *k* to identify the optimal number of clusters. Different solutions were compared using Bayesian Information Criterion (BIC) ideally resulting in a minimal BIC as the optimal number of clusters. In reality, the actual number retained is merely a question of personal taste, because the minima achieved are not always completely clear. Nevertheless, the user usually defines a number of clusters and the discriminant analysis is performed, which results in a barplot of eigenvalues, and, finally, a scatterplot can be obtained representing the individuals as dots and the groups as inertia ellipses.

Such a scatterplot was arranged in this study to visualize grouping of the strains tested. Strains, which appeared as a discriminant group in one region were excluded step by step until only two strains remained. To confirm the result strains were organized according to their prearranged grouping and the procedure was repeated.

2.4.6 Mass spectrometric protein characterization by LC-ESI MS/MS

S. c. S288C (genome sequenced) was cultivated aerobically and anaerobically and the extraction was conducted according to the optimized protocol (cf. 2.4.1.3). The appropriate time point for harvesting the cells was determined by HPLC analysis of the metabolites (cf. 2.4.6.2). The generated MALDI-TOF mass spectra were summarized and all peaks that appeared in at least 80 % of the spectra were taken into account to provide reproducibility. *In silico*, the masses were compared with the genome sequence obtained from the National Center of Biotechnology Information (NCBI) DB (http://www.ncbi.nlm.nih.gov/) considering chemical modifications such as methylation, hydroxylation or multiple charges.

In parallel, the extracts were analyzed using a GeLC-MS/MS approach to prove the accuracy of the theoretical calculation. Therefore, the extraction was conducted in the same manner as for MALDI-TOF MS measurements and was separated by a one dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Gels were stained to localize the proteins/peptides with the estimated size, which were then destained, isolated, and prepared for deNovo sequencing by LC-ESI MS/MS (cf. 2.4.6.1). Raw data were provided as *.mgf files (Perkins et al., 1999), which were processed to peptide sequences using PepNovo (Frank and Pevzner, 2005; Frank et al., 2005; Frank, 2009; Frank et al., 2006) and, in parallel, using an open sharedroot computer cluster (ATIX; http://opensharedroot.org/) employing an MPI (message passing interface) for job control (Gabriel et al., 2004). BASH scripts and programs

written in C language generated and submitted catenations of amino acid sequences to MS-BLAST for protein/peptide identification according to Behr et al. (2007). Finally, protein functions were mapped on an interactive metabolic pathway map by iPath explorer 2.

2.4.6.1 SDS-PAGE

One dimensional SDS-PAGE was employed for protein separation and size estimation. The selected protocol is optimized for low molecular weight proteins (Schägger and von Jagow, 1987) and was modified according to Schurr (TUM, personal communication). All individual operation steps were conducted as clean as possible, e.g. the material surfaces were treated with acetone to avoid human protein contaminations.

The separating gel (17 % polyacrylamide) was produced with ingredients listed in Tab. 2.6 using freshly prepared APS and TEMED to initiate cross-linking, which was completed after one hour. The poured gel was overlaid with isopropanol, which was discarded after polymerization. The separation gel was overlaid with the stacking gel (4 % polyacrylamide), where a comb was plunged to prepare sample cavities, and the gel was allowed to polymerize overnight.

Two cultures of the genome sequenced *S. c.* TMW 3.308 were inoculated aerobically and anaerobically, respectively, as previously described (cf. 2.4.1.3). Three tubes â 900 µL culture suspension for each condition were extracted as for MALDI-TOF MS analysis and combined. The solvent residues of the assembled protein extracts were removed by a rotary evaporator. The dried samples were resuspended in 200 µL SDS buffer, and 10 µL were placed on the freshly prepared gel. Additionally, protein standard and bromphenol blue as a running marker were placed on the gel, and anode and cathode buffer (composition listed in Tab. 2.6) were filled in the electrophoresis chamber. To focus the proteins, a constant electric current of 80 mA was used until the samples reached the separation gel. Subsequently, the current was set to 10-15 mA and the gel ran for about 10 h. Afterwards, the gel was fixed for 20 min in 12 % trichloroacetic acid (TCA), washed with deionized H_2O (dH_2O), and Coomassie-stained for 1 min in 200 mL of the staining solution gently heated in a microwave. This was followed by incubation at room temperature for 2 up to 10 h to achieve a satisfying staining result. After washing with dH_2O the gel was treated with the destaining solutions I and II (cf. Tab. 2.6) until the background was not colored anymore. The bands, which were located in a mass range identical to range covered by MALDI-TOF MS analysis (2-20 kDa) were marked, and the gel was destained completely. Proteins of around 5, 6, 8, 10, and 12 kDa, were excised

from the gel, reduced and alkylated according to the protocol recommended by the Zentrallabor für Proteinanalytik (ZfP) of the Ludwig-Maximilian-Universität, Munich, Germany,
(http://www.proteinanalytik.abi.med.uni-muenchen.de/service/sample_submission/index.html) including a final destaining step with ACN and a 45 min treatment at 56 °C with 0.1 M AmBic buffer I (cf. Tab. 2.6). Subsequently, two washing steps with ACN were followed by 30 min incubation in the presence of AmBic buffer II containing 55 mM iodoacetamide in the dark. Finally, the samples were washed several times with AmBic buffer and ACN, and deNovo sequencing by LC-ESI MS/MS was performed at the ZfP in Munich.

2.4.6.2 Determination of metabolic products by HPLC-RI

All measurements were conducted in duplicate. 18 h yeast precultures were used to inoculate aerated YNB medium with 20 g/L glucose. Aliquots of the inoculated broth were then transferred to 24 conical flasks and 24 screw cap tubes to incubate cultures aerobically and anaerobically, respectively, so that at each of 12 measurement time points, two flasks and two tubes with screw caps could be used to measure optical density and glucose and EtOH contents using HPLC measurement. The optical density (at 590 nm) was measured in 1 mL cuvettes in the linear range of the photometer used (0.05 to 0.5), i.e. diluted with YNB medium if necessary. For HPLC analysis, the supernatants of the samples prepared for MALDI-TOF MS analysis were treated with 50 µL perchloric acid/mL sample and incubated at 4 °C overnight for protein precipitation. Afterwards, samples were centrifuged (14,000 × g, 4 °C, 30 min) and the supernatant was membrane filtered (0.2 µm, RC membrane). EtOH and glucose contents were determined by ion-exclusion HPLC (for details see Tab. 2.19) and quantified externally against a six point calibration line in a range from 12.5 to 500 mmol EtOH/L and 3 to 120 mmol glucose/L, respectively.

Tab. 2.19: HPLC system and conditions

Autosampler	Gynkotek Gina 50
Pump	Dionex P680
Detection	RI (Gynkotek RI-71)
Column	Phenomenexx Rezex ROA Organic Acid H^+ HPLC ion exclusion (8 % crosslinked resin), 300 x 7.8 mm
Mobile phase	5 mN H_2SO_4, membrane filtered (0.2 µm) and degased with He
Flow	0.6 mL/min
Injection volume	10 µL
Oven temperature	85 °C
Software	Chromeleon V6.6

3 RESULTS

3.1 Optimization of the sample preparation method

The performance of the MALDI-TOF MS method and its potential to reliably identify yeast at their species levels relies on its capability to acquire reproducible spectra. This, however, strongly depends on the sample preparation method. This dependency was assessed using three test yeasts (*S. cerevisiae* var. *diastaticus* (TMW 3.236), *Wickerhamomyces (W.) anomalus* (TMW 3.237), and *Debaryomyces (D.) hansenii* (TMW 3.238)) cultured on YM and YPG medium. In Tab. 3.1, tested protocols for sample preparation are listed including the respective percentages of successfully acquired spectra.

Tab. 3.1: Number of successfully acquired spectra [%] depending on the used methods: (D) direct transfer, (FA) on-target extraction, (E) extraction, (EZ) extraction without centrifugation, (GB) glass beads, (FP) fast prep, (UB) ultrasonication bath, (U) ultrasonication stick. The three test strains TMW 3.236, 3.237, and 3.238 cultured on YM and YPG medium were used

	Method	**Number of acquired spectra [%]**							
		TMW 3.236		**TMW 3.237**		**TMW 3.238**		**Ø**	
		YM	YPG	YM	YPG	YM	YPG	**YM**	**YPG**
Agar	D	66.7	50	66.7	50	94.4	88.9	**75.9**	**63**
	FA	72.2	55.6	77.8	61.1	88.9	100	**79.6**	**72.2**
Broth	E	100	100	100	100	92.6	100	**97.5**	**100**
	EZ	0	0	0	0	0	44.4	**0**	**14.8**
	GB	100	94.4	94.4	61.1	100	66.7	**98.1**	**74.1**
	FP	100	94.4	100	100	72.2	44.4	**90.7**	**79.6**
	UB	94.4	100	5.6	66.7	83.3	44.4	**61.1**	**70.4**
	U	94.4	100	66.7	100	61.1	11.1	**74.1**	**70.4**

On average, the methods E, FP, and GB were the three most successful methods to acquire MALDI-TOF MS spectra regardless of the cultivation medium used. Subsequently, the quality of the resulting spectra was compared by summarizing the generated spectra and counting the number of peaks considering only peaks, which both have an intensity of > 1,000 arbitrary units (a.u.; effective range from 0 up to approx. 5×10^4 a. u.) and appear in at least 80 % of the single spectra obtained to allow for reproducibility (pdr=0.8) (cf. Fig. 3.1).

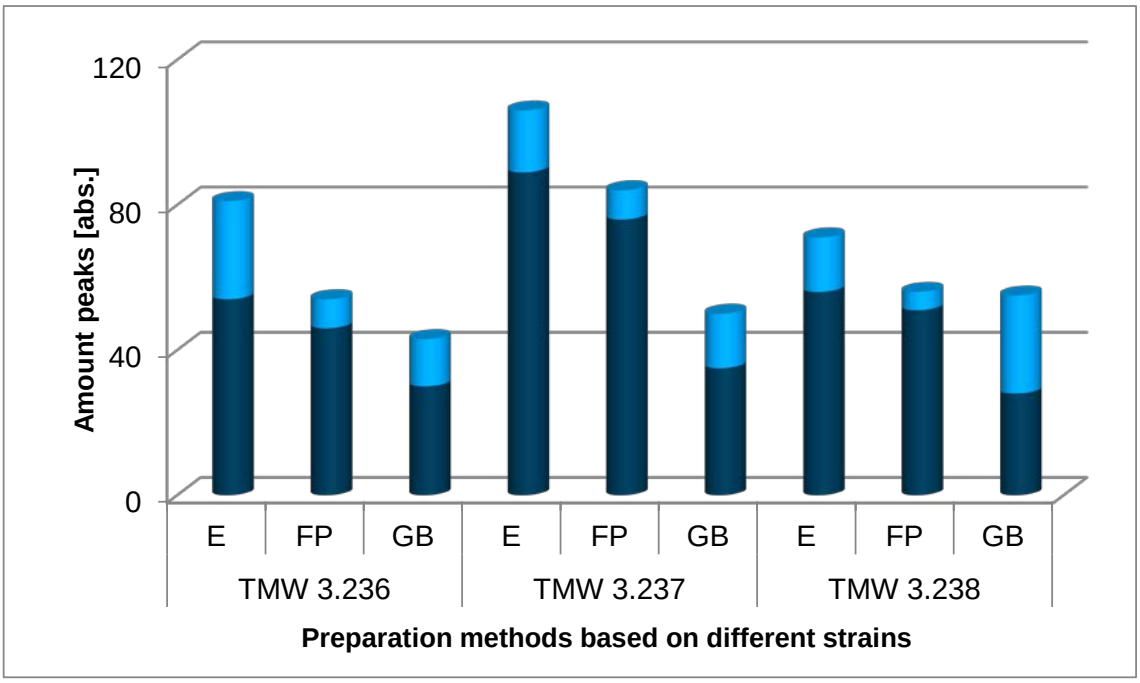

Fig. 3.1: Number of counted peaks with an intensity of > 1,000, subdivided into an intensity range from 1,000 to 5,000 (dark blue ●) and intensities of > 5,000 (light blue ●) depending on the three test strains grown in YM. The three most effective preparation methods (cf. Tab. 3.1) E, FP, and GB are compared

The absolute number of signals differed from species to species, however a tendency in methodology is evident: for yeasts grown in YM, E compared with FP led to a more than 20 % increase in total peak numbers, whereas GB always displayed the lowest peak detection. The larger peaks above an intensity of 5,000 a. u. were outnumbered in all cases. Similar results were obtained by using YPG broth (data not shown). Consequently, E was used as the standard preparation method for further experiments.

Additionally, the MALDI-TOF MS parameterization was optimized. The default setting for laser energy ranged from initially 30 % to maximum 40 % and increased in 2 % steps during the 240 single shots used to acquire one spectrum. However, significantly better results measuring for yeast cultures can be obtained using a range of initially 50 % to 65 % as pointed out in Fig. 3.2. The total number of peaks (> 1,000 a. u., dark blue) increased by 28 % for TMW 3.236, 35 % for TMW 3.237, and 52 % for TMW 3.238. Furthermore, a shift to higher peak intensities was observable, i.e. a decrease in peak numbers in the intensity range from > 1,000 to 5,000 a. u. (blue column) – except for TMW 3.238 – accompanied by an increase in the range above 5,000 a. u. (light and pale blue in Fig. 3.2).

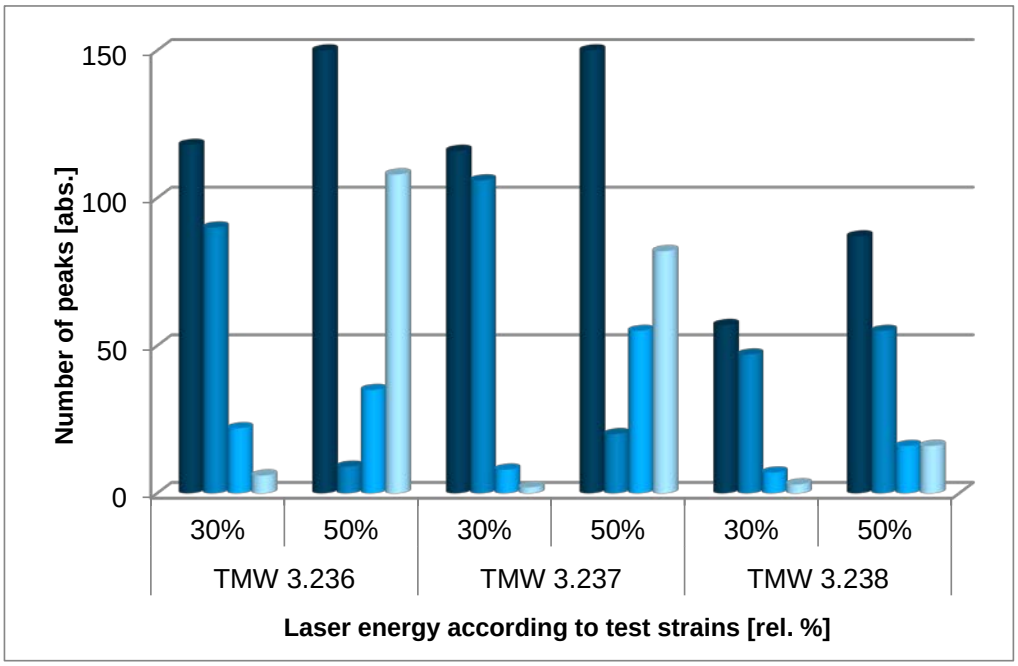

Fig. 3.2: Number of peaks of one sample cultivated in YM for 24 h at 30 °C, extracted according to method E, and measured at two different laser energy levels (30 % / 50 %). Different colors indicate different peak intensities of > 1,000 a. u. (dark blue ●), > 1,000 – 5,000 a. u. (blue ●), > 5,000 – 9,000 a. u. (light blue ●), > 9,000 a. u. (pale blue ●)

3.2 Effect of different culturing conditions

Environmental or physiologic parameters including the availability of oxygen, different growth phases, nutrients, and cell concentration were tested to check whether the species identification is affected by such parameters to evaluate the influence on the resulting mass spectra and to appraise whether different metabolic states could be differentiated by MALDI-TOF MS.

3.2.1 Oxygen

Fig. 3.3A exemplarily depicts the results for strain *S. cerevisiae* var. *diastaticus* TMW 3.236, where the sum spectrum of cultures incubated with air (black) is compared with that obtained after cultivation under anaerobic conditions (grey). For this comparison a mass range from 3-12 kDa is shown, since larger regions did not display any further signals. In Fig. 3.3B, the mass/charge (m/z) range displayed is further decreased to illustrate small differences in detail. The majority of peaks appear in both spectra ensuring reliable species identification, but the distribution of intensities obviously differs. Additionally, the presence of some larger peaks generated only under aerobic conditions

at a m/z ratio of 3081, 3704, 5297, 5777, and 6640 Da or anaerobic conditions (e.g. at m/z = 3494 Da) are clearly visible.

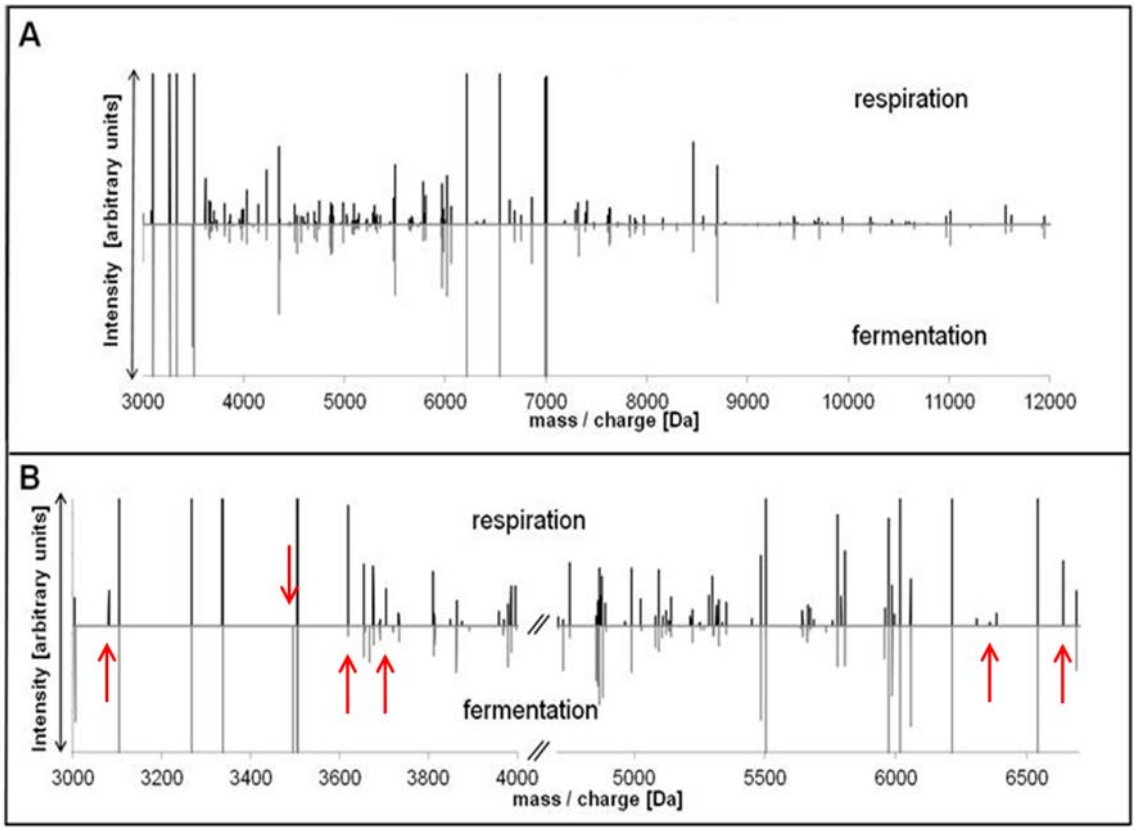

Fig. 3.3: Sum spectra of *S. cerevisiae* var. *diastaticus* (TMW 3.236) grown in Wort for 12 h at 30 °C. Upper black peaks indicate cultivation under aerobic conditions. Lower grey peaks indicate cultivation without air. (A) shows the entire mass range from 3 to 12 kDa used for the evaluation. (B) depicts a limited m/z range to point out the differences more clearly.

Although, differences appear to be species and nutrient-dependent, single mass spectra (n=360) compared with DB entries (sum spectra created from 30 spectra per condition (aerobic and anaerobic in five media (YPG, YM, W, Sab, ME), respectively (n=900)) can be precisely differentiated (99.6 %) along the small, oxygen availability-dependent differences (biomarkers) in the peptide mass fingerprint.

3.2.2 Nutrients

The result of a principal component analysis (PCA) derived from MALDI-TOF MS spectra of strain TMW 3.237 grown in five different media under aerobic condition is illustrated in Fig. 3.4. The 30 single spectra per growth medium displayed here (marked by five colors) show a clear grouping according to the used cultivation broth.

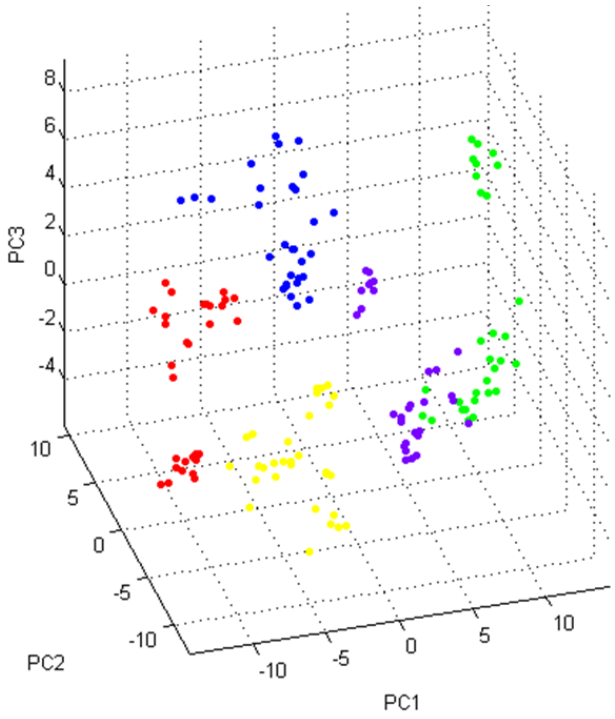

Fig. 3.4: Principal component analysis (PCA) created by BT software of *Wickerhamomyces anomalus* (TMW 3.237) grown in YPG (●), YM (●), W (●), Sab (●), ME (●).

Tab. 3.2 summarizes the hit rates of TMW 3.236, TMW 3.237, and TMW 3.238 grown aerobically in different media depending on different algorithms created by ClinProTools, where 21 spectra were validated with 9 single spectra per cultivation condition (nutrient broth used). The DB comparison was done with in-house software. For this purpose, 21 spectra per condition were summarized to create a DB entry and 9 spectra per condition were reconciled against this DB.

On average, the Support Vector Machine (SVM) algorithm led to best results (a match of 96 %) for strain TMW 3.236, which was followed by the Genetic Algorithm (GA) and the Supervised Neural Network (SNN) both with 76 %, whereas the Quick Classifier (QC) algorithm yielded significantly lower rates of 56 %. Highest hit rates for strain TMW 3.237 were obtained using the algorithms SNN and QC (82 %). Similar to the results obtained for strain TMW 3.236, SVM (Ø 80 %) was the best choice for analyzing TMW 3.238 spectra. On average, the use of SVM was suitable to obtain the highest hit rates for all three strains from different species (Ø 81 %).

Focusing on the hit rates depending on the media used, it is striking that – independent of the algorithm used – yeasts cultured in YPG could be retraced to the used broth to 100 % in most cases, except when the SNN algorithm was used to analyze TMW 3.236 spectra (44 % match). Second highest hit rates were found when spectra from strains

cultivated in W were analyzed. The best matches (%) of all other broths are randomly distributed among strain/algorithm/cultivation medium combinations.

In contrast to the four algorithms created by ClinProTools, the DB comparison using in-house software showed highest match rates with data obtained from both YM and YPG cultures (100 %) regardless of the strain used. The overall performance using the in-house software (87 % average for all strains and media) as well as the reliability (lowest standard deviation of hit rates (28 %)) were superior to the performance and reliability of the other algorithms tested.

Tab. 3.2: Each algorithm was tested creating a test data set from 21 spectra per nutrient broth generated from TMW 3.236 / 3.237 / 3.238 cultures and using 9 additional spectra to validate the class. All classes were calculated with four algorithms (by ClinProTools 2.2) and the average hit rates of the five different media were compared. DB comparison was done using in-house software with 21 spectra to create a DB entry and 9 spectra for the DB comparison

Class	Algorithms [matched %]				
	Supervised learning (modeling)				Database
	Genetic Algorithm (GA)	**Support Vector Machine** (SVM)	**Supervised Neural Network** (SNN)	**Quick Classifier** (QC)	**In-house software**
ME	100/33/11	100/44/0	100/11/0	33/78/0	100/89/0
SAB	100/100/11	100/0/100	89/100/89	22/67/67	44/100/100
W	78/100/100	78/89/100	56/100/100	100/100/78	78/100/100
YM	0/33/44	100/100/100	89/100/100	22/67/0	100/100/100
YPG	100/100/100	100/100/100	44/100/100	100/100/100	100/100/100
Ø	**76 / 73 / 53**	**96 / 67 / 80**	**76 / 82 / 78**	**56 / 82 / 49**	**84 / 98 / 80**

3.2.3 Growth phase

Fig. 3.5 exemplarily displays the cluster analysis (in-house software) visualizing similarities of MALDI-TOF MS spectra obtained after periodical measurements of TMW 3.236 cultures grown in W broth over 48 h. The maximum distance level representing complete diversity in the spectra obtained is 1.0. Two main clusters were identified, which differ clearly with a distance level of up to 0.45. The first cluster contains samples grown for 2.5 to 12.5 h, whereas the second one comprises samples measured after 15 up to 48 h. Both can be divided into two sub-clusters resulting in four growth

phase-dependent groups: 2.5 - 7 h, 10 - 12.5 h, 15 - 24 h, and 36 - 48 h. It can be assumed that below a distance level of 0.1, a differentiation of spectra is not possible anymore. The clusters changed slightly when YM broth was used, where sub-clusters of 2.5 h, 5 - 10 h, 12.5 - 18 h, and 24 - 48 h were identified (cf. p. XL, Fig. S2).

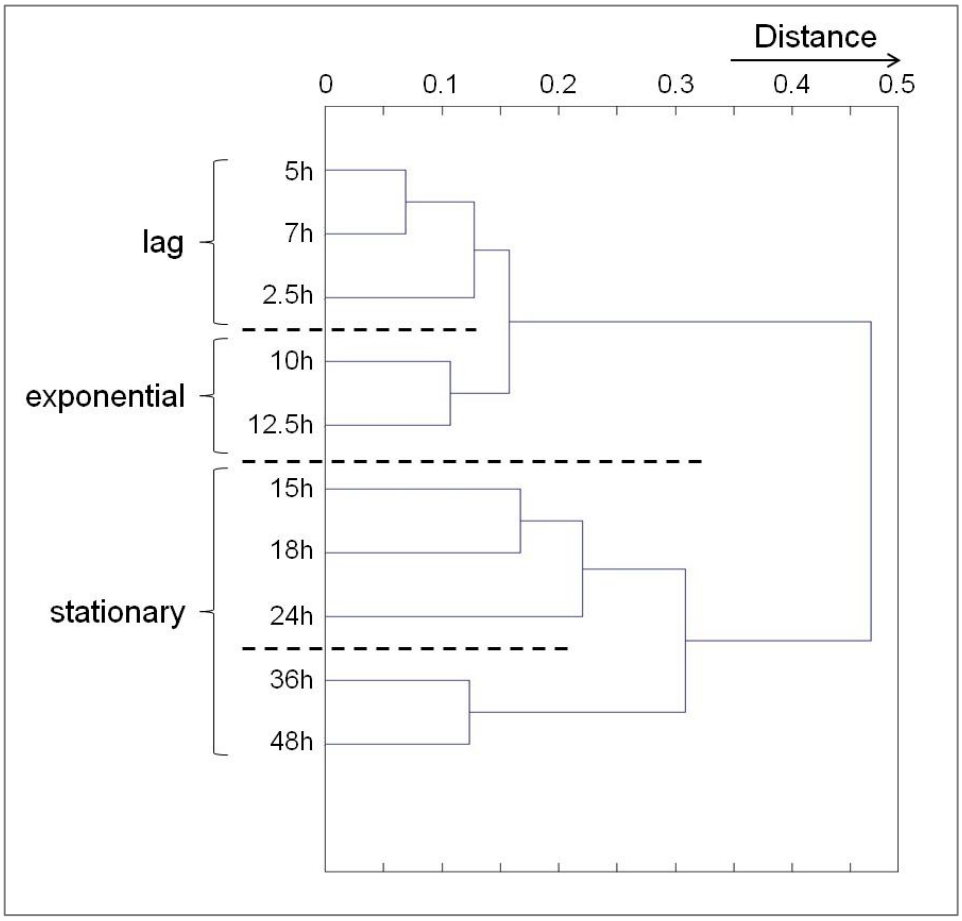

Fig. 3.5: Dendrogram of strain *S. cerevisiae* var. *diastaticus* TMW 3.236 sum spectra at different growth phases in Wort broth. Maximum distance level is 1.0 (analysis by in-house software)

Two main clusters with only one splitting in two sub-clusters were observed for strain TMW 3.237 grown in in W (2.5 – 7 h, 10 – 15 h / 18 – 48 h) (for details see p. XLI, Fig. S3-A) and YM (2.5 - 7 h, 10 - 24 h / 36 - 48 h) as well as TMW 3.238 in W (15, 18, and 2.5 h merged stepwise to 5 – 12.5 h / 24 – 48 h) (for details see p. XLI, Fig. S3-B) and YM (5 - 12.5 h / 15 – 24 h, 36 – 48 h).

3.2.4 Cell concentration

Cell suspensions were diluted to final optical densities (OD_{590nm}) of 0.25, 0.5, 0.75, and 1.0 in YPG and YM medium to evaluate the influence of the cell concentration on the quality of mass spectra. It was possible to generate spectra down to a dilution degree of

0.75. At OD 0.5, it was increasingly difficult to obtain spectra, especially for strain TMW 3.237 grown in YM, where this was possible in only 56 % of the measurements. At OD 0.25, the number of spectra obtained from measurements was limited for strains TMW 3.236 and TMW 3.237, i.e. for TMW 3.236 in YM limited to 89 % and in YPG to 78 %, for TMW 3.237 in YM to 44 % and in YPG to 67 %. In contrast, the generation of spectra from strain TMW 3.238 was never restricted. For further analysis, the results using very low cell concentrations (OD 0.25) were not considered due to deficiencies in their comparability. The generated spectra of samples adjusted to ODs of 0.5, 0.75, and 1.0 were summarized and the number of peaks that appear in 60 % (pdr = 0.6) of the spectra was counted (cf. Fig. 3.6). Slight variations are detectable in all cases. However, a uniform tendency was observed.

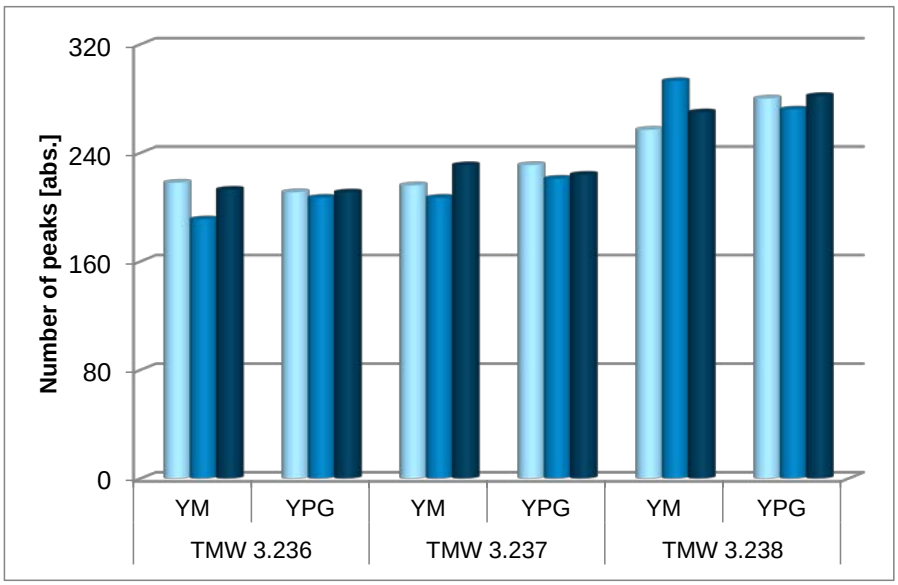

Fig. 3.6: Spectra were summarized by LIMPIC to count the number of peaks (pdr=0.6) at OD_{590nm} 0.5 (pale blue ●), 0.75 (blue ●), and 1.0 (dark blue ●)

3.3 Database enlargement and validation

A total of 190 yeast isolates from the beverage and brewing industry was identified by 26S rDNA sequence analysis (data provided by the BLQ institute, Weihenstephan) and by MALDI-TOF MS. An overview of all encountered species is given in Tab. 2.8. The overall distribution is given in Fig. 3.7, where green segments symbolize the correct identifications (MALDI-TOF MS results match the 26S rDNA sequencing results) and red colored segments indicate mismatches. In total, 172 isolates (90.6 %) yielded concordant results at species level for both methods applied. Five isolates (2.6 %) were identified at species level by MALDI-TOF MS but could only be identified at genus level using 26S rDNA sequencing, which was identical to that identified by MALDI-TOF MS.

However, since no reference results were available the correctness of the MALDI-TOF MS identifications could not be confirmed.

Rhodotorula (R.) mucilaginosa was incorrectly identified twice (1.1 % of all isolates) as *W. anomalus* by MALDI-TOF MS. For isolates belonging to either one of these two species, it was observed that even when the identification by MALDI-TOF MS (best match) was correct, the second best match often represented the respective other species with only marginal differences in the log-score.

Another isolate (0.5 % of all isolates), i.e. *Candida (C.) mesenterica* as identified by 26S sequencing, was misidentified as *C. inconspicua* by MALDI-TOF MS. Four more isolates (2.1 %) yielded discrepant results at genus level. *R. mucilaginosa* (26S sequencing) was identified as *D. hansenii* (MALDI-TOF MS), *Dekkera (Dk.) anomala* and *W. anomalus* were identified as *Pichia (P.) membranifaciens*, and *C. krusei* was identified as *P. membranifaciens*.

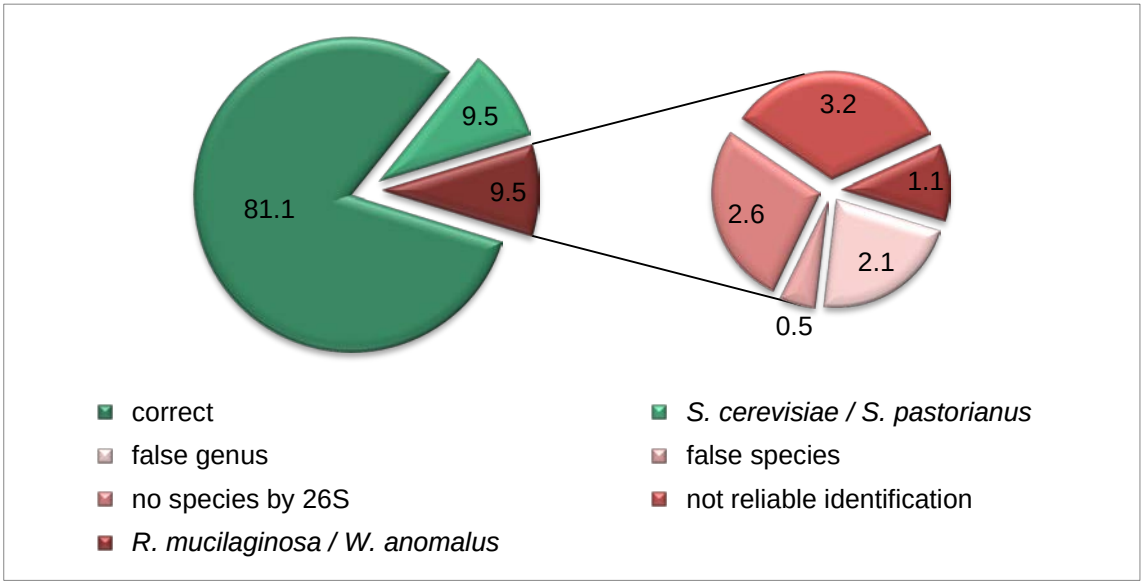

Fig. 3.7: Distribution of identifications. In the left pie chart the distribution of all samples is displayed. Correct identifications (MALDI-TOF MS results match the 26S rDNA sequencing results) are indicated by green color. The right chart splits the incorrect identifications (red color) according to their different reasons

Finally, six isolates (3.2 %) belonging to *Cryptococcus (Cr.) luteolus*, *Kwoniella (K.) mangroviensis, C. railenensis* (2), and *W. anomalus* (2) could not be identified reliably by MALDI-TOF MS, not even when applying a optimized protein extraction procedure.

3.4 Brewing yeast typing

3.4.1 Reference methods

All used brewing strains were kindly provided by the strain collection of the BLQ Weihenstephan, where 26S rDNA sequencing, real-time PCR, and PCR-DHPLC (Hutzler et al., 2010) were employed to guarantee reliable strain assignment and purity of the cultures, which is described in detail in chapter 2.4.3.

3.4.2 Differentiation of top- and bottom-fermenting yeasts

3.4.2.1 Approach using MALDI-TOF MS Biotyper software 3.0

A summary of the first approaches using the BT software for the identification of yeast isolates by MALDI-TOF MS is displayed in Fig. 3.8. The dark green colored boxes (diagonal) show the hit rate [%] by which a strain exactly matched with its previously created DB entry. Numbers in other boxes indicate hit rates [%] by which the BT software assorted a specific strain to the DB entry of another strain. Box areas with a specific blue color indicate the industrial application or specific properties of the yeast strains used, i.e. include strains that are used for the same beer type such as wheat beer or ale for top-fermenting (TF) strains or equal flocculation behavior for bottom-fermenting (BF) strains. These groups are referred to as “biotypes” in the following text. The differentiation of top- and bottom-fermenting strains worked for 98.6 % of the tested yeast isolates. At strain level, the correct identification rate for top fermenters was 47.3 % on average. For bottom fermenters, the average hit rate was only 22 %, although there were fewer BF strains tested. Within the top fermenters, 94 % of the wheat beer strains matched with other wheat beer strains, and all the outliers were Altbier strains. However, 70 % of Altbier strains matched Altbier strains, and mismatching strains mostly fell into the class of Kölsch strains (almost 27 %). Kölsch strains matched in 65 % of all cases with Kölsch strains, and the others were allotted to Altbier strains. The group of ale strains matched themselves in 93.3 % of all cases, while only a few were identified as Altbier and wheat beer strains.

The BF strains tested (TMW 3.275 through 3.284; all *S. pastorianus*) were mostly identified correctly as bottom fermenters (96.3 %). However, only an average of 22 % was identified correctly at strain level. Conspicuously, the DB entry of TMW 3.284 had no single match and even TMW 3.280 only matched in 10 % with itself. Highest hit rate

was obtained by TMW 3.277 with 70 %. Those spectra that were wrongly assigned as TF strains matched in approximately 82 % of the cases with wheat beer library entries.

%			top fermenting strains																	bottom fermenting strains									
	Biotype	TMW	50	51	53	55	58	52	57	59	36	37	56	32	60	61	62	38	39	75	76	77	78	79	80	81	82	83	84
top fermenting strains	Wheat beer	3.250		10		83	7																						
		3.251	3	33		50		7			7																		
		3.253		23		67		3			7																		
		3.255	7		13	63	17																						
		3.258			3	70	20	3			3																		
	Alt beer	3.252		3				17		3	10	10		57															
		3.257	3			7			53	3	13			20															
		3.259							3	53	13			30															
		3.336				3			40	27	10			20															
		3.337								30		63		7															
	Kölsch	3.256							23	20				57															
		3.332								3	7	17		73															
	Ale	3.260													97		3												
		3.261										3			10	87													
		3.262															100												
		3.338						3				17						43	37										
		3.339					3			7									90										
bottom fermenting strains	high flocculation	3.275				3														3		23	10			10	47	3	
		3.276																				23	3				43	30	
		3.277				3														10		70		7			3	7	
		3.278									3											10	20	17		23	3	23	
		3.279				7	3													3		23	17	10		3	7	27	
		3.280			3															3	20	17		3	10	43			
		3.281	3			7															3	27	3	10		27	10	10	
		3.282						3														7	20	3		3	43	20	
		3.283																			7		17	7		7	27	37	
		3.284																		13	3		7	17		10	43	7	

Fig. 3.8: Similarity calculation by BT software 3.0 – test spectra reconciled against DB entries of 27 strains composed of TF strains (blue) and BF high flocculating strains (light green) calculated according to the SOP of the manufacturer with 24 spectra per strain from one extraction procedure (BT1*24). The resulting hit rate [%] is displayed: green boxes show the exact strain match, whereas the color grounded parts indicate different biotypes

An enlargement of the dataset with spoilage *S. c.* var. *diastaticus* (violet), BF type strains (dark green), and BF strains, which exhibit low flocculation properties in comparison to those already measured (middle green) was performed to include a set of BF strains with apparently different properties assuming that this would result in more diverse MALDI-TOF MS spectra and facilitate the assignment of various isolates (results summarized in Fig. 3.9).

Despite the fact that the library entries were enlarged, the TF strains matched very similar except the ale yeast TMW 3.339, which correlated with DB entries of the strains TMW 3.273 and 3.274 in 23 % of the cases. Additionally, strain TMW 3.338, where the strain level match generally increased, was identified as TMW 3.339 (20 %).

For the BF strains the hits were more widespread but matched with other bottom fermenters in nearly all cases, except TMW 3.279, which complied in 10 % with wheat beer strains. Only TMW 3.289 and 3.358 could be sorted to their corresponding DB entry in 100 %, followed by TMW 3.288 in 97 % of all cases. Some library entries were more frequently matched than others, i.e. most prominently, almost 26 % of all BF search

entries matched with the DB entry for strain TMW 3.358, followed by TMW 3.359 with almost 12 %. In contrast to this, the DB entries of TMW 3.284 and 3.352 did not get one single match, and TMW 3.280 was matched only once. Strain level and beer type level hit rates are summarized in Tab. 3.3.

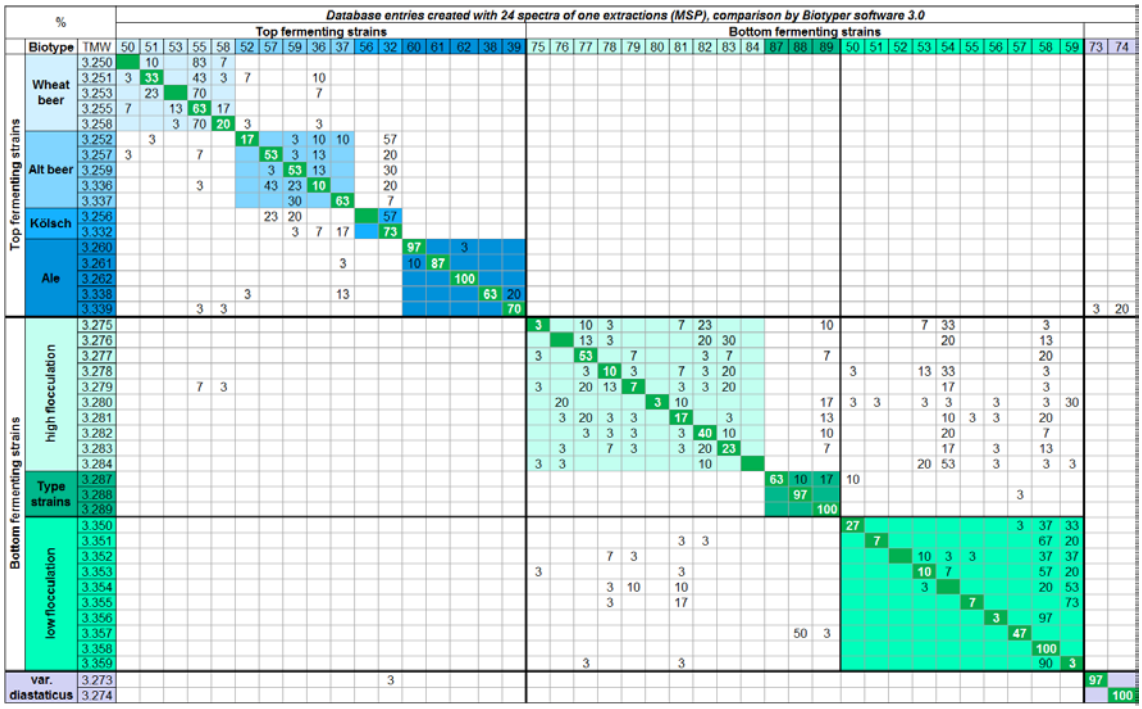

Fig. 3.9: Similarity calculation by BT software 3.0 – test spectra reconciled against DB entries of 42 strains, calculated according to the SOP of the manufacturer with 1*24 spectra per strain from one extraction procedure (BT1*24) of TF strains (blue), BF strains with high (light green) and low flocculating properties (middle green), BF type strains (dark green) and two spoilage strains (violet). The resulting hit rates [%] are displayed: diagonal green boxes show the exact strain match, whereas the color grounded parts indicate different biotypes

Fig. 3.10 exemplarily shows the spectra of two of three replicate measurements of strain TMW 3.275 samples out of the same extraction and measured timely in immediate succession (A), and comparing independent extractions from two different days (B). Although, the peak intensities in (A) are different, the overall relation is constant, whereas the pink spectrum in picture (B) lacks the two marked peaks completely. The differences between both approaches are clearly visible and might have influenced identification results. Hence, in contrast to the recommended procedure of the manufacturer, library entries by the BT software were created on the basis of 4 triplicates measured independently (n=12) to better account for biological and technical differences (BT4*3).

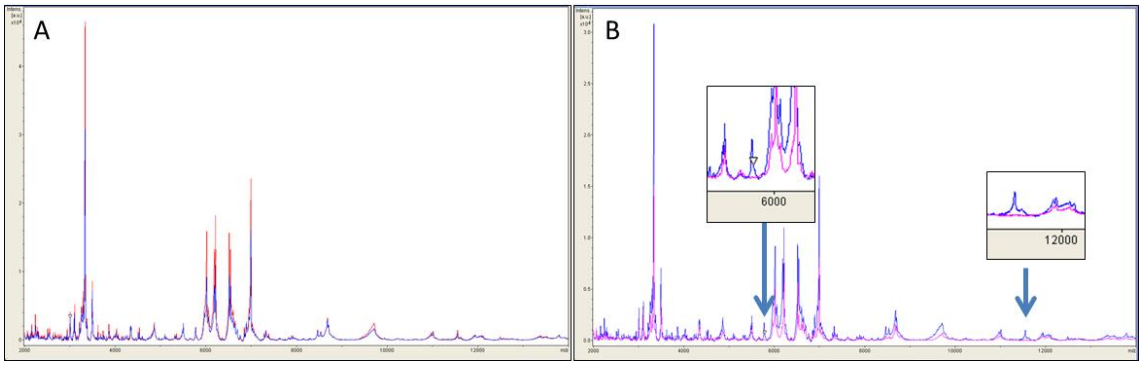

Fig. 3.10: TMW 3.275 measured on one day out of the same extraction (A) compared with separately extracted measurements on two different days (B), which differ in the two marked peaks

The MALDI-TOF MS results, i.e. hit rates for strain identification, using the adjusted procedure for making DB entries (BT4*3) are depicted in Fig. 3.11 and summarized in Tab. 3.3. Almost all results changed, which makes it difficult to describe any general trends. When the average over all samples is considered, a slightly higher correct strain identification match score of 45.3 % was observed using the BT4*3 method as compared with the 38.6 % achieved using the BT1*24 (manufacturer SOP). For wheat beer strains TMW 3.250 and 3.253, the strain level identification hit rate increased from zero to 43 % and 63 %, respectively, whereas the rates decreased from 33 % to 7 %, from 63 % to 0 %, and from 20 % to 10 % for the strains TMW 3.251, 3.255, and 3.258, respectively.

At the biotype level, 94 % of measured wheat beer strain preparations were correctly assorted to wheat beer strains using the BT1*24 method, whereas only 90 % obtained beer type level concordance with DB entries using the BT4*3 protocol. All Altbier strains matched with the Kölsch strain, TMW 3.332, applying the BT1*24 method. However, using the BT4*3 method it also resulted in a Kölsch strain (TMW 3.256) matching with 4 out of the 5 Altbier strains tested. For the Altbier strains, TMW 3.252, 3.336, and 3.337, much higher matches with themselves were achieved using BT4*3. However, the remaining two Altbier strains, TMW 3.257 and 3.259, did not match once in contrast to a correct strain identification in 53 % of all measurements applying the standard BT1*24 method. For both Kölsch beer strains better results were obtained using the BT4*3 method, i.e. the hit rate for strain TMW 3.256 was improved from 0 % (BT1*24) to 70 % (BT4*3), and that for TMW 3.332 was increased from 73 % to 90 % using the BT4*3 protocol.

Finally, hit rates for the ale beer strains, TMW 3.260, 3.261, and 3.262, were decreased from 97 % to 70 %, from 87 % to 13 %, and from 100 % to 77 %, respectively, when the BT4*3 protocol was used. Rates for TMW 3.338 and 3.339 improved from 63 % to 80 %

and from 70 % to 93 %, respectively. Correct beer type assortment for ale beer strains worsened from originally 90 % down to 68.7 %. Most of the wrong assortments were assigned to Altbier and Kölsch strains.

The type strain matches differed slightly depending on the protocol used to create DB entries: the strain level match obtained for strains TMW 3.287 and 3.288 were not affected, but the wrongly assigned results changed somewhat using the BT4*3 method. In contrast, the 100 % successful strain identification for TMW 3.289 using the BT1*24 protocol was lost and this strain was wrongly identified as TMW 3.358 in 13 % of the measurements when the BT4*3 method was used. On the other hand, none of the high and low flocculating strains matched with TMW 3.289 anymore.

The BF strains were still not very organized when the BT4*3 method was used instead of the BT1*24 protocol. The “dominant” DB entry using the BT1*24 protocol, TMW 3.358, which matched with nearly all strains except TMW 3.355 and 3.357, lost nearly all of the wrongly assigned hits of other low flocculating strains, unfortunately including matches with itself when using the BT4*3 method. The DB entry for TMW 3.359, which was matched by the strains TMW 3.280 and 3.350 - 3.355 in 11.9 % of the cases (BT1*24), was now hit with only 5.8 % (BT4*3). BF strains, where a proper strain level identification was possible and/or DB entries were rarely hit by other strains include strain TMW 3.280, which matched only with itself (53 %), TMW 3.357, which had 100 % strain level identification and its DB entry was only hit once by another strain, and TMW 3.281, which also had only one incorrect match when another strain was tested (TMW 3.279).

Identification results for the spoilage strains TMW 3.273 and 3.274 were almost identical regardless of the protocol used (BT1*24 or BT4*3).

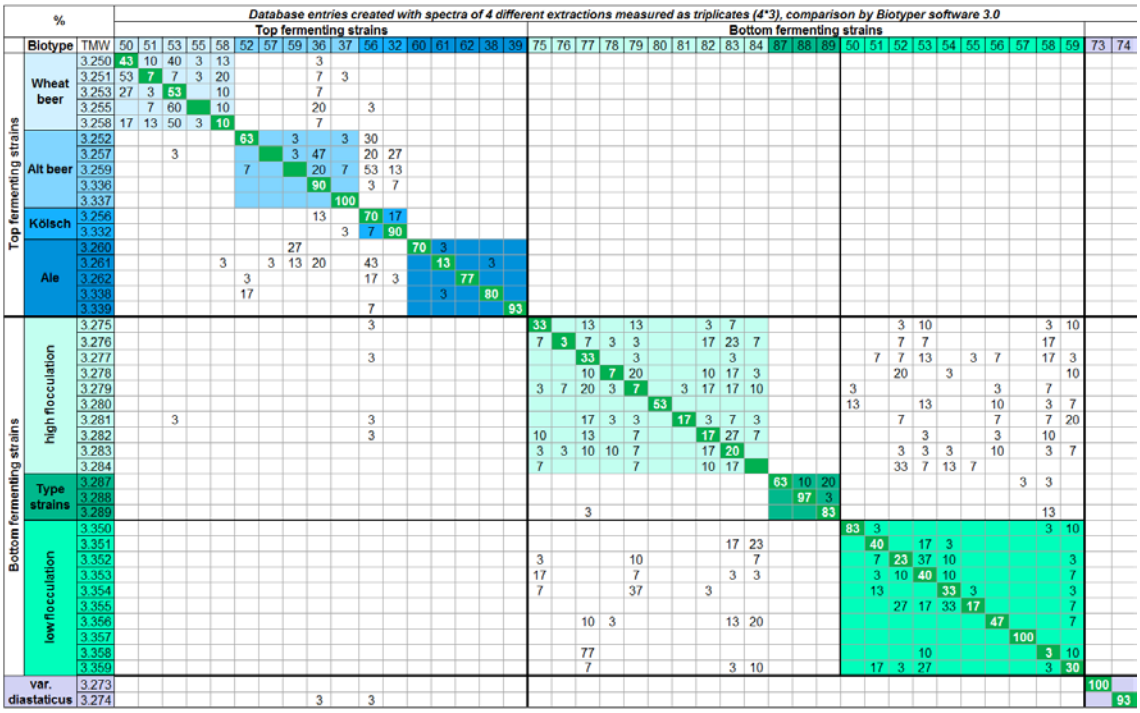

Fig. 3.11: Similarity calculation by BT software 3.0 – test spectra reconciled against DB entries of 42 strains, calculated with 4 independently measured triplicates (n=12) (BT4*3) of TF strains (blue), BF strains with high (light green) and low flocculating properties (middle green), BF type strains (dark green) and two spoilage strains (violet). The resulting hit rates [%] are displayed: diagonal green boxes show the exact strain match, whereas the color grounded parts indicate different biotypes

Additionally, the subtyping procedure of the BT software was performed with the BT1*24 as well as BT4*3 DB entries. This resulted in a further shift in strain identification. For example, 14 matches changed for TMW 3.275 when the BT1*24 method was used to create DB entries. All of these changes lead to other wrong identifications. When the BT4*3 method was used, 9 identifications changed for this strain, two of which now assorted to the correct DB entry (originally identified as TMW 3.279 and 3.353) and another two previously correctly identified matches were wrongly assigned afterwards (TMW 3.277, 3.283). Similar results were obtained with the other strains.

To visualize the resemblance of the BF strains, a PCA dendrogram, displayed in Fig. 3.12, was generated by BT software based on triplicate measurements of 20 strains including TMW 3.275 – 3.284 (high flocculants) and TMW 3.350 - 3.359 (low flocculants) (n=60). All strains were cultured simultaneously under identical conditions, extracted in parallel and measured on the same target. These very comparable conditions eliminated the influences of the technical variability as good as possible to enhance the comparability of the results. Three main clusters appeared, which are marked in green, blue, and red in Fig. 3.12. In the green cluster, high flocculating strains dominate (except TMW 3.350), the blue cluster comprises mostly low flocculating strains (apart from TMW

3.277), and the red cluster contains various strains. For only nine of the twenty strains spectra from the triplicate measurements were similar enough to cluster together in one arm, which impressively indicates that the measurement-related differences are frequently larger than strain-specific differences even when special care is taken to minimize the technical variability during cultivation, sample preparation and handling.

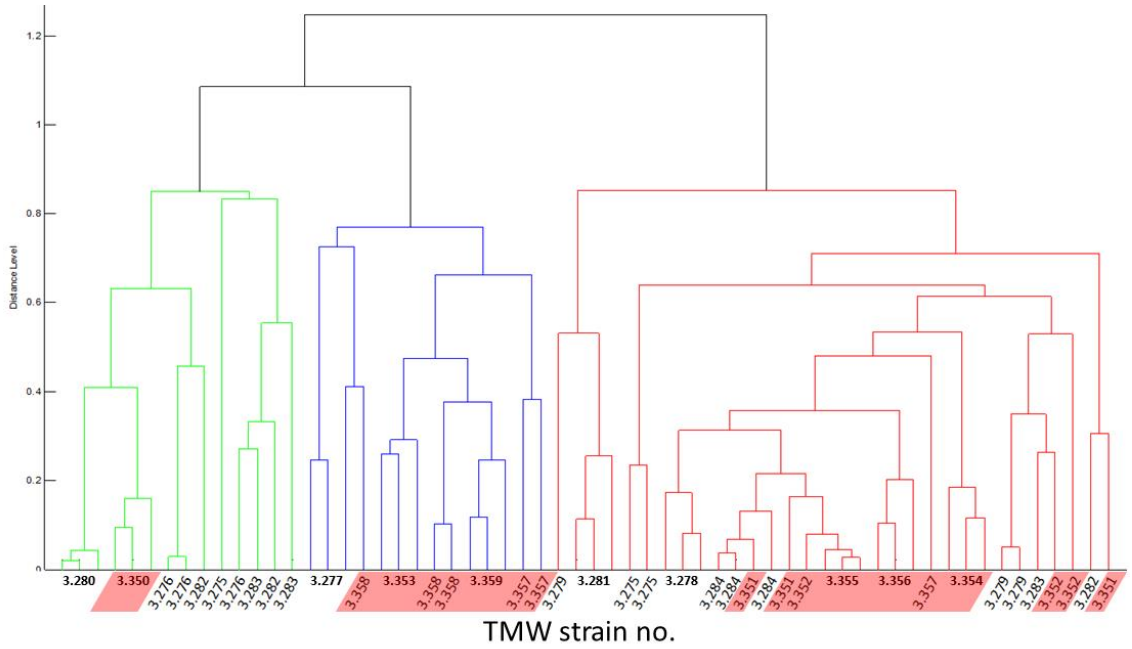

Fig. 3.12: PCA dendrogram based on MATLAB integrated in MALDI BT software 3.0. High and low (highlighted with a red background) flocculating strains, measured as triplicates on one target.

Fig. 3.13 displays a color-coded heatmap of single spectra from the 20 BF strains all measured simultaneously created by BT software. The scale runs from dark red illustrating 100 % conformity, which appears when the spectrum is compared with itself, down to dark blue indicating no concordance of two spectra. In comparison with the dendrogram in Fig. 3.12, the heat map is beneficial to display the relationship between spectra from all strains beyond a yes/no answer. Unfortunately, it was only possible to create the heatmap with single spectra so it is kind of a snapshot and might therefore create an appearance not coincide with the results of the other tools, visible e.g. within the results of TMW 3.359. In the dendrogram in Fig. 3.12 the nearest neighbor is TMW 3.358, but in the heatmap (Fig. 3.13) it is displayed that both strains are completely different (dark blue). However, TMW 3.357 stood out with a high number of light blue fields, which is in accordance with the similarity calculation in Fig. 3.11, where TMW 3.357 matched only with itself.

Very high similarities were observed within TMW 3.352 and 3.354 with a nearly 100 % color conformity, closely followed by TMW 3.281, 3.278, and 3.279.

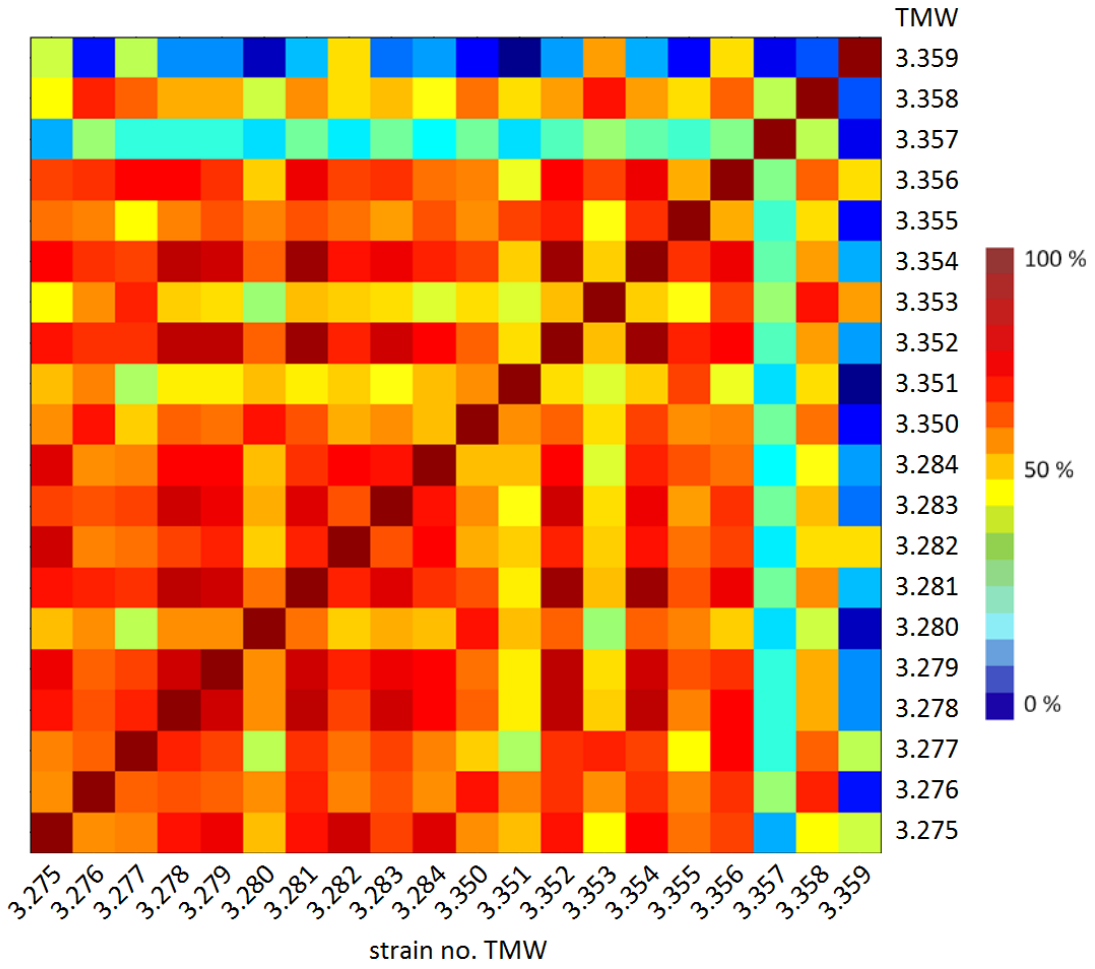

Fig. 3.13: Heatmap of single spectra of BF strains TMW 3.275 – 3.284 (high flocculants) and TMW 3.350 – 3.359 (low flocculants) created by BT. The color code from dark red (match for 100 %) to dark blue (0 % conformity) gives an impression of the relationship between all strains

3.4.2.2 Approach using in-house software

In parallel to the analysis using the BT software, an in-house software application based on MASCAP implemented in octave was set up and run in parallel on an open shared-root computer cluster (ATIX; http://opensharedroot.org) (MAS).

The results of a similarity calculation [%], where the MSP (main spectrum) measurement was used to generate a DB entry and this entry was then compared with single test spectra, are shown in Fig. 3.14 and summarized in Tab. 3.3. Again, the green marked

boxes indicate the correctly identified samples at strain level, whereas the colored areas demonstrate different beer types according to their application in the production of different beer flavor types such as wheat beer or ale. From the 1260 single spectra, 99 % could be assorted to the correct species. At strain level, 34.1 % were identified correctly. A couple of strains were conspicuous, since they did not even get one single match, i.e. TMW 3.278, 3.280, 3.281, 3.350, 3.356, 3.357, and 3.358. On the other hand, TMW 3.354 got over 16 % of all matches – this equals to 33.7 % of all BF strains – followed by TMW 3.353 with 5.6 % (11.8 % of BF), and TMW 3.258 with 5.4 % (45.3 % of all wheat beer strain spectra).

At beer type level, a high number of wheat, Altbier, and ale beer strains were correctly assorted to their respective beer type, i.e. 99.3 %, 92.0 %, and 97.3 %, respectively. Kölsch beer was an exception within the TF group, since the strains matched with Altbier strains much better than with the other Kölsch strain, which is reflected by only 15.0 % beer type concordance and 8.3 % correct strain level identification obtained for the two Kölsch strains. On average, the approach using the in-house software achieved correctly identified strains in 44.7 % of the spectra and 86.7 % correctly assigned beer type levels for top fermenters.

For BF strains, identification at strain level was successful in 18.0 % of the cases, where 12.0 % were assigned correctly in the group of high flocculants and 24.0 % in the group of low flocculants. For this property, overall 64.7 % correct assignment was obtained, 43.3 % for high and 86.0 % for low flocculants. The type strains were clearly separated as one group with a strain level identification hit of 53.3 %. Finally, 76.7 % was achieved on strain level for the two spoilage strains.

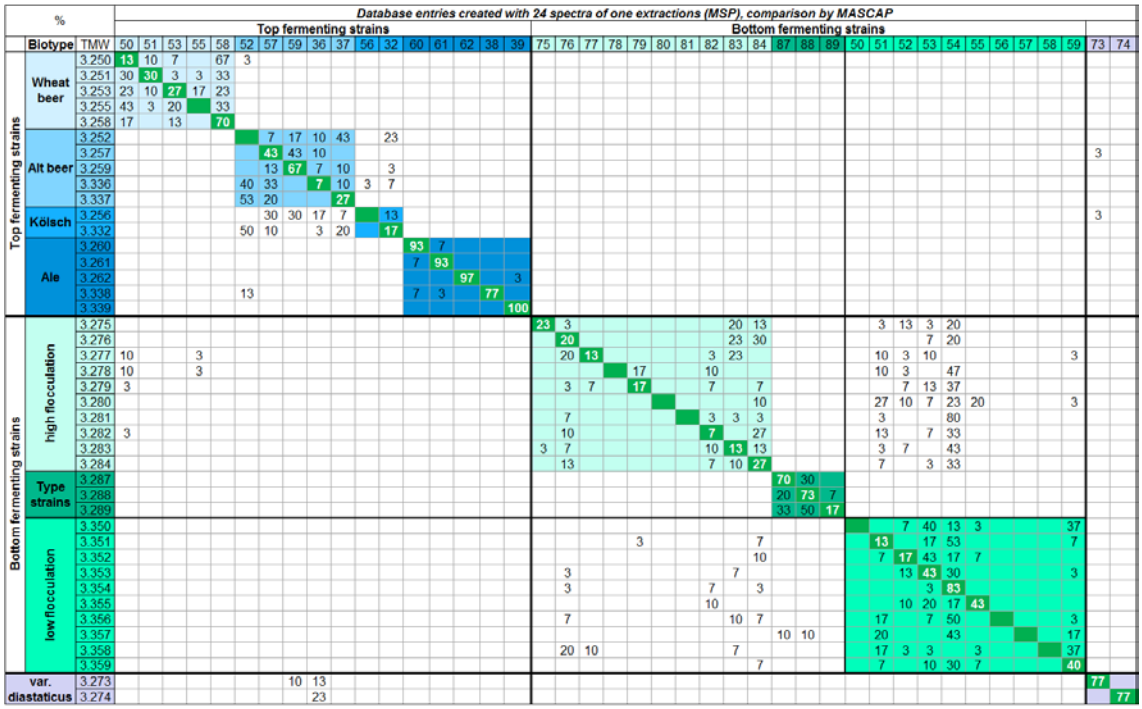

Fig. 3.14: Similarity calculation by MASCAP – test spectra reconciled against DB entries of 42 strains with 24 spectra per strain from one extraction procedure (MAS1*24) of TF strains (blue), BF strains with high (light green) and low flocculating properties (middle green), BF type strains (dark green) and two spoilage strains (violet). The resulting hit rates [%] are displayed: diagonal green boxes show the exact strain match, whereas the color grounded parts indicate different biotypes

To improve the DB search, the library entries were created with four independently measured triplicates (4*3) in parallel to the BT approach. The results are summarized in Fig. 3.15 and Tab. 3.3. The differentiation at species level was achieved in 100 % of the cases.

For top-fermenting *S. c.* strains, an identification rate of 65.0 % was achieved. Considering all samples, a score of 51.5 % could be realized. For wheat beer, the mismatches at strain level were always other wheat beer strains resulting in a biotype match of 100 %. At strain level, an average of only 40 % could be achieved, where the best match was obtained for TMW 3.250 (57.0 %). A large number, i.e. 36.2 % and 34.2 %, of all wheat beer strains matched the DB entries of the strains TMW 3.253 and 3.255, respectively. The worst strain identification rate among the wheat beer strains was obtained by TMW 3.258, where only 10.0 % were identified correctly. Spectra from Altbier strains correlated in 95.2 % with DB entries from other Altbier strains, wrong identifications at the biotype level were only found for TMW 3.336, which was assigned to either one of the two Kölsch strains. On average, Altbier strains showed a higher strain level identification rate compared with wheat beer strains with an average value of 68.6 %. Strain TMW 3.259 frequently mismatched with TMW 3.252 and 3.257. The

Kölsch strains matched with Kölsch strains in 59.5 % of all cases. At strain level, 35.7 % correct identification could be achieved on average. Mismatches occurred always with Altbier strains. The approach used here (in-house software and measurement of four independent triplicates (4*3)) delivered very good results for the differentiation of ale strains. Such strains seemed to be the most unique ones as reflected by an average strain level identification rate of 98.1 %. Only 5.0 % of the strains TMW 3.260 and 3.339 were identified as ale yeast TMW 3.261 and *S. c.* var. *diastaticus* TMW 3.273, respectively. However, the *S. c.* var. *diastaticus* strains matched in 100 % with each other.

Differences in the identification rates when the in-house software approach was used were also observable for *S. p.* strains (BF), but not as distinctive as for *S. c.*. The average identification rate at strain level was 38.3 % (type strains included). Without taking the type strains into account, only 32.4 % matched with their own DB entries. The achieved discrimination at biotype level, i.e. between high and low flocculants, was 63.3 % for high and 81.4 % for low flocculating strains. Among the BF strains, the library entry of TMW 3.354 was the most frequently matched with 15.5 % of the spectra from the 20 BF strains matching this entry, followed by TMW 3.352 (12.6 %), 3.284 (9.5 %), and 3.276 (9.1 %). Interestingly, TMW 3.284 did not even match once with itself. Furthermore, the DB entries of TMW 3.280 and 3.357 were only matched by themselves.

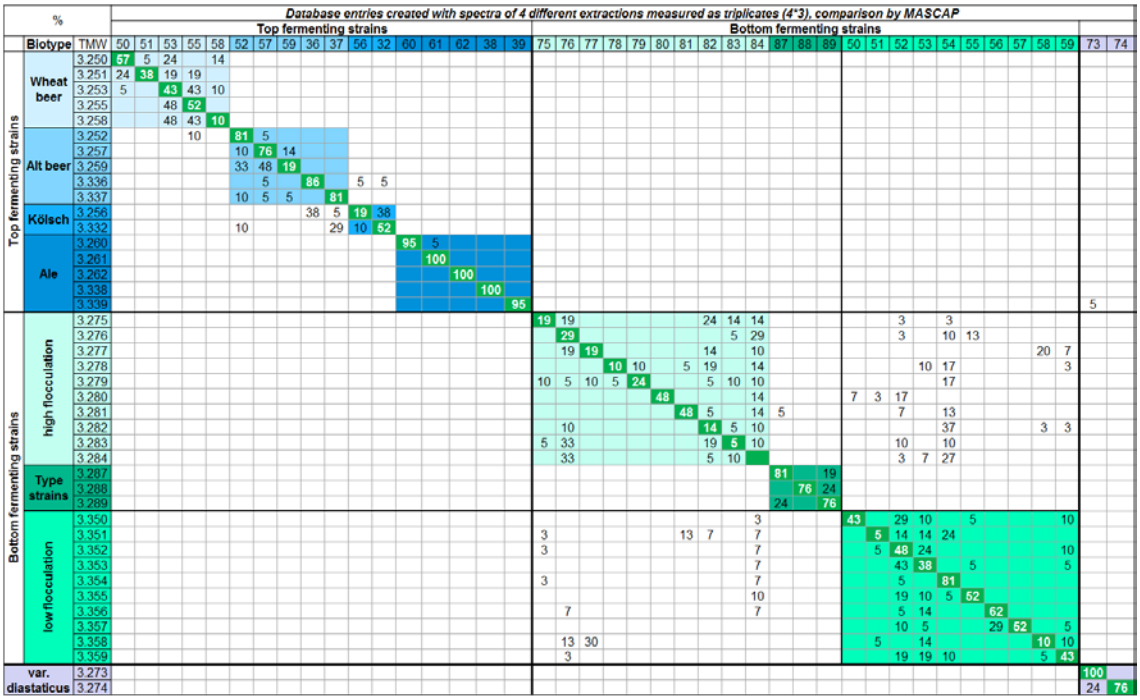

Fig. 3.15: Similarity calculation by MASCAP – test spectra reconciled against DB entries of 42 strains, calculated with 4 independently measured triplicates (n=12) (MAS4*3) of TF strains (blue), BF strains with high (light green) and low flocculating properties (middle green), BF type strains (dark green) and two spoilage strains (violet). The resulting hit rates [%] are displayed: diagonal green boxes show the exact strain match, whereas the color grounded parts indicate different biotypes

The direct comparison of the two methods (1*24 and 4*3) to generate library entries by MAS showed higher match rates for nearly all ecotypes when the 4*3 approach was used. The only exception was the BF low flocculation group, where 86.0 % correct assignment was obtained with the 1*24 method compared with 81.4 % with the 4*3 approach. The same tendency was visible at strain level, where the hit rates for TMW 3.258, 3.259, 3.339, 3.274, 3.275, 3.283, 3.284, 3.351, 3.353, and 3.354 were higher when spectra were reconciled with 1*24 DB entries. In contrast, the 33 remaining strains achieved higher results when the 4*3 DB entry calculation was applied. The summary of average strain level hit rates per biotype and the average correct identification of a specific biotype (Tab. 3.3) illustrate the superiority of the 4*3 compared with the 1*24 method, which becomes apparent comparing the last two columns of Tab. 3.3. Strain level hit rates for wheat beer strains increased from 28 % to 40 %. Altbier strains nearly tripled their score from 28.7 % to 68.6 %. More than four times better rates were achieved for Kölsch strains, i.e. from 8.3 % to 35.7 %. Hits for high flocculants were improved from originally 12.0 % to 21.4 %, and strain level hit rates for low flocculants increased from 24.0 % up to 43.3 %. Matches were also enhanced for the two *S. c.* var. *diastaticus* strains (from 76.7 % to 88.1 %) and the type strains (from 53.3 % to 77.8 %). All in all, the strain level identification rate increased from 34.1 % up to 51.5 %.

Tab. 3.3: Overview of hit rates [%] achieved with different approaches of commercially available MALDI BT and in-house software based on MASCAP implemented in octave. Analysis was conducted with 24 spectra of one extraction (1*24) or 12 spectra of 4 extractions (4*3). The library was composed of 42 strains (TMW 3.250 – 3.262, 3.273 – 3.289, 3.332, 3.336 – 3.339, 3.350 – 3.359)

Software			Biotyper 3.0 [%]		MASCAP [%]	
Library entry generated with			1*24	4*3	1*24	4*3
Differentiation TF / BF			99.2	99.6	99.0	99.9
S. cerevisiae (TF)	strain level	Wheat	23.3	22.7	28.0	40.0
		Altbier	39.3	50.7	28.7	68.6
		Kölsch	36.7	80.0	8.3	35.7
		Ale	83.3	66.7	92.0	98.1
	Ø strain level hit rate		**47.3**	**50.6**	**44.7**	**65.0**
	biotype level	Wheat	94.0	90.0	99.3	100
		Altbier	70.0	68.7	92.0	95.2
		Kölsch	65.0	91.7	15.0	59.3
		Ale	90.0	68.7	97.3	99.0
	Ø biotype level hit rate		**82.4**	**77.6**	**86.7**	**93.6**
S. pastorianus (BF)	strain level	High flocculating	15.7	19.0	12.0	21.4
		Low flocculating	20.3	41.7	24.0	43.3
	Ø strain level hit rate		**18.0**	**30.3**	**18.0**	**32.4**
	biotype level	High flocculating	52.7	63.0	43.3	63.3
		Low flocculating	87.3	72.0	86.0	81.4
	Ø biotype level hit rate		**70.0**	**67.5**	**64.7**	**72.4**
	strain level	*S. c.* var. *diastaticus*	98.3	96.7	76.7	88.1
		Type strains	86.7	81.1	53.3	77.8
	biotype level	*S. c.* var. *diastaticus*	98.3	96.7	76.7	100
		Type strains	95.6	92.2	100	100
Total strain level identification			*38.6*	*45.3*	*34.1*	*51.5*

3.4.2.3 Comparison Biotyper vs. MASCAP (1*24)

A comparison of the 1*24 method in combination with both software applications (cf. Fig. 3.9, Fig. 3.14, and Tab. 3.3) shows that strains were identified correctly in 38.6 % of all cases applying the BT approach (better results achieved for 23 strains) versus 34.1 % using the MAS software (18 strains got better scores here). Only TMW 3.256 did not get one hit regardless of the method used. For wheat beer, the overall correct hits were 23.3 % using BT (with higher scores for TMW 3.251 and 3.255; 94.0 % accordance with biotype level), whereas 28.0 % were obtained using MAS (higher scores for TMW 3.250, 3.253, and 3.258; 99.3 % correct biotype level). When taking a closer look at the misidentifications, considerable software-dependent shifts became visible: only 2.0 % of all 150 wheat beer single spectra matched with the reference entry of TMW 3.250 when BT software was used, whereas 30.7 % of the search entries corresponded with the reference entry for TMW 3.250 when using the MAS approach. Similar observations

could be made with TMW 3.258, where 9.3 % matched with the library entry according to the BT approach, and 45.3 % were assorted to this reference by MAS. For the strain TMW 3.255, a contrary effect was observed, i.e. 66.0 % of all strains matched with this strain using BT, but only 5.3 % matched with this reference entry when MAS was applied.

For Altbier strains this effect was not that obvious. The strain identification rate by BT was generally higher (except for TMW 3.259), but the mismatches with the Kölsch strain TMW 3.332 were also considerably higher. Consequently, the identification score at biotype level dropped from 92.0 % with MAS down to 70.0 % using BT. However, the match of the mentioned Kölsch strain, TMW 3.332, with itself was much better by BT (73.0 % vs. 17.0 % (MAS)). For ale strains, identification worked much better using MAS on both the average strain (92.0 % compared with 83.3 %) as well as the biotype level (97.3 % compared with 90.0 %).

For the BF high flocculating strains BT obtained better results at biotype and strain level (cf. Tab. 3.3). The misidentifications were unevenly scattered, where BT showed highest hit rates with TMW 3.354 (20.7 % of all high flocculating search entries corresponded with it), followed by TMW 3.277 and 3.282 (12.3 % each). On the other hand, TMW 3.284 did not get one match. This is in contrast to the in-house approach, where TMW 3.354 also represented the most frequently hit DB entry, but TMW 3.284 was also frequently hit (13.0 %). Not matches were found for TMW 3.278, 3.280, and 3.281 using the MAS approach. For the low flocculating BF strains, much higher biotype level compliances were obtained with both approaches (BT 87.3 %, MAS 86.0 %). However, at strain level the attribution was comparably poor with 20.3 % by BT and 24.0 % by MAS. The strain TMW 3.358 was the most frequently hit reference entry with over 29.7 % matches of all BF strains, followed by TMW 3.359 with 13.7 %. The used MAS software resulted in a completely different observation, where TMW 3.354 was the most frequently hit DB entry again, followed by TMW 3.353 with 11.8 %, and TMW 3.359 with 7.5 %. TMW 3.350, 3.356, 3.357, and 3.358 entries did not get any match.

For *S. c.* var. *diastaticus* BT enabled an almost completely correct identification obtaining 98.3 % strain hit rate, whereas MAS just achieved 76.7 % on strain and biotype level. Type strains (BF) were assorted to other type strains in 100 % by MAS. However, the rate at strain level was much higher when BT software was employed, i.e. 86.7 % compared with only 53.3 % when MAS was used.

3.4.2.4 Comparison Biotyper vs. MASCAP (4*3)

The comparison of the two software approaches with DB entries created with four different triplicates (4*3) (cf. Tab. 3.3, Fig. 3.11 and Fig. 3.15) in most cases revealed a better performance of the MAS approach regarding average strain identification hit rates. However, Kölsch strains are excluded, since BT scored with 91.7 % compared with 59.5 % (MAS) at biotype level and 80.0 % (BT) compared with 35.7 % (MAS) at strain level. Similarly, strain level identification for type (96.7 % (BT) vs. 88.1 % (MAS)) and spoilage strains (81.1 % vs. 77.8 %) worked better using BT.

Within the wheat beer group, TMW 3.253 seemed to be the “dominant” DB entry in both approaches with a match score of 42.0 % of all wheat beer single spectra when BT was used and 36.2 % for MAS. In contrast, the reference spectrum of TMW 3.255 correlated with the search entries in only 2.0 % employing BT software, but 31.4 % when MAS was used.

For the Altbier group, the misidentification by BT with both Kölsch strains was remarkably high with 30.7 %. In contrast, almost no spectra matched with the DB entry for the Altbier strain TMW 3.257 and TMW 3.259 was hit by just about 1.3 % of all spectra. In comparison with BT, only 1.9 % of all Altbier single strains were assorted to the Kölsch group by MAS, further 2.9 % mismatched with TMW 3.255, and TMW 3.257 led the ranking of matches with 27.6 %.

For Kölsch strains, a proper biotype identification rate was achieved using the BT software (91.7 %), whereas a significantly lower rate was found when MAS was used (59.5 %). Similarly, the use of BT led to better results at the strain level, i.e. 70.0 % for TMW 3.256 and 90.0 % for TMW 3.332 compared with 19.0 % and 52.0 % when MAS was used. For the latter software tool, the false classification of TMW 3.256 was mostly TMW 3.336 (38.0 % of all cases; equal rate as the other Kölsch strain TMW 3.332). Furthermore, TMW 3.332 frequently corresponded to TMW 3.337 (29.0 %).

Ale strains were sorted very distinctly by MAS with 99.0 % correct biotype level hits, whereas a correct assortment of the biotype was achieved only for 68.7 % of all ale strains when BT was employed due to the fact that many samples were assorted to Altbier or Kölsch strains.

One single spectrum of the high flocculation BF strains TMW 3.275, 3.277, 3.281, and 3.282, respectively, matched the DB entry for the Kölsch strain TMW 3.256 applying BT. Additionally, the mentioned strain TMW 3.281 matched once with a wheat beer strain (TMW 3.253). These were the only misidentifications at species level, which did not

occur analyzing MALDI-TOF MS spectra of BF strains with the MAS software. Furthermore, spectra derived from *S. c.* var. *diastaticus* strain TMW 3.274 matched a *S. c.* DB entry twice (TMW 3.336 and 3.337, both Altbier strains). MAS sorted only one spectrum of a *S. c.* strain to a *S. c.* var. *diastaticus* reference. For BF strains, hits on strain and biotype level were slightly better using MAS, which was most clearly visible for low flocculants at biotype level, where 81.4 % (MAS) compared with 72.0 % (BT) correct biotype assortment was achieved.

3.4.2.5 Alternative procedures

Since both the applied BT (Biotyper) and MAS (in-house, MASCAP-based) approach did not yield satisfactory results in discriminating high and low flocculating bottom-fermenting (BF) yeast strains, further variations of the in-house bioinformatics were tested to improve the discrimination between such biotypes. In Fig. 3.16 results are depicted, where triplicates of the test spectra were summarized to one search entry. A discrimination between high and low flocculent strains was now possible for 100 % of the entries made. At strain level, an average of 25.0 % correct identification was reached for the high flocculants, while this rate reached almost 32.0 % for the low flocculants. The most frequently hit reference entry belonged to strain TMW 3.284 with 28.0 % concordance with all search entries generate from the spectra of high flocculating BF strains. Strain TMW 3.358 was the “top scorer” of the low flocculating BF strains with 27.0 %, followed by TMW 3.352 with 23.0 %. In contrast, TMW 3.356 did not even get one match, whereas TMW 3.357 matched in 100 % with itself and with no other entry. In contrast to the approach of the comparison of single spectra against the reference entries, the average strain level score rose for high flocculants from 21.4 % to 25.0 % and dropped for low flocculating strains from 43.3 % down to 32.0 %.

%		high flocculation										low flocculation									
	TMW	75	76	77	78	79	80	81	82	83	84	50	51	52	53	54	55	56	57	58	59
high flocculation	3.275	**10**				10			10		70										
	3.276		**20**		20				30	10	20										
	3.277		10	**20**		20			20	30											
	3.278				**20**	10		10	10	10	40										
	3.279			10	30	**10**				30	20										
	3.280	20			10	10	**30**				30										
	3.281	10	10	10	30			**20**		20											
	3.282	10	10	10					**20**	30	20										
	3.283	10	20		10	20				**30**	10										
	3.284	20								10	**70**										
low flocculation	3.350											**20**			10					30	40
	3.351												**10**	40	20					20	10
	3.352													**50**	20						30
	3.353													40	**20**	10					30
	3.354												10	50	10	**10**					20
	3.355													50			**10**				40
	3.356														20					80	
	3.357																		**100**		
	3.358														10					**90**	
	3.359														40					50	**10**

Fig. 3.16: Similarity calculation by MAS – triplicate measurements of 30 test spectra per strain were summarized (n = 10) of 20 BF *S. p.* strains subdivided in high and low flocculating strains and were checked against DB entries (4*3). The results are displayed as percentages

3.4.2.6 MDS

MDS (Multidimensional Scaling) facilitates the visualization of the level of similarity of individual cases of a dataset. The plane of such graphs can be partitioned into regions based on distances to points in a specific subset of the plane (Voronoi diagram), which enables the visualization of distinct subgroups within a dataset. Fig. 3.17 shows a Voronoi diagram, where sum spectra of two datasets per strain, one for the reference entry and one as a test dataset (marked with “t”) are spread over a 2D map according to their similarities. The advantage, compared with a cluster analysis with a visualization of the results as a dendrogram, is that correlation between all points can be depicted whereas the dendrogram can visualize only single arms. Entries for wheat beer strains generated a distinct area, which is in accordance with the DB comparison displayed in Fig. 3.15. Kölsch and Altbier strains were slightly mixed, which also corresponded with the DB comparison. Ale and “Diastaticus” were mixed as well, although Fig. 3.15 showed a clear assignment of the Ale strains (99 %) and “Diastaticus” (100 %). Among the strains tested and depicted in Fig. 3.17, it was remarkable that test and DB spectra of one strain were not always the nearest neighbors.

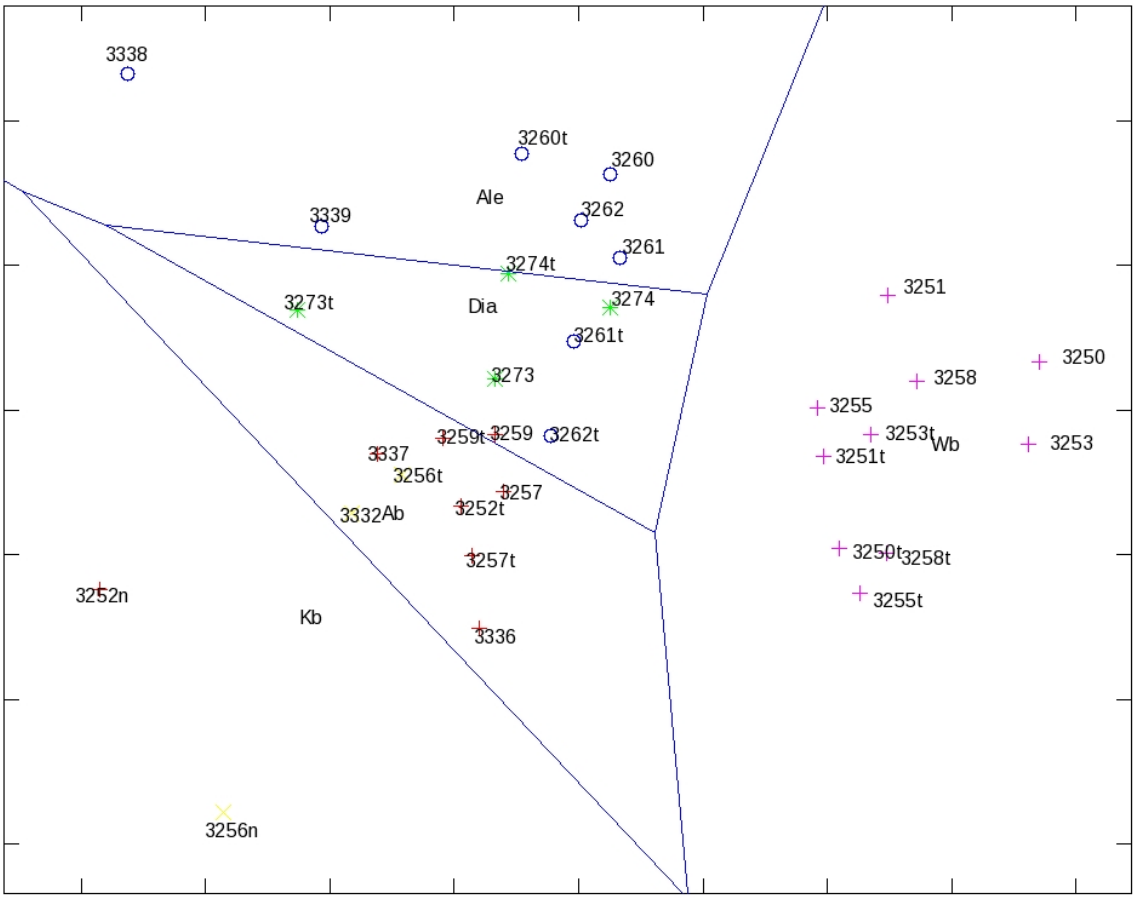

Fig. 3.17: Voronoi diagram based on a decomposition of metric space. Sum spectra of top-fermenting strains (n = 30) were compared (Ab: Altbier, Kb: Kölsch, Wb: wheat beer, Dia: *S. c.* var. *diastaticus*). Numbers indicate strain numbers omitting the dot after the first digit (TMW X.XXX). Data points marked with a "t" indicate entries created on the basis of a test dataset measured on ten different days (30 spectra / strain). Remaining data points represent entries based upon MSP measurements (1*24 / strain)

Fig. 3.18 depicts the MDS of the BF strains, where (A) was calculated on the basis of MSP measurements and (B) compares the test dataset generated on ten different days. Part (A) of the figure enables a distinct separation of the low and high flocculating strain groups, where only strain TMW 3.351 is located at the "border" between both areas. The calculation based on the test spectra (B) shows an overlap of the two groups, where only strains TMW 3.350, 3.355 – 3.358, 3.275, 3.276, and 3.280 are not included in the intersecting set. In both cases (A and B), TMW 3.357 was arranged far away from other strains. Additionally, both diagrams show TMW 3.275 and 3.276 located next to each other, which also applies for the pairs TMW 3.352 / 3.353 and 3.283 / 3.284.

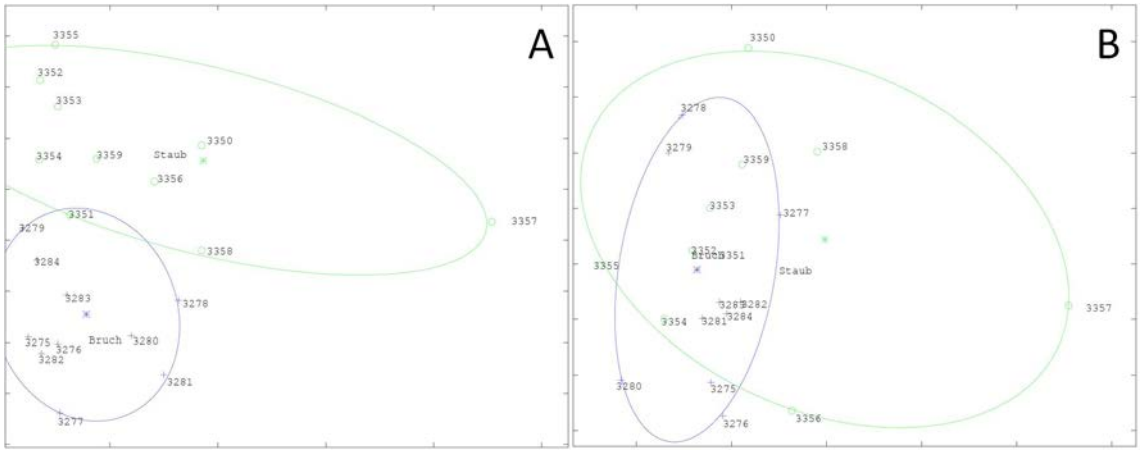

Fig. 3.18: Voronoi diagrams based on a decomposition of metric space of BF strains with “Bruch” (blue): TMW 3.275 – 3.284 (high flocculants) and “Staub” (green): 3.350 – 3.359 (low flocculants). (A) is generated with MSP measurements (1*24), (B) is calculated with a test dataset measured on ten different days (n=30 / strain)

3.4.2.7 DAPC

The DAPC (discriminant analysis of principal components) grouping of the TF (top-fermenting) strains performed by R employing “find.clusters” function resulted in a scatterplot, which is displayed in Fig. 3.19-A. Group 1 and 2 are separated very well, whereas group 3, 4, and 5 are located next to each other and overlap more or less completely.

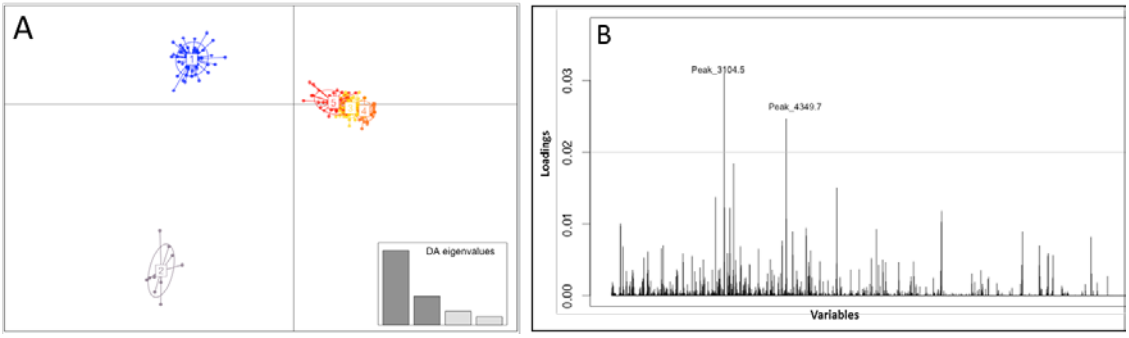

Fig. 3.19: (A) DAPC performed with the “find.clusters” function of R, without any guidelines for grouping and the corresponding loading plot (B)

The composition of the five groups is detailed in Tab. 3.4, where a remarkable distribution can be observed, which does not concur with the biotype sorting except for group 4, which solely comprises wheat beer strains. The multiple spectra per strain are, in most cases, distributed over several groups. The only exceptions are the strains TMW 3.256, 3.338, and 3.343.

Group 1 predominantly contained Altbier strains and 86 % of the spectra derived from Kölsch strains. Additionally, all spectra from the ale beer strain TMW 3.338 were assigned to this group. Group 2 combined only three strains including all spectra from wheat beer strain TMW 3.343. In group 3, a conglomeration of nearly all strains except Kölsch strains was found. Group 4 presented the only group that was exclusively comprised of strains belonging to a distinct biotype group (application in brewing wheat beer). Five of six wheat beer strains were assigned to this group. The predominant representatives of group 5 were Altbier and ale beer strains (four of five strains from each of these two biotypes). In contrast to groups 3 and 4, almost no spectra from wheat beer strains were assorted to group 5.

Tab. 3.4: Detailed composition of the five groups generated by “find.clusters” of R software with 11 summarized triplicate measurements per strain. The absolute number of spectra sorted to a group is listed next to the percentage by which strains assigned to a specific biotype (industrial application) appear in a certain group

	TMW	Group 1		Group 2		Group 3		Group 4		Group 5	
		abs.	%	abs.	%	abs.	%	abs.	%	abs.	%
Wheat beer	3.250					2		9			
	3.251					2		8		1	
	3.253				78.6	6	26.8	5	100		2.8
	3.255					4		7			
	3.258					5		7			
	3.343			11							
Altbier	3.252	9		2							
	3.257					7				4	
	3.259		39.3		14.3	6	18.3			5	44.4
	3.336	7								4	
	3.337	8								3	
Kölsch	3.256	11	31.1								8.3
	3.332	8								3	
Ale	3.260					9				2	
	3.261					9				2	
	3.262		29.5		7.1	9	33.8			2	27.8
	3.338	11									
	3.339	6		1						4	
var. *diastaticus*	3.273					6	21.1			5	16.7
	3.274					8				3	

In addition to the described cluster analysis (Fig. 3.19 and Tab. 3.4), which did not result in a meaningful strain differentiation according to the proposed biotypes, a hierarchical cluster analysis was performed, which is shown in Fig. 3.20. For clarity the cluster arms

in Fig. 3.20 are not labelled individually. Colored bars indicate grouping of spectra, where bar A marks the grouping according to the biotype and bar B displays the grouping of the unsupervised (no cross-validation or predefined classes) clustering by “find.clusters”. As Fig. 3.20 clearly demonstrates, this approach did not lead to a convincing result.

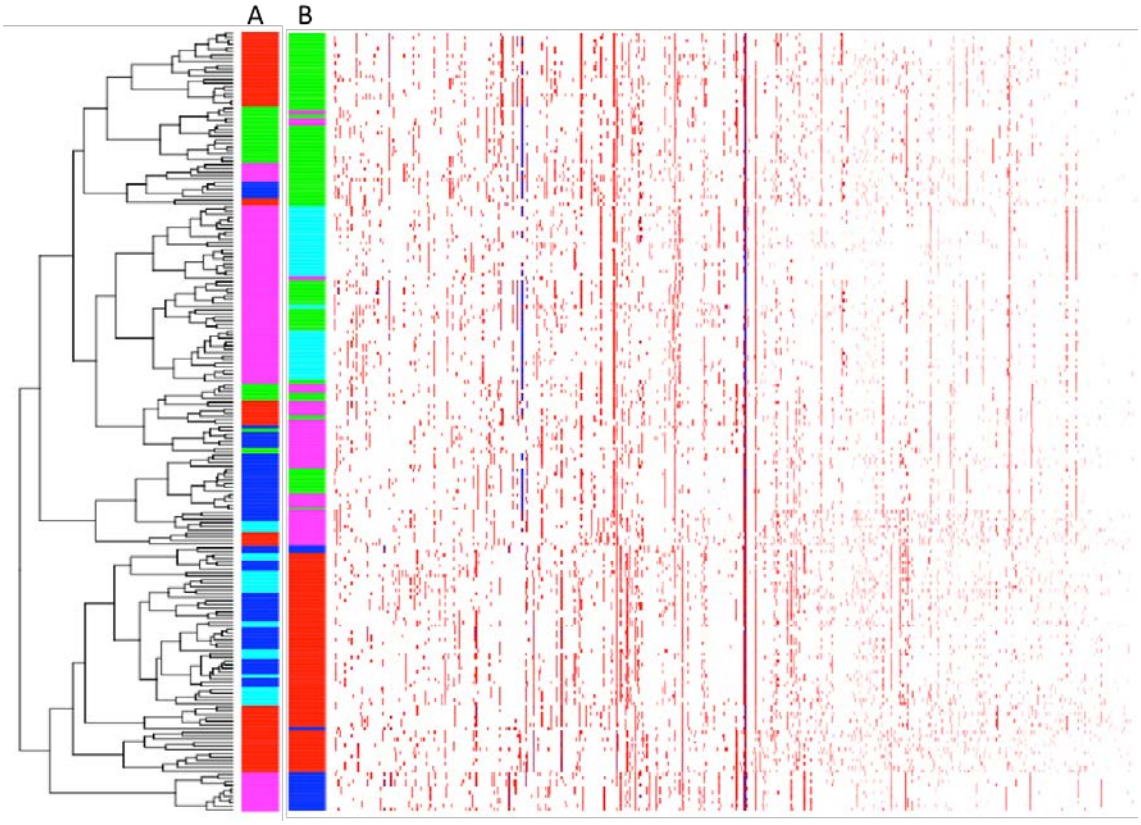

Fig. 3.20: Hierarchical cluster analysis. Colored bars mark the affiliation of MALDI-TOF MS spectra from different yeast strains according to the biotype of a strain (A) and the grouping calculated by the (unsupervised) “find.clusters” function (B) Color code A: ale beer (●), Altbier (●), diastaticus (●), Kölsch (●), wheat beer (●) Color code B: group 1 (●), group 2 (●), group 3 (●), group 4 (●), group 5 (●)

To carry out a supervised DAPC, a cross-validation was performed first to provide an optimized number of PCs (principal components) to be retained. Therefore, the dataset was randomly stratified split into a test dataset (30%), whereas the remaining data was used as training dataset of which again 10 % were used for cross-validation “xvaldapc” by R software. This ensured that finally one member of each group was represented in both datasets. Fig. 3.21-A displays the results of the cross-validation, where the number of PCs retained in each DAPC varies along the x-axis and the proportion of successful outcome prediction varies along the y-axis. Each point on the graph represents individual replicates, and the blue cloud symbolizes the density of such points in different regions of the plot (Jombart, 2013). Approximately 50 was identified as the optimal number of PCs to be retained, which was used for the subsequent DAPC calculation with the

training dataset considering the actual biotype grouping of the strains tested. The model calculated is visible in Fig. 3.21-B, where clearly separated groups of ale (blue), *S. c. var. diastaticus* (yellow), and wheat beer (red) could be identified. Only Altbier (grey) and Kölsch (brown) groups were slightly mixed. The loading plot in Fig. 3.21-C indicates which masses of the underlying MALDI-TOF MS spectra primarily contributed to the discrimination between the biotype groups. Finally, Fig. 3.21-D displays the DAPC of the test dataset with a similar separation of wheat beer, ale, and diastaticus and overlapping groups containing Kölsch and Altbier strains, which is very similar to the grouping for the training dataset and indicates the validity of the system.

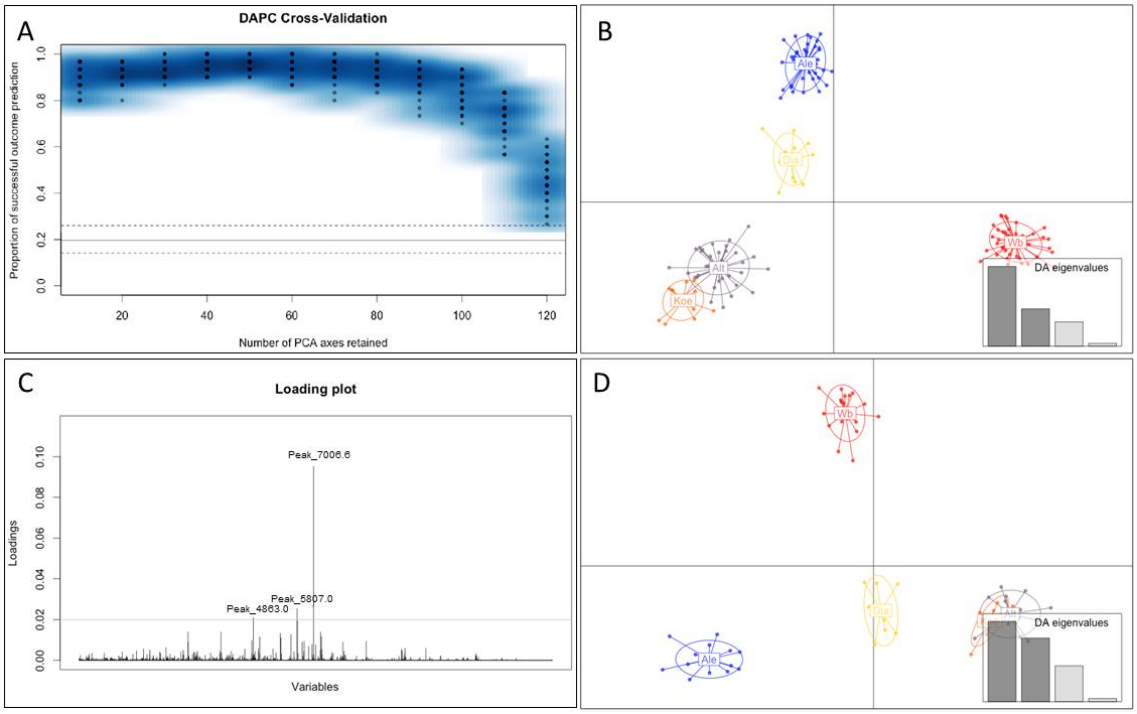

Fig. 3.21: Supervised DAPC with (A) the results of the cross-validation to derive the optimized number of PCs to be retained, (B) the DAPC model of the training dataset, (C) the loading plot of the DAPC model with the training dataset, and (D) the DAPC model of the test dataset.

With the function “predict.dapc”, group memberships were predicted based on the DAPC results. The predicted groups corresponded to the actual biotype groups for all wheat beer, ale, and *S. c.* var. *diastaticus* strains. Only 25 % of the spectra obtained from the two Kölsch strains (TMW 3.256 and 3.332) were sorted to the Altbier group, and 19 % of the Altbier strain spectra (from strains TMW 3.252, 3.336) were identified as Kölsch strains. The trend that the Altbier and Kölsch groups cannot be separated completely is

confirmed even when only these two groups are considered for calculation (data not shown).

The DAPC analysis of the BF strains resulted in a maximum differentiation value of 83 %. In the approach visualized in Fig. 3.22, the spectra that were separated incorrectly belong to the strains TMW 3.277, 3.280, 3.281, and 3.351. Since the groups for training and validation datasets were split randomly, the possibility could not be ruled out that, in some cases, only one spectrum per strain is organized in the validation dataset, which puts the definition of a mismatching data pool into perspective. However, even when the groups were organized differently, the hit rate remained more or less constant.

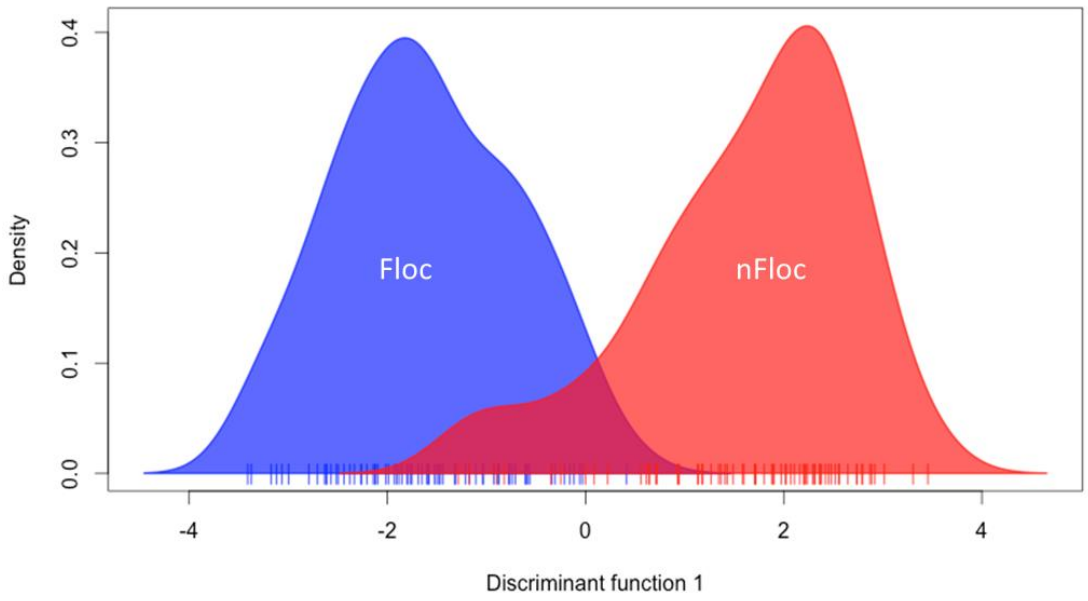

Fig. 3.22: DAPC analysis of ten high flocculant (Floc) and ten low or non-flocculant (nFloc) BF *S. p.* strains, all measured in triplicate on ten different days (n=30). Triplicates were summarized (n=10) and the whole dataset (n=200) was separated randomly in two groups, first one for modelling, the second group to verify the calculated system. This procedure was conducted several times to ensure that the results of the grouping are independent of the chosen dataset. A correct classification for flocculent and non-flocculent strains could be achieved in 83 % of the cases.

3.5 Wine yeast typing

3.5.1 Reference method

The numerical analysis of generated, digitized delta-PCR patterns of the 33 wine yeast strains tested (cf. Fig. 3.23) yielded the dendrogram displayed in Fig. 3.24. Considering 95 % similarity as the arbitrary threshold to define biotypes, the 33 strains could be divided into 23 groups based on their fingerprint similarity. According to this definition, the Champagne strains *S. c. x S. uvarum* (cluster U) could not be separated along their delta-PCR fingerprint just as well as Lalvin W46/Lalvin W27 (cluster A), Lalvin V1116/Uvaferm WAM (cluster I), Lalvin ICV D47/Lalld19 (cluster M), and, finally, Lalld2/Lalvin ICV OKAY® (cluster N).

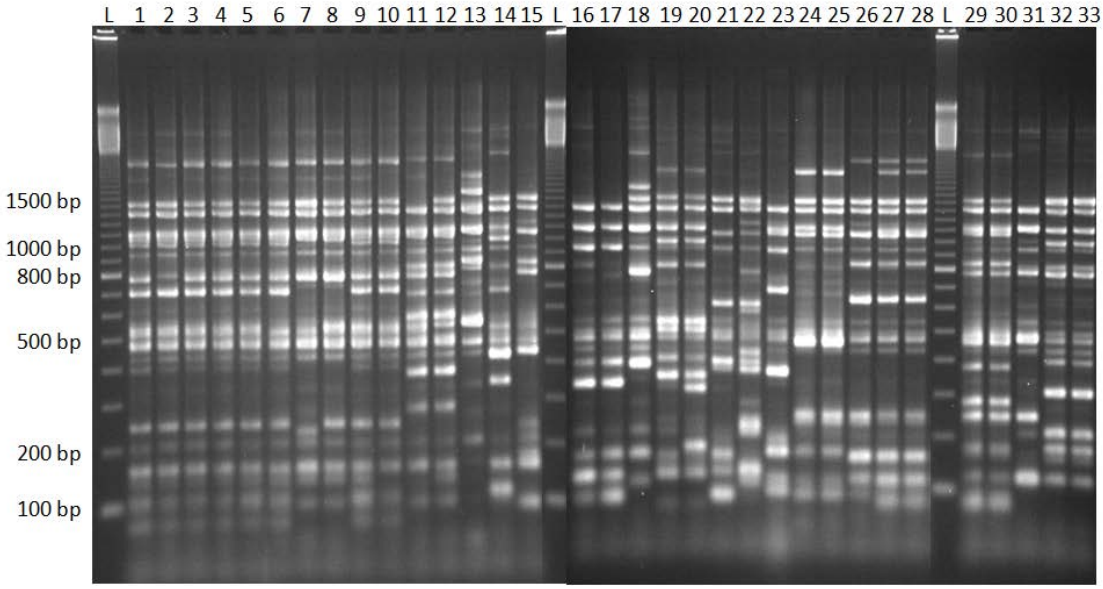

Fig. 3.23: Fingerprint of 33 *Saccharomyces* strains using delta-PCR with primers delta12-delta21. Electrophoresis was performed on a 1.5 % agarose gel, which was run at 100 V for 90 minutes. Lanes: L, ladder 100 bp (GE Healthcare); 1, Lalvin EC1118®; 2, Lalvin CH14; 3, Lalvin QA23®; 4, Lalvin DV10; 5, Levuline CHP; 6, Vitilevure Quartz; 7, Zymaflore X5; 8, Anchor VIN13; 9, Affinity$_{ECA5}$; 10, Lalld6; 11, CrossEvolution; 12, Vitilevure Elixir; 13, Lalvin S6U; 14, Lalvin CY3079; 15, Zymaflore F15; L; 16, Lalvin ICV D254; 17, Uvaferm HPS; 18, Vitilevure C; 19, Levuline ALS; 20, Anchor VIN7; 21, Lalvin 71B; 22, Siha levactif 6; 23, Levucell SB20; 24, Lalvin ICV OKAY®; 25, Lalld2; 26, Lalvin ICV K1M; 27, Uvaferm WAM; 28, Lalvin V1116; L; 29, Lalvin ICV D47; 30, Lalld19; 31, Lalvin W15; 32, Lalvin W27; 33, Lalvin W46 (performed by Lallemand Inc., Montréal, Canada)

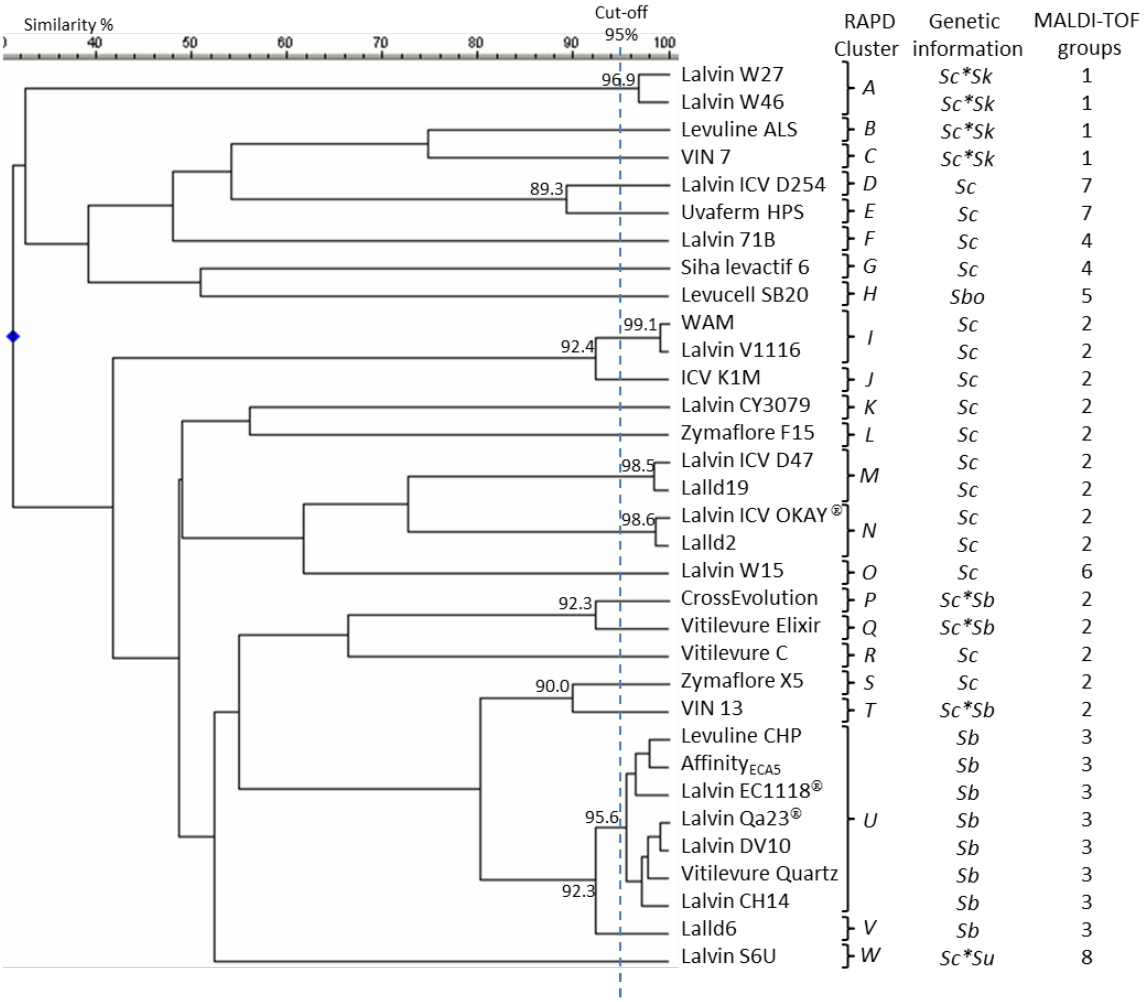

Fig. 3.24: Strain clustering with GelCompar II V5.0 according to delta-PCR fingerprints (performed by Lallemand Inc., Montréal, Canada). Delta-PCR clusters were defined according to the 95 % similarity cut-off. MALDI-TOF MS groups were defined according to the results of the MDS calculation (cf. Fig. 3.26). Genetic information: *Sc*: *Saccharomyces cerevisiae*; *Sk*: *S. kudriavzevii*; *Sb*: *S. bayanus* var. *bayanus*; *Su*: *S. bayanus* var. *uvarum*; *Sbo*: *S. cerevisiae* (*boulardii*)

3.5.2 Analysis of MALDI-TOF Mass Spectra

MALDI-TOF MS spectra of the first 15 strains (initially blind-coded strain numbers Lalld1 – 15) were recorded as described above (cf. section 2.4.5.1). The results of a similarity calculation are shown in Tab. 3.5. The hit rates [%] of the test spectra are displayed in comparison to the DB entries. The grey marked boxes are the correctly identified samples at strain level. An average hit rate at strain level of 68.1 % could be achieved when the DB entry was generated with 21 spectra of seven biological replicates using BT software (Biotyper software, Tab. 3.5-A). The DB comparison resulted in an average of only 51.1 % correct matches at strain level when 24 spectra of a single extraction were summarized according to the SOP (standard operating practice of manufacturer, Tab.

3.5-B). In-house software predicated on MAS (MASCAP-based) for sum spectra generation combined with a DB-based spectra identification tool performed with seven biological replicates as DB entries resulted in a hit rate of 63.7 % (Tab. 3.5-C).

Tab. 3.5: DB comparison. The DB entries of the training dataset (Lalld1-15) were generated by 21 summarized single spectra of seven biological replicates using BT software (A), by 24 summarized single spectra of one replicate using BT software (B), and by 21 summarized single spectra of seven biological replicates calculated with MAS (C). Entries were matched against nine test spectra (three biological replicates). The resulting hit rate is displayed [%]. Grey boxes indicate correct matches at strain level

A	Database strains (21 spectra of 7 bioreplicates) by Biotyper [hit rate, %]														
Test Strains	Lalvin W46	Lalld2	DV10	Levuline CHP	Lalvin W15	Lalld6	Lalvin ICV D47	Lalvin V1116	Lalvin 71B	Lalvin EC1118®	Levucell SB20	Lalvin QA23®	Lalvin ICV OKAY®	Affinity$_{ECA5}$	Lalvin W27
Lalvin W46	100														
Lalld2		22.2											77.8		
DV10			33.3	44.4		22.2									
Levuline CHP			22.2	44.4		22.2						11.1			
Lalvin W15					100										
Lalld6						100									
Lalvin ICV D47							100								
Lalvin V1116								77.8					22.2		
Lalvin 71B									100						
Lalvin EC1118®						22.2				11.1		11.1		55.6	
Levucell SB20						11.1					77.8			11.1	
Lalvin QA23®			11.1	33.3		22.2				11.1		0		22.2	
Lalvin ICV OKAY®		11.1											88.9		
Affinity$_{ECA5}$														100	
Lalvin W27	33.3														66.7
Ø	68.1														

B	Database strains (24 spectra of 1 bioreplicate) by Biotyper [hit rate, %]														
Test Strains	Lalvin W46	Lalld2	DV10	Levuline CHP	Lalvin W15	Lalld6	Lalvin ICV D47	Lalvin V1116	Lalvin 71B	Lalvin EC1118®	Levucell SB20	Lalvin QA23®	Lalvin ICV OKAY®	Affinity$_{ECA5}$	Lalvin W27
Lalvin W46	100														
Lalld2		88.9											11.1		
DV10			77.8			22.2									
Levuline CHP			77.8	0		22.2									
Lalvin W15					100										
Lalld6			33.3			66.7									
Lalvin ICV D47		11.1				22.2	66.7								
Lalvin V1116		66.7						0					33.3		
Lalvin 71B									100						
Lalvin EC1118®		22.2	11.1			33.3				0		33.3			
Levucell SB20			22.2			44.4					33.3				
Lalvin QA23®			55.6			11.1						33.3			
Lalvin ICV OKAY®		77.8											22.2		
Affinity$_{ECA5}$				22.2										77.8	
Lalvin W27	100														0
Ø	51.1														

C	Database strains (21 spectra of 7 bioreplicates) by MASCAP [hit rate, %]														
Test Strains	Lalvin W46	Lalld2	DV10	Levuline CHP	Lalvin W15	Lalld6	Lalvin ICV D47	Lalvin V1116	Lalvin 71B	Lalvin EC1118®	Levucell SB20	Lalvin QA23®	Lalvin ICV OKAY®	Affinity$_{ECA5}$	Lalvin W27
Lalvin W46	44.4														55.6
Lalld2		22.2											77.8		
DV10			22.2	22.2		22.2						33.3			
Levuline CHP				33.3								66.7			
Lalvin W15					100										
Lalld6			11.1			44.4						44.4			
Lalvin ICV D47							100								
Lalvin V1116								100							
Lalvin 71B									100						
Lalvin EC1118®			11.1							0		66.7		22.2	
Levucell SB20											100				
Lalvin QA23®						11.1				22.2		55.6		11.1	
Lalvin ICV OKAY®													100		
Affinity$_{ECA5}$												33.3		66.7	
Lalvin W27	33.3														66.7
Ø	63.7														

Wrongly assigned entries provided valuable information about the similarity of the tested strains according to their peptide fingerprint so that it was possible to prearrange the strains in six groups for further analysis. With little variations depending on the software tool and summarizing strategy, the strains could be assigned to group 1 consisting of Lalvin W46 and Lalvin W27, group 2 including Lalld2, Lalvin V1116, and Lalvin ICV OKAY®, and group 3 consisting of Lalvin DV10, Levuline CHP, Lalld6, Lalvin EC1118®, Lalvin QA23®, and Affinity$_{ECA5}$. Strains Lalvin W15, Lalvin ICV D47, and Lalvin 71B could be each considered as forming a separate group, and Levucell SB20 could be assigned

to group 3, but probably rather presented a single group. In order to check the manual grouping, an MDS with Voronoi triangulation based on similarity computations was created (cf. Fig. 3.25). This visualized the differences between the strains on a 2D-map, where longer distances between data points indicate stronger differences in MALDI-TOF MS fingerprints (sum spectra of all 30 measurements per strain). For group 1 (Lalvin W46, Lalvin W27), group 2 (Lalld2, Lalvin V1116, Lalvin ICV OKAY®), group 3 (Lalvin DV10, Levuline CHP, Lalld6, Lalvin EC1118®, Lalvin QA23®, Affinity$_{ECA5}$), the prearranged organization was confirmed. Additionally, strains Lalvin W15 and Lalvin 71B considered as forming a separate group as well as strain Levucell SB20 (where it was not completely complete clear whether it falls into a distinct group), were found separated from the other strains and designated as group 4 (Lalvin 71B), group 5 (Levucell SB20), and group 6 (Lalvin W15). Only strain Lalvin ICV D47 (previously considered as forming a separate group) seemed to be similar to group 2 rather than being separated in a single group. A subsequent comparison revealed that the genetic background and the known industrial (recommended) application of the strains tested correspond very well with this grouping (cf. Fig. 3.25, italic letters). Thus, the number of blind-coded strains (n=18) to be included in the analysis was enlarged. Samples were measured in an analogous manner as done before. The different possibilities to evaluate the results obtained are discussed in section 4.6.1. The MDS calculation based on the recommended industrial application of the strains is displayed in Fig. 3.26. Strain grouping as a result of the DB comparison is summarized in Tab. 3.6.

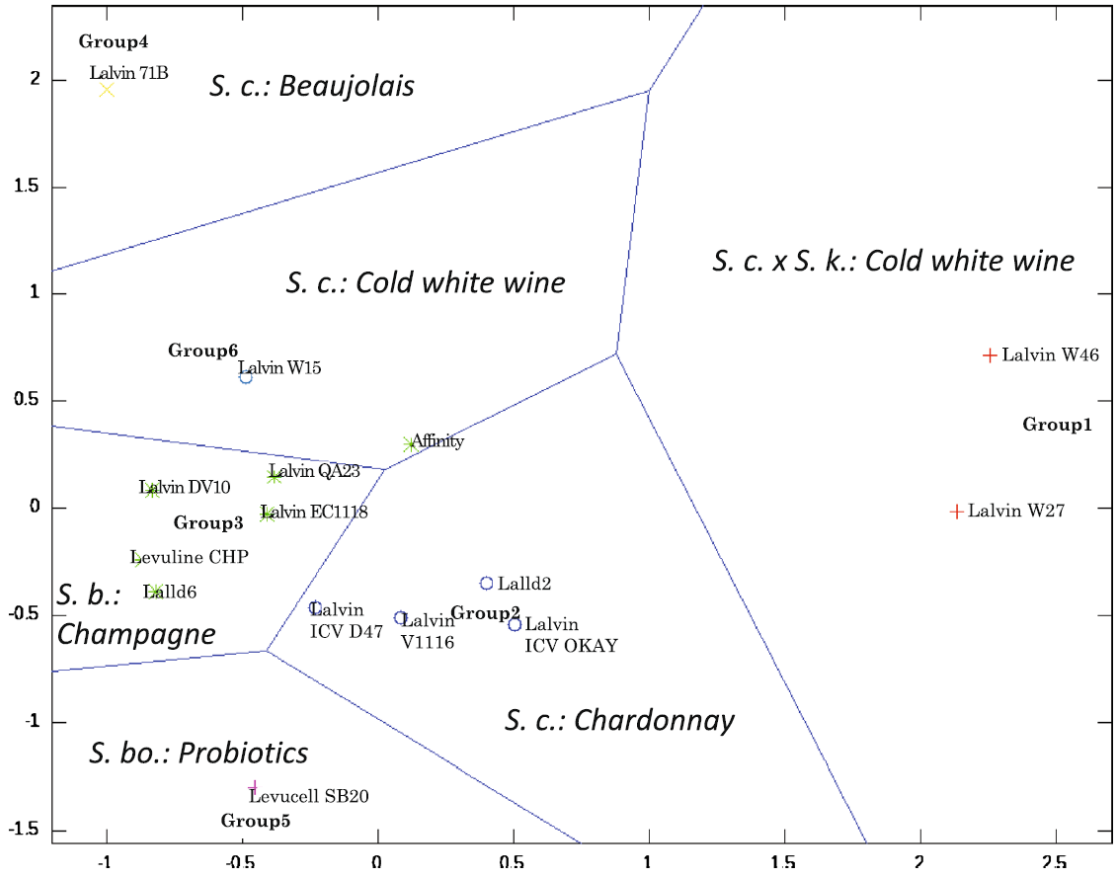

Fig. 3.25: MDS of the training dataset based on 30 replicates (10 biological * 3 technical) per strain. The graph is based on a decomposition of metric space by distances between sets of points, i.e. here the prearranged groups (indicated in italics). Each of the “cells” contain one focus, which is marked by the group number

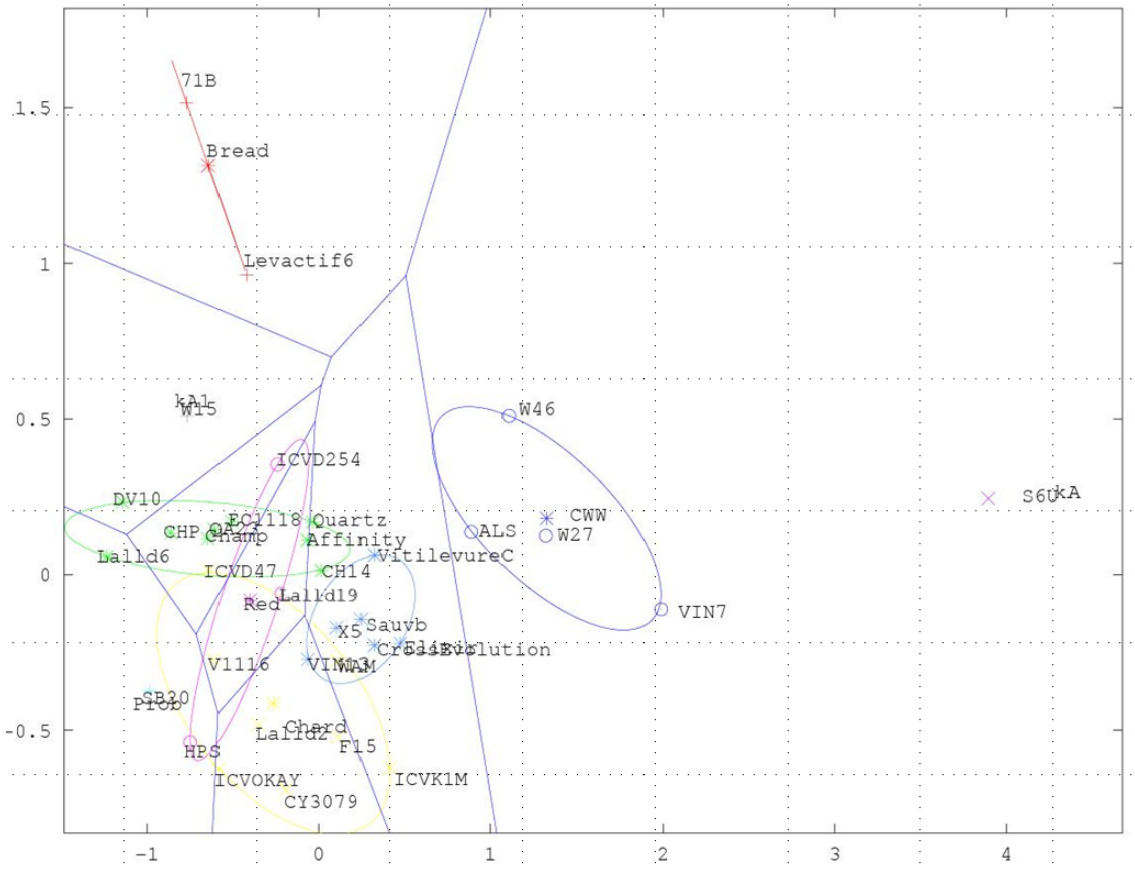

Fig. 3.26: MDS of the complete dataset based on 30 replicates (10 biological * 3 technical) per strain. The graph is based on a decomposition of metric space by distances between sets of points, i.e. here the prearranged groups. Each of the “cells” contain one focus, which is marked by the group name (abbreviation of the recommended industrial application, i.e. CWW, Bread, Sauvb, Chard, Champ, Prob and Red)

Tab. 3.6: Grouping according to the MDS analysis including strain numbers, genetic background information, and recommended industrial application. With *S. c.: S. cerevisiae, S. k.: S. kudriavzevii, S. b.: S. bayanus var. bayanus, S. u.: S. bayanus var. uvarum,* and *S. bo.: S. cerevisiae (boulardii)*

Group	Strains training	Strains validation	Genetic information	Recommended Application
1	Lalvin W46, Lalvin W27	Anchor VIN7, Levuline ALS	*S. c. x S. k.*	Cold white wine
2	Lalld2, Lalvin ICV D47, Lalvin V1116, Lalvin ICV OKAY®	Lalvin CY3079, Lalld19, Lalvin ICV K1M, Uvaferm WAM, Zymaflore F15, Zymaflore X5, Vitilevure C	*S. c.*	Chardonnay, Sauvignon Blanc
		CrossEvolution, Anchor VIN13, Vitilevure Elixir	*S. c. x S. b.*	
3	Lalvin DV10, Lalvin EC1118®, Lalvin QA23®, Levuline CHP, Lalld6, Affinity$_{ECA5}$	Lalvin CH14, Vitilevure Quartz	*S. c. x S. u.*	Champagne
4	Lalvin 71B	Siha levactif 6	*S. c.*	Miscellaneous
5	Levucell SB20		*S. bo.*	Probiotics
6	Lalvin W15		*S. c.*	Cold white wine
7		Uvaferm HPS, Lalvin ICV D254	*S. c.*	Red wine
8		Lalvin S6U	*S. c. x S. u.*	White / Rosé wine

The groups CWW (= group 1, cold white wine, *S. c. x S. kudriavzevii*), Bread (= group 4, Lalvin 71B, Levactif 6, *S. c.*), Prob (= group 5, probiotics, *S. boulardii*), W15 (= group 6, *S. c.*, cold white wine strain Lalvin W15), and S6U (= group 8, white/rosé wine, *S. uvarum* strain Lalvin S6U) could be precisely separated from each other and strains not belonging to one of these groups. To focus on the groups of Champagne (Champ), red wine (Red), Sauvignon Blanc (Sauvb), and Chardonnay (Chard) strains, which were more difficult to distinguish, the above mentioned strains were omitted for further analysis. The result is shown in Fig. 3.27.

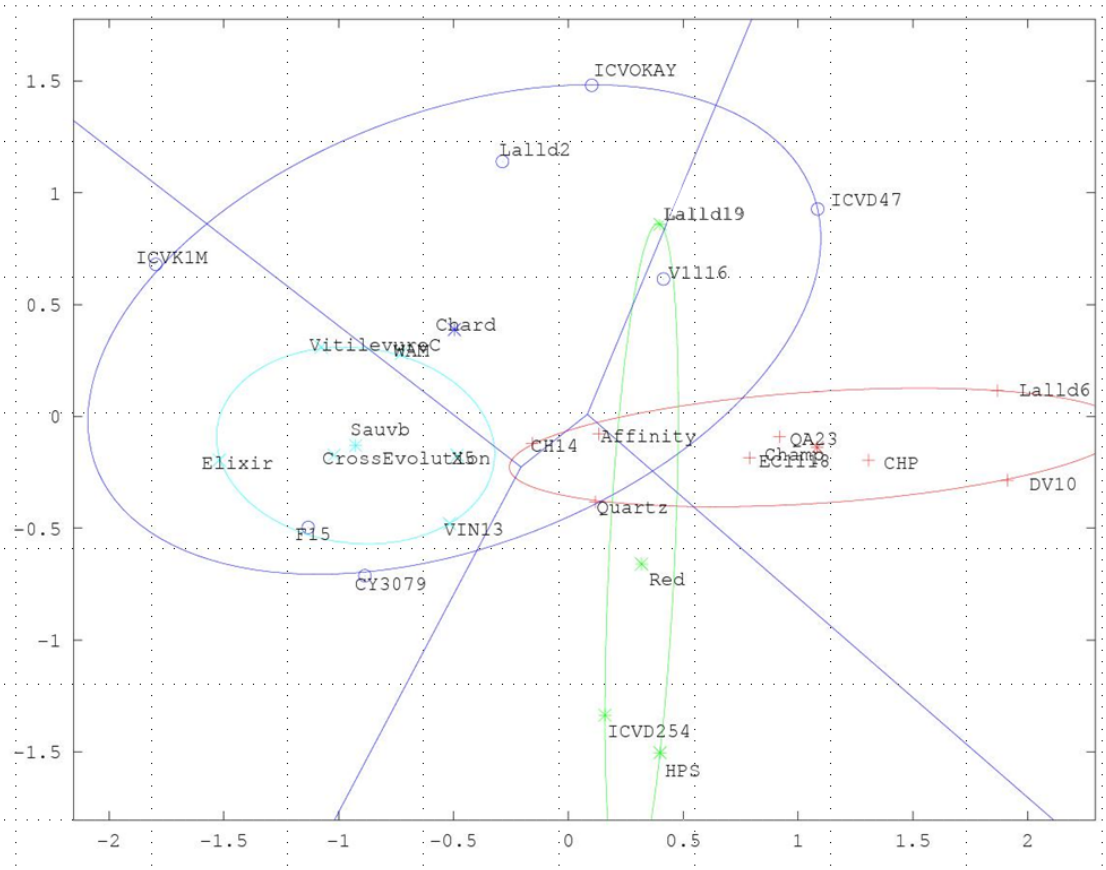

Fig. 3.27: MDS solely containing strains that were hardly differentiable in Fig. 3.26. Data points represent the analysis of 30 replicates per strain. The graph is based on a decomposition of metric space by distances between sets of points, i.e. the prearranged groups. The MDS is divided into “cells” each containing one focus, which is marked by the group name, i.e. Champ (Champagne, brown), Red (red wine, green), Sauvb (Sauvignon Blanc, teal), and Chard (Chardonnay, purple)

Champagne strains (marked with a brown plus) were found as a more or less distinct group in the middle of the right quadrant. The center of the ellipse is indicated by an asterisk and the group annotation, in this case “Champ”. The red wine strains Uvaferm HPS and Lalvin ICV D254 (marked with green asterisks) grouped separately in the lower region. However, strain Lalld19 was found far away, i.e. located in the Chardonnay group next to the Chardonnay strain Lalvin V1116. Finally, the Sauvignon Blanc strains (Uvaferm WAM, Vitilevure C, CrossEvolution, Anchor VIN13, Vitilevure Elixir, and Zymaflore X5), visualized with a teal “x” were clearly located in the left part of the MDS but superposed by the group of Chardonnay strains (circles), which were widely spread in this presentation.

In addition to this initial visualization of strain similarities/differences by MDS graphs based on the prearranged groups, a procedure to select the optimal number of clusters was performed on the basis of the lowest associated BIC (Bayesian Information

Criterion) number analyzed by DAPC (discriminant analysis of principal components) using the *adegenet* package for the R software (Jombart, 2013; Team, 2011). For this analysis, all 30 measurements of the 33 strains were taken into account (n=990). The resulting scatterplot allowed for a graphical assessment of the structures between groups and enabled the stepwise elimination of clearly separable groups. Firstly, Lalvin S6U could be recognized as a distinct group (G1). A new calculation without the Lalvin S6U data was performed isolating Lalvin 71B and Siha levactif 6 as group 2 (G2). In the next step, Uvaferm HPS and Lalvin ICV D254 (G3) were separated, followed by Lalvin W15 (G4). The last clearly differentiable group (G5) included Lalvin W46 and Lalvin W27 as well as VIN7 and Levuline ALS. For the remaining strains, the measurement variability (variability among the different MALDI-TOF MS spectra obtained for one strain) was partly higher than the differences between strains. Thus, data considered for further analysis comprised not all measurements of one strain to facilitate strain discrimination and strains were assigned to a specific group when a two-third majority of matching spectra was achieved. Doing so, further groups could be identified: G6 contained Lalld2, Lalvin ICV D47, Levucell SB20, Lalvin ICV OKAY®, and Lalld19. G7 included Lalvin V1116, Lalvin ICV K1M, and Uvaferm WAM. G8 could be comprised Lalvin DV10, Levuline CHP, Lalld6, Lalvin EC1118®, and Lalvin QA23®. Vitilevure Quartz as a single strain formed G9, and Lalvin CY3079, Zymaflore F15, Vitilevure C, CrossEvolution, Anchor VIN13, Vitilevure Elixir, and Zymaflore X5 formed G10. Finally, $\text{Affinity}_{\text{ECA5}}$ and Lalvin CH14 were assigned to G11. The strains were renamed according to this new prearranged grouping and tested in the same manner as described above to confirm the groups by checking the clusters for their accuracy. Fig. 3.28-A displays the grouping with all spectra organized in 11 groups. It was a subjective decision whether G1 or G5 should be eliminated as a first step, since both groups were clearly separated from the main group. However, G1, which consisted of the spectra from the strain Lalvin S6U (*S. uvarum*) were excluded first. The stepwise elimination continued with G3, which included the *S. c.* red wine strains and was also clearly separated after the first exclusion (cf. Fig. 3.28-B). This was followed by G5 comprised of spectra from the four cold white wine *S. kudriavzevii* strains (cf. Fig. 3.28-C to -D). In the next step, the strains Lalvin 71B and Siha levactif 6 (G2), both used e.g. in bread making, were eliminated as a separate group (cf. Fig. 3.28-D), just like strain Lalvin W15 (G4, cf. Fig. 3.28-E), which all confirmed the MDS analysis. However, after this step (G4 exclusion), the results of the MDS and the DAPC analysis differed slightly, which can be observed in Fig. 3.28-F and Tab. 3.6. The strains included in G10 of the DAPC were all part of group 2 according to the MDS, used for Chardonnay and Sauvignon Blanc, and genetically partly *S. c.* and

S. c. x S. bayanus. Fig. 3.28-F shows a clearly confined G10 (according to the DAPC analysis) with a few outliers, which were located nearest to the center of G10, but overlap with one data point assigned to G11. The clear separation of G6 and G7 (Fig. 3.29-G) enabled their deletion (cf. Fig. 3.29-G to -H and -H to -I). These two groups contained *S. c.* strains all belonging to group 2 of the MDS results, except *S. boulardii*, which was organized separately in group 5 by MDS. The final three groups G9 (Vitilevure Quartz), G11 (Affinity$_{ECA5}$ and Lalvin CH14), and G8 (Lalvin EC1118®, Lalld6, Levuline CHP, Lalvin DV10, Lalvin QA23®) were all Champagne strains (cf. Fig. 3.29-I, J).

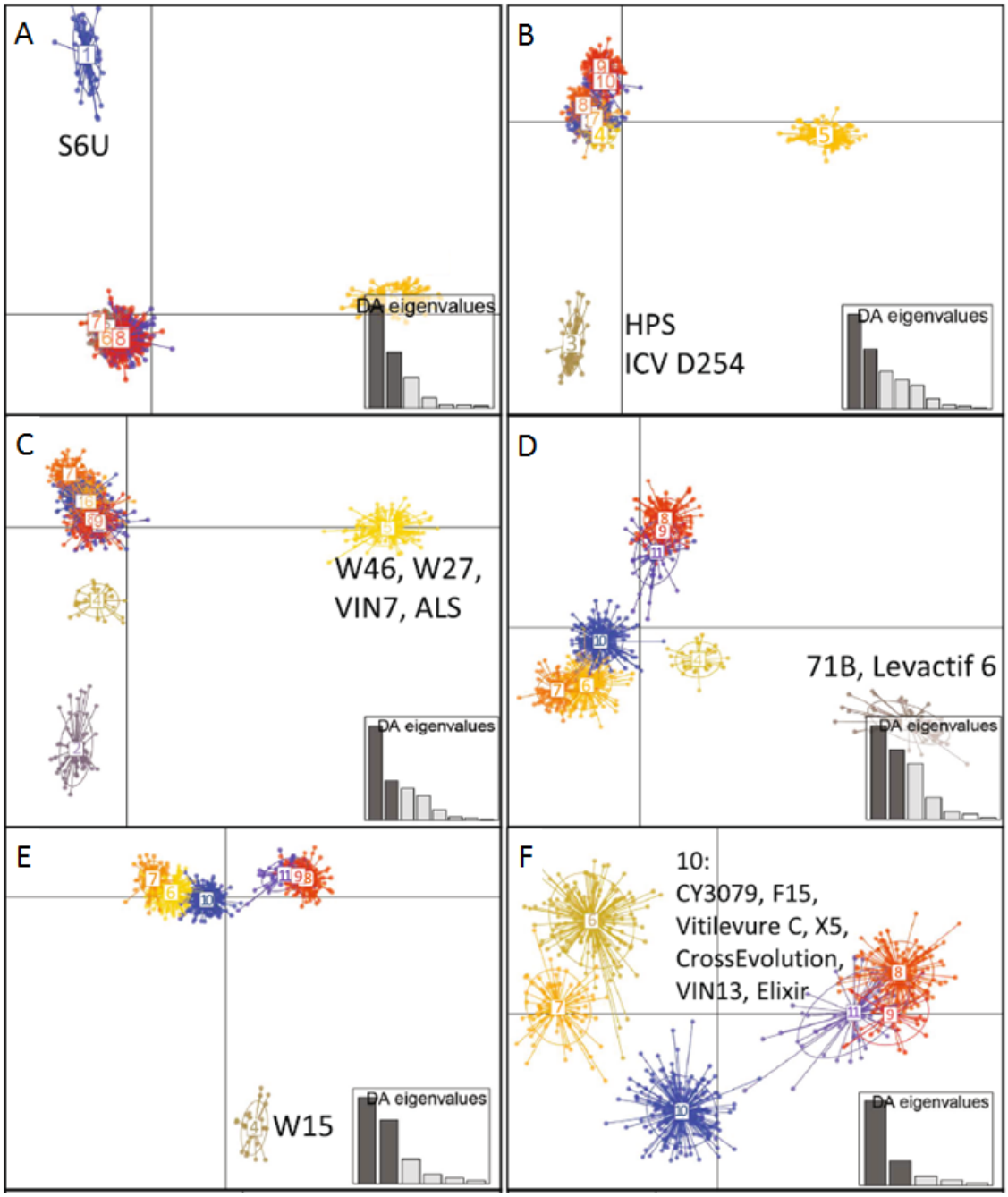

Fig. 3.28: DAPC analysis performed with the *adegenet* package for the R software displayed as a scatterplot, where the groups that could be clearly distinguished were stepwise eliminated. Names of the included strains in the groups are marked by in the graph prior to their elimination. (A) All 990 measurements organized in 11 groups (B) after

elimination of G1 (C) after elimination of G3 (D) after elimination of G5 CAN after elimination of G2 (F) after elimination of G4 (second part of the figure see Fig. 3.29)

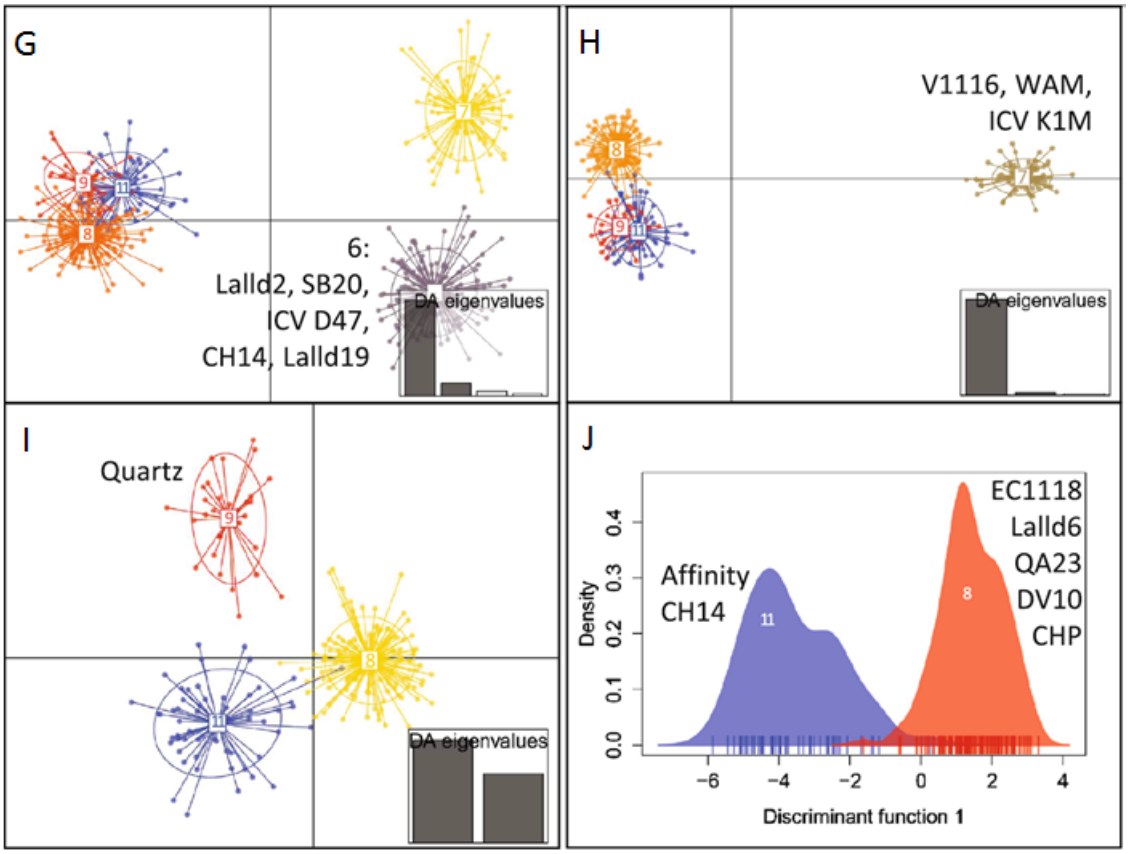

Fig. 3.29: Part 2 of Fig. 3.28: (G) after elimination of G10 (H) after elimination of G6 (I) after elimination of G7 (J) and after elimination of G9

3.6 Identification of biomarkers

A total of 12 MALDI-TOF MS spectra of the genome sequenced *S. c.* strain TMW 3.308 (S288C) after growth under aerobic and anaerobic conditions for 24 h was acquired. The metabolic activity was observed by HPLC and OD measurements displayed in Fig. 3.30. Under aerobic conditions, glucose was depleted at the time point where the spectra were acquired (24 h) and the diauxic shift took place where the yeast cells turned from fermentation to respiration using (oxidizing) the EtOH produced during fermentation as a carbon source.

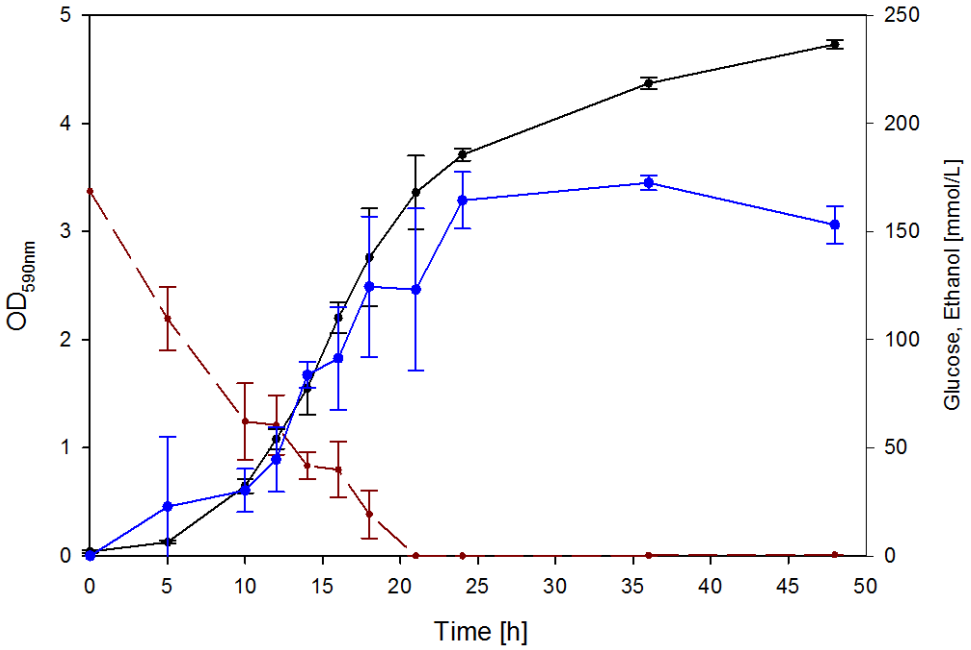

Fig. 3.30: Determination of optical density (–●–), glucose (- -●- -), and EtOH (–●–) contents in the supernatant of the *S. c.* culture (TMW 3.308) under aerobic conditions monitored for 48 h.

To identify proteins by de-novo sequencing via GeLC-MS/MS, masses detected in the first MS in combination with the fragmentation pattern in the second MS behind the collision cell can be analyzed by reconstructing amino acid sequence of peptides from overlapping fragments. Peptide sequences or peptide fragments were “blasted” (compared using the Basic Local Alignment Search Tool) against the published sequence for *S. c.* strain TMW 3.308 (S288C) available in the NCBI DB. Similarly, the summarized mass peaks generated by MALDI-TOF MS (cf. Fig. 3.31) were compared with this genome sequence to create a list of putative proteins.

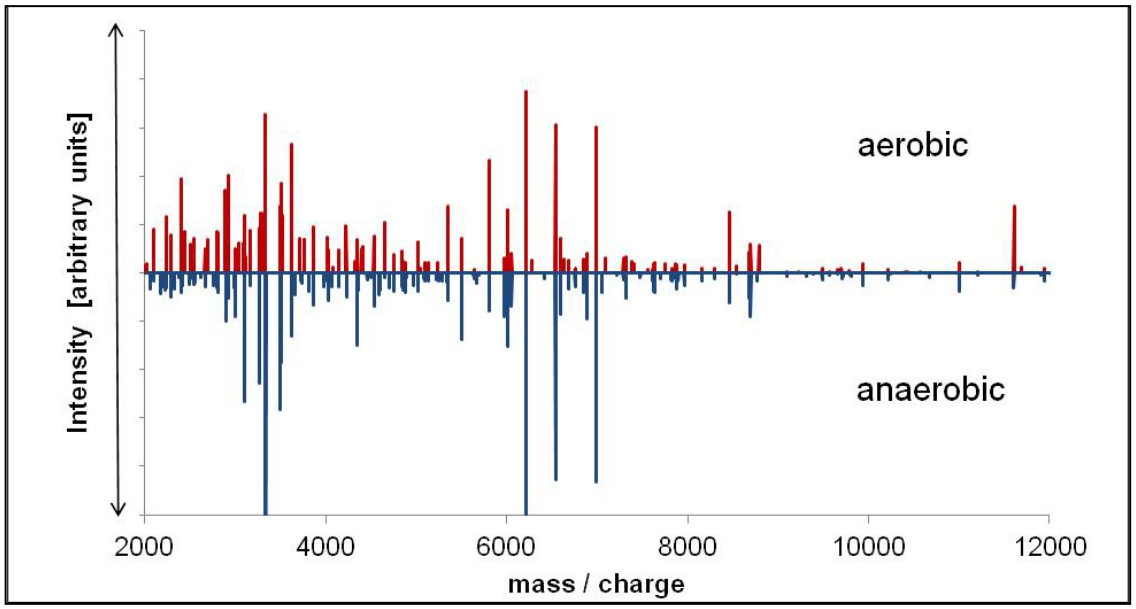

Fig. 3.31: Sum spectra of TMW 3.308 grown with (upper, red) and without (lower, blue) oxygen. Oxygen availability was determined correctly in 100 % (in-house software).

Some of the masses detected by MALDI-TOF MS or GeLC-MS/MS could not be assigned to proteins. Additionally, there were masses that could only be detected by one technique, where the GeLC-MS/MS provided four times more information than MALDI-TOF MS. Tab. 3.7 presents the computed peptides/proteins detected by MALDI-TOF MS that were also confirmed by GeLC-MS/MS and where an assignment to proteins was possible.

Tab. 3.7: List of masses detectable via both GeLC-MS/MS and MALDI-TOF MS including their identification by BLAST analysis. The different growth conditions of the yeast cells are indicated, namely aerobically (a) and anaerobically (aa). The extension _sp denotes signal peptides.

	m/z [Da]	**Kind of peptide** [SwissProt|Uniprot ID]	**Description**
a	3092.5	sp|Q12096|GNT1_sp	Glucose-N-acetyltransferase
	4859.2	sp|P00128|QCR7	Cytochrome b-c1 complex subunit
	6854.8	sp|P46990|RL17B	60S ribosomal protein
	8160.2	sp|Q12511|PP2C5_sp	Protein phosphatase 2C homologe
	10214.8	sp|O74700|TIM9	Mitochondrial import inner membrane translocase subunit
	11692.8	sp|P22943|HSP12	12 kDa heat shock protein, glucose and lipid-regulated protein
	11940.2	sp|P40479|SDP1	Dual-specificity protein phosphatase Sdp1
aa	2226.5	sp|Q04924|GLU2B_sp	Glucosidase 2 subunit beta
	2226.5	sp|P32334|MSB2_sp	Signaling mucin Msb2
	2674.0	sp|P53059|MNT2_sp	Alpha-1
	4867.0	sp|P40975|PMP2	Plasma membrane ATPase proteolipid 2
	5081.0	sp|P25594|VBA3_sp	Vacuolar basic amino acid transporter 3
	6855.0	sp|Q03096|EST3	Telomere replication protein Est3
	7323.0	sp|P39516|RS14B	40S ribosomal protein
	8764.2	sp|P38869|SVP26	Protein Svp26

The complete mass lists generated were mapped on an interactive pathway of *S. c.* by iPath explorer2 to visualize the protein function if possible. Fig. 3.32 displays the metabolic pathway map where peptide or protein fragments, which could be detected with both techniques are marked using thick lines. Detailed explanations are summarized in Tab. 3.8. Regulatory pathways are visualized in Fig. S6.

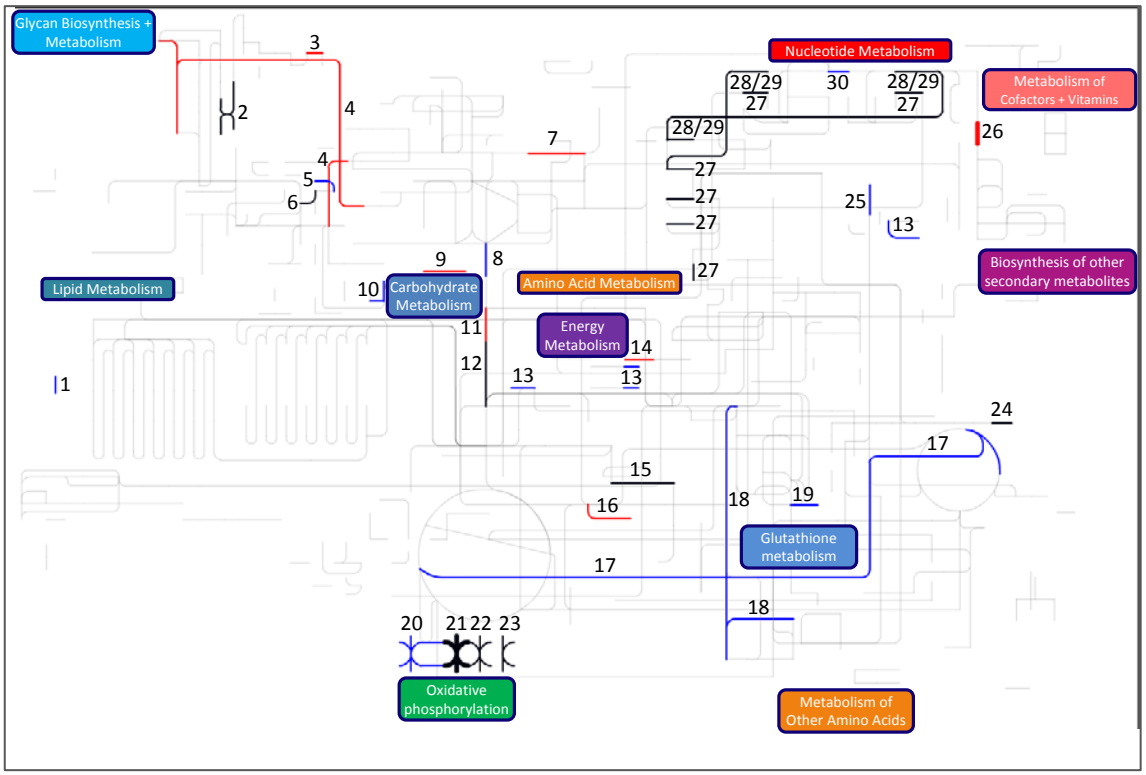

Fig. 3.32: Protein functions mapped on an interactive metabolic pathway of *S. c.* (by iPath explorer2). Displayed are the proteins or peptides belonging to masses detected via both MALDI-TOF MS and GeLC-MS/MS under aerobic conditions (—) and under anaerobic conditions (—). Masses that were detected independently of the growth conditions are also indicated (—). Thin lines (same color code) represent masses, which could be detected with only one technique. More details on proteins are listed in Tab. 3.8.

Tab. 3.8: Description of the proteins mapped in Fig. 3.32

	Name	Pathway	Definition
1	CYOE	Glycolysis / Gluconeogenesis	Alcohol dehydrogenase, propanol-preferring
2	E.2.4.1.119, QOXA	N-glycan biosynthesis	oligosaccharyl transferase complex subunit OST4, OST2
3	KTR6	High mannose type N-glycan biosynthesis	mannosylphosphate transferase
4	DPM2	GPI-anchor biosynthesis	dolichyl-phosphate mannosyltransferase polypeptide 2, regulatory subunit
5	GPP1	Glycerolipid metabolism	Glycerol-3-phosphatase 1
6	AGPAT1_2	Glycerolipid metabolism	lysophosphatidate acyltransferase
7	Pgl, devB	Pentose phosphate pathway	6-phosphogluconolactonase
8	GAPD	Glycolysis / Gluconeogenesis	Glyceraldehyde-3-phosphate dehydrogenase
9	glmU	Amino sugar and nucleotide sugar metabolism	UDP-N-acetylglucosamine/UDP-N-acetylgalactosamine diphosphorylase

	Name	Pathway	Definition
10	IMPA	Inositol phosphate metabolism	myo-inositol-1(or 4)-monophosphatase
11	gpmA	Glycolysis / Gluconeogenesis	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase
12	eno	Glycolysis / Gluconeogenesis	Phosphopyruvate hydratase ENO1
13	AdhP	Glycolysis / Gluconeogenesis	alcohol dehydrogenase, propanol-preferring
14	E3.1.3.15B	Histidine metabolism / Biosynthesis of amino acids	Histidinol-phosphatase (PHP family)
15	LYS2	Lysine biosynthesis	L-aminoadipate-semialdehyde dehydrogenase
16	asnB	Alanine, aspartate and glutamate metabolism	asparagine synthase (glutamine-hydrolysing)
17	argH	Alanine, aspartate and glutamate metabolism	Argininosuccinate lyase
18	DUG1	Glutathione metabolism	Cys-Gly metallodipeptidase DUG1
19	ahcY	Cysteine and methionine metabolism	adenosylhomocysteinase
20	NDUFAB1	Oxidative phosphorylation	NADH dehydrogenase (ubiquinone) 1 alpha/beta subcomplex 1
21	QCR7	Oxidative phosphorylation	ubiquinol-cytochrome c reductase subunit 7
22	COX1, 4, 6A, 6B, 7, 7C	Oxidative phosphorylation	Cytochrome c oxidase subunit
23	ATPeF1A, B, ATPeFD, G, H, J	Oxidative phosphorylation	F-type H^+-transporting ATPase subunit alpha, beta, d, g, h and j
24	coaE	Pantothenate and CoA biosynthesis	dephospho-CoA kinase
25	allB	Purine metabolism	Allantoinase
26	phoA	Folate biosynthesis	alkaline phosphatase
27	ndk	Purine metabolism	nucleoside-diphosphate kinase,
28	RBP6	Purine metabolism	DNA-directed RNA polymerases I, II, and III subunit RPABC2
29	RPC11, RPC25	Purine metabolism	DNA-directed RNA polymerase III subunit RPC10, RPC8
30	guaB	Purine metabolism	IMP dehydrogenase

4 DISCUSSION

This thesis demonstrates the potential of MALDI-TOF MS to differentiate yeasts at a previously not achievable level. As a consequence, yeasts can be assessed for their spoilage potential or predicted for their potential application in specific brewing and wine making processes, which has not been possible with any other previously used (molecular) techniques with so little effort.

This achievement is based on the optimization of sample preparation, evaluation of the impact of culturing conditions, enlargement of the proteome database, and development of new bioinformatics algorithms for data analyses and biomarker definition, which are described in the following.

4.1 Preparation method

In the beginning of this study, different strategies for sample preparation aiming at a rapid cell wall lysis were tested for their effects on the outcome of MALDI-TOF MS analysis (spectra acquired). For this purpose, cells of the beverage spoiling yeast strains *Saccharomyces cerevisiae* var. *diastaticus* (TMW 3.236), *Wickerhamomyces anomalus* (TMW 3.237), and *Debaryomyces hansenii* (TMW 3.238) were grown in/on two different media and treated according to the protocols specified in section 2.4.1 before they were prepared on MALDI-TOF MS target slides. Preparation methods included (D) direct transfer, (FA) on-target extraction using formic acid, (E) extraction, (EZ) extraction without centrifugation, (GB) glass beads, (FP) fast prep, (UB) ultrasonication bath, and (U) ultrasonication stick.

Formic acid, as a strong, reductive organic acid, improved the outcome of MALDI-TOF MS measurements (spectra), which becomes visible in the comparison of method D (direct application) with method FA (on-target extraction using formic acid) (see Tab. 3.1). It was previously reported that formic acid treatments are capable of destructing yeast cell walls (Amiri-Eliasi and Fenselau, 2001), and protein release form cells was reported in combination ACN (acetonitrile) as an aprotic polar solvent (Williams et al., 2003). Similar to the results obtained here, a positive effect of formic acid on the outcome of MALDI-TOF MS measurements was also reported previously by van Herendael et al. (2012) and Haigh et al. (2011).

Ethanol was used in the preparation of all samples except when methods D and FA were applied. The addition of alcohol to yeast cells is widely believed to lead to membrane

disruption, subsequent interference with metabolism, and cell lysis (McDonnell and Russell, 1999) by forming hydrogen bonds with proteins, which consequently lose their structure and function causing proteins and other macromolecules to precipitate (McDonnell, 2007). The increased efficiency of alcohols in the presence of water to provoke cell inactivation, is assumed to be related to higher concentrated solutions (>80 % alcohol) rapidly coagulating proteins on the cell surface, which prevents further penetration (McDonnell, 2007). A MALDI-TOF MS analysis of the EtOH supernatant showed that no leaching of the proteins out of the cells and into solution occurs (Williams et al., 2003). However, the use of alcohols for sample preparation has several advantages, i.e. alcohol inactivates cells (safety measure), reduces cell aggregation and facilitates the suspension of cells and cellular debris (Madonna et al., 2000), and, finally, improves mass signature quality (Qian et al., 2008; Williams et al., 2003). *In silico* comparison of the proteins' molecular weight (including probable modifications such as methylation, addition of water, or multiple charge) from *S. c.* TMW 3.308 (S288C) obtained by MALDI-TOF MS after EtOH extraction with molecular weights calculated on the basis of the published genome sequence of this strain enabled a tentative identification of approximately 40 peaks. These peaks can be predominantly assigned to ribosomal- and mitochondria-related proteins. The observation of these cytosolic organelle proteins indicates that the cell wall is effectively disrupted and it can be assumed that this is transferable to the beverage spoilage yeasts used to evaluate the effect of sample preparation methods on the outcome of MALDI-TOF MS measurements.

The addition of glass beads during the extraction (methods GB and FP) was employed to enhance the cell wall disruption mechanically without the addition of any further chemicals or enzymes (Wong et al., 2007), which would disturb the analysis by provoking interfering signals in MALDI-TOF MS spectra (Smirnov et al., 2004). For the same reason, ultrasonication (methods UB and U) were tested. However, on average, the methods E, FP, and GB were the three most successful methods to acquire MALDI-TOF MS spectra regardless of the cultivation medium used. Although the cases in which spectra were acquired successfully were high for any of these three methods, with respect of the quality of the acquired spectra, E compared with FP and GB led to a significant increase in total peak numbers. This was the reason why E was used as the standard preparation method for further experiments.

In addition to the extraction method, the choice of matrix is crucial to the success of MALDI- TOF MS experiments. In this study, HCCA (α-cyano-4-hydroxycinnamic acid) was employed as standard matrix, since derivatives of cinnamic acid are widely

recognized as good matrices for proteins (Marklein et al., 2009; Stevenson et al., 2010). This is due to several factors including its potential for strong optical absorption associated with the delocalized π-electron system of the conjugated double bonds (Zenobi and Knochenmuss, 1998), its acidic character encouraging ionization while acting as a proton source (Fitzgerald et al., 1993), and its capability of producing a more homogenous layer on the target than other acids, e.g. sinapinic acid (Williams et al., 2003). As a standard procedure in this study, the matrix was freshly prepared in 50 % ACN and 2.5 % (trifluoroacetic acid) (final HCCA concentration: 10 mg/mL). The use of 2.5 % TFA is recommended (Katz, 1954) to modify the pH of the matrix solvent increasing the spectra quality dramatically (i.e. resolution, number of peaks, intensity, SNR) (Williams et al., 2003). This effect is probably due to the property of TFA to work as a proton donator because of the distinct –I-effect of its three flour atoms, which might also enhance its potential to solve proteins. Furthermore, TFA evaporates very quickly at room temperature and ensures a proper crystallization of the matrix.

Additionally, the laser energy can have an effect on the acquired spectra, which was also tested in this study. Generally, an alteration in the laser energy is only possible within a small range. On one hand, a very low energy leads to an insufficient amount of laser power provoking small peak intensities. Consequently, the software does not recognize the peaks during the preprocessing step, and, therefore, the resulting spectrum, usually composed of/calculated from single 40 shots packages, cannot be detected. On the other hand, very high energy levels result in broadening of peaks. This results in a significant drop in resolution as well as a decrease in overall signal numbers detected caused by overlapping of single peaks and degrading the quality of spectra (Williams et al., 2003). Consequently, an optimized energy level is necessary, which generally is slightly above the desorption/ionization threshold to provide reproducible peaks. In this study, best results (total peak number, peak intensity, Fig. 3.2) were achieved when an increased initial laser energy was used (set to 50 - 65 % of the maximum set point possible; standard value for subsequent experiments).

4.2 Influence of growth phase

The identification at species level was never restricted but a tendency could be observed that younger cultures showed better results. This effect is probably due to the fact that an increasing age of a culture leads to a significant increase in the cell wall thickness and less porosity (de Nobel et al., 1990a; de Nobel et al., 1990b; Hiom et al., 1996; Zlotnik et al., 1984). The thickened cell wall has a modified mannoprotein structure (de

Nobel et al., 1990b) responsible for the resistance to proteolysis and reduction of disulfide bonds (Werner-Washburne et al., 1993), what might hamper the protein extraction. Additionally, ribosome synthesis is ceased when the culture is at <30 % of its maximum density (Ju and Warner, 1994). Subsequently, cells grow less rapidly and fewer ribosomes for protein production are required so that the cell starts to degrade them; a cell in stationary phase has finally <25 % of the ribosome complement of a cell in log phase (Warner, 1999). These facts are a likely explanation of the improved peak detection and species identification in younger cultures.

4.3 Different culturing conditions

To check whether the quality of the spectra is affected by environmental or physiologic parameters and to evaluate the influence on the resulting mass spectra, the three test strains were cultured aerobically and anaerobically, cell harvesting was conducted at different time points after inoculation (2.5 to 48 h), five different media were used (YM, YPG, Sab, W, ME) and different cell concentrations were applied on the target. To exclude the influence of the sample preparation, all experiments were conducted using the extraction (E) method (cf. 2.4.1.3).

Despite being cultured under different conditions, a significant number of ions is common to each yeast, what could be expected, if we act on the assumption that the mass spectra are dominated by ribosomal proteins, which ionize well and are only minimally influenced by microbial growth conditions (Wieser et al., 2011). However, the *S. c.* ribosome has 78 ribosomal proteins (Warner, 1999) but the number of peaks detected is obviously higher (cf. Fig. 3.2). A comparison of genome sequence and mass peaks of *S. c.* S288C reveals proteins with functions as membrane and cell wall proteins, stress response factors or ATP synthase subunits, which might be influenced by varying environmental conditions. Some groups investigated these influences, mainly focused on the limitation of the identification performance (Firacative et al., 2012; Goyer et al., 2012). In contrast to these approaches, we would like to determine if differences in growth conditions were reflected sufficiently consistent in the fingerprint to assign the spectrum to the appropriate culture conditions. For this purpose we have employed further software tools (ClinProTools 2.2, in-house software), which account for identification of small changes in the peak pattern instead of matching the similarities. Premise is a strict protocol, where only one parameter per experiment is varied to correlate reason and impact. This is not always possible, since many parameters are linked to each other e.g. during growth the pH decreases.

In default of the complete genome sequences of the spoilage species it is hard to allot any function to the proteins, clearly present or absent dependent on the different culturing condition. However, proteome analyses of *S. c.* have shown that many proteins, belonging to diverse – mainly mitochondrial – functional categories, could be identified with differential expression levels relating to oxygen availability (Bruckmann et al., 2009; de Groot et al., 2007), which is also reflected in the MALDI-TOF MS peak pattern.

Furthermore, it is widely reported that varying growth conditions significantly affect the glycoprotein composition of the external protein layer of the cell wall (Aguilar-Uscanga and Francois, 2003; de Groot et al., 2005), responsible e.g. for facilitated sterol uptake (Wilcox et al., 2002) or growth under anaerobic conditions (Abramova et al., 2001; Klis et al., 2006). Hence, dependent on the algorithm, the influence of available nutrients is visible (cf. Tab. 3.2). Advantage of the DB comparison is the possibility of an unlimited expansion of the dataset, whereas the modeling feature has to be calculated again if the dataset increases. The inhomogeneity of the most effective algorithms dependent on the used broth and species is likely due to the fact that the selected media are all recommended for the usage of yeast cultivation, therefore it can be supposed that a proper nutrient supply is achieved and prevents environmental stress stimuli, which might be better detectable.

Compared with its growth curve (for details see supplemental material, Fig. S1, XL), the sub-clusters of TMW 3.236 can be correlated to lag (2.5 to 7 h), exponential (10 to 12.5 h) and stationary (15 to 48 h) growth phase. The latter is functionally defined as the time when there is no further increase of cell number (Werner-Washburne et al., 1993) mostly caused by starvation (Lillie and Pringle, 1980). Many well reviewed responses were described, for instance, the storage carbohydrates glycogen and trehalose accumulate (Lillie and Pringle, 1980), resistance to various stresses like heat shock increases (Plesset et al., 1987), cell wall mannoproteins, encoded by cell cycle regulated genes (Bähler, 2005; Klis et al., 2006), which surround the glucan polysaccharide matrix (Cabib et al., 1982; Drebot et al., 1990), exhibit altered N-glycosylations (Valentin et al., 1987) and contain increased numbers of disulphide bridges (Werner-Washburne et al., 1993). With regard to these multidimensional changes it is not surprising that the protein pattern differs during a growth period, what is in accordance with Qian et al. (2008). For TMW 3.237 and 3.238 this kind of correlation is more imprecise, probably caused by the less distinctive growth behavior (for details see supplemental material, Fig. S1, XL). Nonetheless, the results are species and broth independent so that it can be concluded

that the growth phase affects the spectra insofar that it is possible to classify young, middle aged and older cultures.

Analyses of the varying concentration of the cell suspension solution showed that a minimal turbidity of a liquid culture of OD_{590nm} 0.5 is necessary to generate spectra reliably. The corresponding cell number determined by plate counting of *S. c.* for OD 0.5 is $5*10^6$/mL, thus – considering the solvent volume – a minimal supernatant according to $4.5*10^4$ cells/µL was required. However, the absolute number of peaks pointing out that there is no consistent trend observable. Mostly the results of OD_{590nm} 0.5 and 1.0 generate the highest peak amount and samples grown in YPG are more homogenous than cultures grown in YM. In default of exact preparation descriptions this is hardly comparable to other work groups Qian et al. (2008), nevertheless it can be assumed that a lower analyte/reagents ratio can enhance the degree of efficiency.

4.4 Database enlargement and validation

In total 190 isolates were identified by 26S rDNA sequence analysis and MALDI-TOF MS. Some limitations of the obtained results arise from the taxonomic resolution of the reference method. In particular, 9.5 % of the identified isolates referred to *S. p.,* which were not reliably differentiated from *S. c.* based on 26S rDNA sequencing. As 26S rDNA sequencing (Kurtzman and Robnett, 1998) cannot clearly differentiate *Saccharomyces* ssp., further methods like ITS- (internal transcribed spacer) (Kawahata et al., 2007) and IGS2- (intergenic spacer) sequencing or PCR-DHPLC (Hutzler et al., 2010) could have been employed. These methods significantly increase labor input and are usually not performed for routine diagnostic of beverage samples. However, we have shown in this study that MALDI-TOF MS analysis has the potential to differentiate the two species unambiguously (cf. 3.4) based on a dataset obtained from 37 brewing strains of both species. It is thus assumed that MALDI-TOF MS results are correct, even if further differentiation of these isolates using reference methods is lacking, so that in total 90.6 % of the isolates could be identified correctly. Likewise, *P. norvegensis,* a synonym for *C. norvegensis*, and *C. inconspicua* could not be separated by 26S rDNA sequencing. For further differentiation, the ITS region and 5.8S rRNA gene analysis should be employed (Nho et al., 1997). Therefore it remains unclear whether MALDI-TOF MS is able to differentiate these two species.

Another reason for a mismatch is the absence of corresponding DB entries, thus, no positive match could be achieved.

The mix up of *R. mucilaginosa* and *W. anomalus* can only be explained by a not sufficiently distinct peptide mass fingerprint to distinguish these two organisms properly. Nevertheless, the morphology of the colonies however allowed correct differentiation as *R. mucilaginosa* is colored red whereas *W. anomalus* is white. Additionally, the four samples with the wrong identification at genus level were also compared based on their colony morphology, which suggested that MALDI-TOF MS identification was correct, e.g. *P. membranifaciens* shows a typical slightly brown color, irregular mat colonies with crinkly or netlike surface and ridged border when they become older (Back, 1994a). We assume that the two laboratories isolated two different species out of the liquid culture, keeping in mind that the analyses were performed delayed in two different labs. Samples were collected and stored up to six weeks after 26S rDNA sequence analysis before MALDI-TOF MS experiments were done.

The failure to identify *W. anomalus* remains debatable, as this species proved to be easily measureable in 37 other cases (see Tab. 2.8). A possible explanation could be that it was no pure culture and instead of *W. anomalus*, an organism not included in the DB was isolated and analyzed.

4.5 Brewing yeast typing

4.5.1 Differentiation of top- and bottom-fermenting strains

To differentiate brewing strains according to their fermentation behavior, 30 spectra per strain were generated measured as triplicates on ten different days to simulate the real-life conditions as realistically as possible and furthermore one dataset with 24 measurements out of one extraction was generated recommended by the manufacturer. Finally, 42 different strains were tested and evaluated applying different software tools (MALDI Biotyper, MASCAP, DAPC) and varying approaches to generate database entries (out of biological replicates, out of one extraction).

The differentiation of top- and bottom-fermenting brewing strains works with all software approaches and experimental setups for at least 99.0 % (cf. Tab. 3.3). This concurs with the findings in chapter 4.4, where the identification of yeasts at species level is shown and with investigations in several studies in the last years (Agustini et al., 2014; Blättel et al., 2013; Marklein et al., 2009; Qian et al., 2008; van Veen et al., 2010; Welker, 2011).

Nonetheless, it should be highlighted that *S. p.* – formerly determined as a natural hybrid of *S. c.* and *S. bayanus* (Dunn and Sherlock, 2008; Nakao et al., 2009; Ogata et al.,

2008; Rainieri et al., 2006; Yamagishi and Ogata, 1999) and nowadays described as a hybrid of *S. c.* and *S. eubayanus* (Libkind et al., 2011) – supposed to bear a strong resemblance with *S. c.,* demonstrated by Kurtzman (1987) who studied the extent of ribosomal RNA sequence divergence among sibling yeast strains and found 58 % nDNA relatedness between *S. c.* and *S. p.* that might hampers the differentiation. In comparison: *S. c.* and *S. bayanus* show 10 % relatedness what is in accordance with Libkind et al. (2011) who proposed the origin of *S. bayanus* as a triple hybrid of *S. eubayanus* and *S. uvarum* that contains contributions from *S. c.* in at least some cases.

Independent of the evaluation method, the differentiation at biotype level works much better for the TF yeasts, even there are more groups defined. This phenomenon is most likely due to the fact that ale yeasts are genetically more diverse than lager yeasts (Lodolo et al., 2008). The genome associated to each strain is unique and – in addition to raw material and process conditions – ultimately defines the final aroma profile of the product (Pires et al., 2014; Ramos-Jeunehomme et al., 1991). The grouping according to the final application of yeasts enables a sorting, which takes into account that yeast strains are employed for special beer types depending on their metabolic spectrum and their physiological properties. It is already shown that the exometabolome and the metabolic footprinting are powerful tools for high-throughput yeast chemotaxonomy (Pope et al., 2007), substantiate the assertion that the distribution of the metabolic products are unique enough to differentiate at strain level. Basically, the origin of the different metabolites can be found partly in genetic differences, but can also be regulated at transcriptome or proteome level e.g. influenced by metabolic fluxes of substrate (Fiehn, 2001; Nielsen, 2003; Nielsen and Oliver, 2005; Rossell et al., 2005; ter Kuile and Westerhoff, 2001). Therefore, the cultivation conditions have a significant impact and may change the protein footprint detectable by MALDI-TOF MS technique.

This high sensitivity of the method, which reflects small changes of growing conditions can be used as a powerful tool for process monitoring where changes during fermentation conditions e.g. increasing ethanol concentration, decreasing sugar availability, osmotic pressure changes etc., could be depicted by peptide screening and contains helpful information for process control and improves the ability to respond rapidly and flexibly to constantly changing raw materials, which currently relies on the brewers experience and stable proceedings to provide a consistent quality and sensory characteristics of the end product.

4.5.1.1 Top fermenting strains

The TF strains differ in some certain properties: For wheat beer, the flavor described with phenolic or clove-like is appreciated, whereas in most other beers the decarboxylation of ferulic acid to 4-vinylguaiacol caused by a single *POF1* gene, is an undesirable note – known as phenolic off-flavor, with a very low threshold level, so consequently, Pof^+ strains are only applied for strong blond and wheat beers (Coghe et al., 2004).

The typical description of Altbier includes a low ester formation and a broad temperature range, with mild fruitiness if the fermentation process is conducted at higher temperatures. In some cases strains can be used for Kölsch or Altbier fermentation, so it is not surprising that the strain properties are very similar. Their flocculation behavior is mostly bad, so that a post-fermentation filtration is necessary. The level of diacetyl production is very low or not detectable (cf. http://www.blq-weihenstephan.de/hefezentrum/mikroorganismen.html). The vicinal diketones (VDK) diacetyl and 2,3-pentanedione are formed during fermentation as valine and isoleucine synthesis by-products with a buttery flavor, which is regarded as undesirable (Krogerus and Gibson, 2013a). The formation is a spontaneous decarboxylation of α-acetolactate resulting in diacetyl or acetoin, influenced by amino acid concentration (Krogerus and Gibson, 2013b), temperature and pH (Saerens et al., 2008b), pitching rate (Verbelen et al., 2009) or cell density (Verbelen et al., 2008). Subsequently, the diacetyl reduction process during maturation by the uptake into yeast cells is a valuable feature, which also depends on fermentation parameters influencing the yeast membrane composition (Boulton and Box, 2003). *ILV6* is shown to encode a small regulatory subunit of acetohydroxy acid synthase (AHAS) (Cullin et al., 1996), responsible for catalyzing the formation of α-acetolactate from pyruvate (Duggleby and Pang, 2000). Gibson et al. (2014) screened 14 lager brewing strains and found differences up to nine-fold in diacetyl concentrations in equivalent fermentation stages, so finally the whole process is strongly strain-dependent.

Ale or stout beer, originally brewed in Britain, is typical for its very dark color made of pale malt and 7-10 % roasted, unmalted barley, which offers the characteristic flavor due to Maillard products (Coghe et al., 2006) and often with addition of caramel malt or sugar, firstly introduced by Guinness as an extra stout with higher gravity (Rabin and Forget, 2014). In contrast to other beer types, ale is normally tapped with a mixture of CO_2 and N_2 to obtain a creamy whitecap (Burberg and Zarnkow, 2009). The flavor profile is described as malty, with a fruity final product depending on temperature process management.

4.5.1.2 Bottom fermenting strains - Saaz and Frohberg

Typical lager yeast characteristics appear to be due to a combination of traits inherited from both parents, followed by artificial evolution maybe depending on the environmental pressures to which the strain is exposed and/or hybrid genome instability (Gibson and Liti, 2014; Kunicka-Styczynska and Rajkowska, 2011).

The psychrophilic nature of *S. p.*, apparently inherited from the cold tolerant *S. eubayanus* parentage, may be the reason for its success as a fermentative organism in the low temperature environment of lager brewing (Gibson et al., 2013). However, in contrast to hybrids encountered in wine fermentation (Peris et al., 2012), the genetic diversity of *S. p.* is extremely low. The initial high level of diversity due to numerous hybridization events was restricted by yeast researchers such as Louis Pasteur and Emil Hansen in the 19th century, who realized that brewery hygiene and the use of pure cultures avoid beer spoilage (Barnett and Lichtenthaler, 2001; Gibson et al., 2013). The aneuploidy nature of the chromosomes is believed to account for the inability of lager strains to mate and their poor spore viability, leading to genetically isolated species (Jespersen et al., 1999; Morales and Dujon, 2012). Yet, they show increased fitness characteristics under certain environmental conditions such as the ability to grow at low temperatures (James et al., 2008; Querol and Bond, 2009). Dunn and Sherlock (2008) postulated two independent individual lineages, defined by genome rearrangements, copy number variations, ploidy differences and DNA sequence polymorphisms, referred to as Saaz and Frohberg after the locations in Germany and Bohemia in which they were originally used and which are still the same in modern breweries consisting more or less of derived variants (Gibson and Liti, 2014; Liti et al., 2005).

It is proposed that the ancestor of Saaz yeasts was an output of a cross between *S. c.* and *S. eubayanus*, losing a considerable portion of *S. c.* DNA, so finally proportionally more *S. eubayanus* DNA retained, inheriting the greater cold tolerance, however, the exact mechanisms are still unknown (Dunn and Sherlock, 2008; Querol and Bond, 2009; Walther et al., 2014). Yamagishi et al. (2010) perceived that the *S. eubayanus* form of *KEX2* gene, which encodes the kexin protease, supports the growth at low temperatures.

The ability to utilize maltotriose and consequently the fermentation performance is reported differently and is known to be strain-dependent. The study of Gibson et al. (2013) shows an inability to utilize maltotriose of Saaz and *S. eubayanus* strains and they suggest a lack of active maltotriose transport, which prevents the uptake, whereas Duval et al. (2010) identified some Saaz strains capable to utilize maltotriose during an

aerobic cultivation failing a clear response of the performance under fermentation conditions.

Maltotriose is the second most abundant fermentable sugar in wort (15-20 %) (Zastrow et al., 2001) and residual amounts in beer result in atypical flavor and economic problems because of lower EtOH amounts (Zheng et al., 1994). Therefore, the maltotriose fermentation is a desirable strain property, mainly dependent on the transporter profile encoded by *AGT1* (Han et al., 1995) and *MTT1/MTY1* (Dietvorst et al., 2005; Salema-Oom et al., 2005). Furthermore, a strong flocculation potential for Saaz strains is reported historically, a characteristic that is subject to change over time and depends on environmental factors during handling due to the subtelomeric position of the *FLO* genes controlling this property (Sato et al., 2002; Verstrepen et al., 2003b; Vidgren and Londesborough, 2011). The ester production of Saaz yeasts is comparatively small and does not exceed the taste threshold in the final product (Gibson et al., 2013).

In contrast, the Frohberg group, with its most prominent representative Weihenstephan 34/70 (TMW 3.275), the only genome sequenced lager yeast (Nakao et al., 2009), contains two genome equivalents of *S. c.* and one *S. eubayanus* equivalent (Dunn and Sherlock, 2008). The strains are deemed to be stronger fermenters going along with the ability to utilize maltotriose and the ester production is clearly higher (Gibson et al., 2013; Walther et al., 2014) whereas the flocculation potential is lower (Gibson and Liti, 2014).

The dataset of bottom fermenters, received by the BLQ Weihenstephan refers most probably to the Frohberg group (personal communication M. Hutzler) because of its strong fermentation traits, but a classification by Dunn and Sherlock is predominantly missing apart from TMW 3.275.

4.5.2 Performance of Biotyper 3.0 Software

Fig. 3.10 depicts technical replicates generated out of one extraction on one day (A) compared with spectra generated out of biological replicates on different days (B). It is not surprising that version (A) is more reproducible but on the other hand it is only a snapshot of one sample and does not take account for biological or technical daily variations. Until now, the aim of this technique was a reliable, fast identification at species level, which can be guaranteed using the most reproducible peaks for the DB and neglect the small differences emanate from culturing conditions, discussed extensively in chapter 3.2, technical variations, user experience and performance or even the humidity, which influences the drying time on the target and thus ultimately the crystallization

process. Nonetheless, to improve the sensitivity of the method, it is necessary to have a look on these small differences.

The comparability of BT1*24 and BT4*3 is limited, since BT4*3 is based on a modified dataset – missing the copies, adding new triplicates, but a tendency can be observed that those strains, which are more similar to others such as Altbier and Kölsch strains improve the identification performance if the DB is created out of biological replicates whereas those, which are more unique like the type strains or *S. c.* var. *diastaticus* identification changed to the loss. The uniqueness of the type strains may be indicative of their older hybrid origin (Fernández-Espinar et al., 2003) and *S. c.* var. *diastaticus* is only available twice. Discrepancies have been reported that type strains maintained in different culture collections are not identical (Rainieri et al., 2003) (Joubert et al., 2000) – CBS 1538^T (TMW 3.288) was found to be a hybrid and seem to have similar characteristics to *S. carlsbergensis* and lager brewing strains whereas NRRL-Y 1551 (TMW 3.289) does not seem to be a hybrid (Joubert et al., 2000), obtained by different proteomic pattern. These findings are confirmed by Casaregola et al. (2001), who displayed that NRRL-Y 1551 carries the *S. bayanus*-type *MET2* allele whereas CBS 1538^T has a typical *S. p. MET2* pattern. However, the differentiation by MALDI-TOF MS of these two type strains works well with the BT approach. Overall, the results are slightly better using BT4*3, with an improved strain level identification from 38.6 % to 45.3 %.

For *S. p.* the subdivision refers to their flocculation behavior, basically caused by their genetic setting of *FLO* genes and the environmental composition. One reason for the much worse correlation of the *S. p.* strains in Fig. 3.8 might be the imbalance of the number of strains of low (two) and high flocculation (ten). Consequently, the number of low flocculating strains was increased up to ten strains to avoid results predicted on probabilities.

The dendrogram in Fig. 3.12 gives an impression, how close the different *S. p.* strains are related, than the triplicates of one strain are only organized next to each other for 9 of the 20 strains. Interestingly, the triplicates of TMW 3.357, which seems to be very unique considering the heatmap of Fig. 3.13 and the results illustrated in Fig. 3.11, are split in two different cluster arms. Furthermore, the classification according to the flocculation potential is not visible as well. Corresponding values between TF and BF strains in the dendrogram (cf. Fig. 3.12) as well as in the spectrum comparison (cf. Fig. 3.11) can be found within TMW 3.358 and 3.277 with 77 % and TMW 3.351 and 3.284 with 23 %.

One explanation might be that the division of the BF strains in low and high flocculants, which is of course an important feature for brewers, is an attribute that is not consistent enough to consult it as the main trait, known to be largely influenced by propagation, storage, cropping and repitching (Sato et al., 2002; Verstrepen et al., 2003b; Verstrepen and Klis, 2006; Vidgren and Londesborough, 2011). In comparison with other studies (discussed in detail below, cf. 4.8), these problems are also reported (Scherer, 2002).

The BT software, programmed primarily to identify at species level for routine analysis, provides the opportunity for advanced users to create their own DB. The approach with four independent triplicates used as the basis of the DB entries, leads to completely different results (cf. Tab. 3.3, Fig. 3.9, Fig. 3.10) compared with the recommended procedure (BT1*24) with more or less strain-dependent improvement or worsening. The rates improve for Altbier and Kölsch strains, plus high and low flocculating strains but deteriorate for wheat and ale strains as well as for *S. c.* var. *diastaticus* and the type strains. Anyway, on the whole the results are better with BT4*3. Within the MSP creation method there are a couple of parameters for setting up the MSP, including peak frequency, maximal desired peak number, thresholds, intensity corrections or mass shifts, to tolerate the differences occurring because of performance deviations but regard the strain-dependent variations. Furthermore, the software provides the possibility to influence the algorithm slightly by subtyping function, which is an additional step after the usual MSP match taking into account that identical peaks in all spectra get a very small weight (even a zero) for the final score to focus on the discriminating small peaks within very similar spectra. The idea is plausible to firstly use a step to generate identification at species level by comparing the similarities of the DB and search spectra and afterwards eliminate the similarities and focus only on the differences within one similar group. The manufacturer promised a better separation for highly related species (Bruker Daltonik, 2008) but actually at strain level for *Saccharomyces* yeasts, this does not improve the results, maybe due to the fact that yeast spectra per se are not that precise compared with bacteria.

4.5.3 Performance of MASCAP approach

The same dataset was analyzed by in-house software, where the different DB entry calculations have clearly a higher impact on the results. The approach 4*3 achieves in nearly all subgroups higher matches with the only exception at biotype level of the low flocculating BF strains (cf. Tab. 3.3). Further increasing numbers of triplicate datasets did not improve the results significantly (data not shown). The 1*24 approach shows 10

DB entries, which completely are not matched by the corresponding search entry and mostly weak by other search entries, whereas the 4*3 version only displays this phenomenon for TMW 3.284, whereat this refers to other BF strains in 9.5 %, thus, the discrepancy between search and DB entry is obviously too high. The acquired spectra always present complex features, composed by a number of overlapping peaks with different amplitudes and contaminated by artifacts of biological or physical origin (Gras et al., 1999; Mantini et al., 2008) demonstrated in background noise and baseline drift. The processing algorithms separate protein signals from noise, which is an intractable problem, but can partly be solved if multiple iterations of spectra are summarized and reproducible peaks sustain higher ratings so that finally the probability of the appearance of a signal is taken into account, already displayed in several studies to improve the procedure of filtering relevant information from background signals (Coombes et al., 2005; Gras et al., 1999; Mantini et al., 2008; Mantini et al., 2007; Satten et al., 2004; Yasui et al., 2003). The 4*3 approach includes this consideration that makes it more precise in most cases. For TMW 3.258 the results of 1*24 are better as all wheat beer strains match with this DB entry, therefore it is a question of probabilities. For the strains TMW 3.339, 3.275, 3.353 and 3.354 the higher strain level identification rate achieved by MAS1*24 constitutes equal or less than five points, which is equivalent to approximately one measurement.

4.5.4 Comparison of Biotyper vs. MASCAP (1*24)

The differences in the algorithms used with the dataset of 1*24 become clear if the results of TMW 3.358 are compared (cf. Fig. 3.9, Fig. 3.14), where 29.7 % of all BF strain measurements correspond with the DB entry generated by BT, but no single hit is registered by MAS. Both versions are non-satisfying in this situation, the first with false positive results the latter with false negative ones. In total, the MASCAP strain level identification (34.1 %) is worse than the BT (38.6 %). It should be mentioned that the BT approach was conducted according to the recommendations of the manufacturer. In routine analysis the preparation of DB entries by users themselves might be unwanted, and tendencies can be observed in software updates that the user has more and more restricted possibilities to intervene the measurement – probably due to the fact that the most dominant application is clinical microbiology and the GMP requires authenticity of the data. Nevertheless, studies are published, where the DB creation is modified as well, e.g. 72 spectra of three independent measurements were used and the most

reproducible 24 ones were taken to create the DB entry (Moothoo-Padayachie et al., 2013).

4.5.5 Comparison of Biotyper vs. MASCAP (4*3)

The comparison of the two approaches executed with 4*3 DB entries, shows in total a higher hit rate at strain level by MASCAP with 51.5 % vs. 45.3 % by BT. Better results by BT could be achieved for Kölsch on strain and biotype level, and for *S. c.* var. *diastaticus* and type strains at strain level. Strain background, already discussed, displayed a high homology of Altbier and Kölsch strains in their field of application – clearly visible in Fig. 3.15: except of 10 % of the Altbier strain TMW 3.252, which is organized to wheat beer strain TMW 3.255, all the other wrong assigned measurements belong to Altbier and Kölsch in both directions.

To summarize the results of the similarity calculation it should be pronounced that the highest strain level hit could be achieved by in-house software using biological replicates for DB creation. The biotype sorting works for TF strains, but reliable strain identifications could not be achieved. The separation of *S. c.* and *S. c.* var. *diastaticus* works well by BT1*24 and 4*3 as well as by MAS4*3. The classification along the flocculation behavior of the BF strains leads to greater difficulties with the best results obtained by MAS4*3.

4.5.6 Alternative procedures

A new collection of all 20 bottom-fermenting strains was measured as triplicates on ten different days (n=600) again. The repetition of the measurements of TMW 3.275 to 3.284 is necessary for a maximized standardization so that a possible discrimination between high- and low-flocculating strains does not correlate with the date the measurement was carried out, maybe based on small aberrations on calibration, laser energy, different lots of chemicals or something else that might influence the acuteness of mass spectra. The results are displayed in Fig. 3.16, depicting that a clear classification along the flocculation potential is possible. Applying this system to the TF strains does not improve the results compared with single spectra (data not shown). The equal data processing employed for DB and search entries seems to impact the results positively. Further attempts to summarize 10 single spectra to one search entry do not improve the attribution, maybe due to the fact that the 10 spectra, which are generated originally as triplicates, are more diverse so that during the processing step the small relevant differences, which are not reproducible enough, are deleted and therefore restrict the

information content of the sum spectrum. The identification at strain level does not improve significantly over all 20 strains.

4.5.7 Performance of MDS

The Voronoi offers the beneficial possibility that differences between all members of the data set can be illustrated in distances on the 2D map, an objective method that provides more information than a DB comparison where an assignment of search and DB entry is mandatory with a yes/no answer but neglects the intensity of conformity.

The Voronoi diagram for TF strains based on two data points per strain, one created with MSP measurements (1*24) and the other marked with “t” out of 30 spectra measured on ten different days. Obviously, both procedures lead to slightly different results, since the data points are not always located next to each other such as TMW 3.250 or 3.262. Furthermore, a separation of wheat beer strains is clearly visible, but mixtures can be observed within Ale and *S. c.* var. *diastaticus* as well as Altbier and Kölsch strains. This is in accordance with the results obtained by DB comparison, where the MAS4*3 presents the one wrong assigned ale beer strain to *S. c.* var. *diastaticus* and mixed Altbier / Kölsch strains.

As discussed before (cf. 4.5.1.1) the occurrence of Pof^+ phenotypes among TF wheat beer strains and the enzymatic activity elucidate the volatile phenol concentration in wheat beer and strong blond beers, which is partly influenced by raw material and process management but mostly by yeast selection, displayed by Vanbeneden et al. (2008).

A clear separation along the flocculation potential of the BF strains can be observed by the approach where MSP measurements are employed (cf. Fig. 3.18-A). In contrast, the summed multiple measurements (cf. Fig. 3.18-B) seem to eliminate the differences between the strains, so finally the two groups overlap nearly completely. This is in accordance with the results of the TF strains, where both variants are displayed in one MDS, which shows the two data points of one strain partially in higher distances to each other. In fact, (A) is a snapshot and works in this case more or less randomly. A second dataset of MSP measurements provides worse results (data not shown), an evidence of the limited reproducibility of single MSP measurements.

4.5.8 Performance of DAPC

The DAPC grouping without any guidelines sorts the strains in five groups of different sizes (cf. Fig. 3.19, Tab. 3.4), where group 1 and 2 build distinct areas, group 3 to 5 overlap in the scatterplot. Only in three cases the approach organizes the complete spectra of one strain to one group. Additionally, group 4 is organized along the application of their members as wheat beer strains. This is a further indication of the uniqueness of wheat beer strains, already visible during the other approaches. Apart from TMW 3.253, the majority of the single measurements are sorted in group 4. Considering the hierarchical cluster analysis of Fig. 3.20 the two main cluster arms divide the dataset in group 1 and 2 for the lower arm (red and blue) and in group 3, 4 and 5 (green, cyan and magenta) in the upper part. In the cluster presentation Kölsch and Altbier strains of group 1 appear very close and the complete TMW 3.343 of group 2 is located separately in the lowest arm. Group 3 and 5 are mixed in the cluster as well, both groups include the strains TMW 3.251, 3.257, 3.259, 3.260, 3.261, 3.262, 3.273 and 3.274 of which all except 3.251 are solely represented in these two groups (cf. Tab. 3.4). Summarizing the above, it can be assumed that the cluster presentation is stretched to the limit in trying to visualize the complexity of the relationship of a dataset in this extent without any stipulated specific requirements.

Afterwards, a supervised DAPC was carried out, splitting the dataset in test and training. The concordance of test and training dataset indicates the validity of the system. DAPC analysis partitions the dataset into a between group and a within group component, the variation of the first has to be maximized while minimizing the second (Jombart et al., 2010). This works well for wheat beer, ale beer and *S. c.* var. *diastaticus,* which can be clearly separated in the scatterplot whereas Kölsch and Altbier are partly mixed.

Altogether, the separation of Kölsch and Altbier yeast strains represents a challenge for every used approach. It also suggests that Kölsch and Altbier strains have little difference on the level of sub-proteomes as determined by MALDI-TOF MS correlated with an obviously smaller impact on the final beer of these types than that one resulting from malt types and the brewing processes used in Cologne and Düsseldorf. The strain collection of the BLQ Weihenstephan, Freising, for instance, provides strain TUM 165 recommended for both beer types (http://www.blq-weihenstephan.de/fileadmin/user_upload/PDF/Mikroorganismen/TUM_165.pdf), merely depending on the type of wort.

Assessing the unsupervised sorting depicted in Fig. 3.20 and summarized in Tab. 3.4 it is shown that without any guidelines the spectra do not correlate necessarily with beer types. So on the one hand, the system is able to learn from a training dataset and to organize new spectra according to these learned characteristics, on the other hand, the unsupervised approach might be a chance to group strains along properties, which were not recognized so far. These could be valuable information to seek for new brewing strains or the apply known strains for another type of beer, simply to categorize them before time and resource consuming efforts were conducted for test series.

The DAPC analysis of the BF strains provides the separation depicted in Fig. 3.22 with a correct classification of the flocculation potential on 83 %. Further subtyping does not take place, because the flocculation is the only given property to split the dataset.

4.6 Wine yeast typing

In general, it is observed that the variability of the measurements of one yeast strain could differ greatly (cf. XLII, Fig. S4). The intra-group variance was shown to be strain-dependent, which made a consulting of biological replicates essential to achieve valid results instead of just a snapshot. This might be due to the fact that yeasts were distinctly influenced by various growth conditions (cf. 3.2) and therefore it was necessary to retain the culturing parameters very strictly.

4.6.1 Evaluation of the results

Two different approaches where applied for the initial grouping of strains. Firstly, a DB-based approach was optimized for strain level identification. Therefore, we calculated DB entries with 21 spectra from 7 biological replicates of all 33 strains and compared them with a test dataset made up of 9 single spectra. We have seen that the DB establishment based on biological replicates yielded a significant higher strain level hit rate (68.1 %) compared with an approach on multiple measurements (24-fold) of only one sample (51.1 %) (cf. Tab. 3.5). High second-hit similarities or wrong identifications were used for grouping and enabled a simple overview of the strains' relationship among each other.

Secondly, ClinProTools was used (Ketterlinus et al., 2005), a post-processing software, which generates classification models and validate them using different sophisticated mathematical and bioinformatic algorithms. Included are GA, which copies evolution in nature, SVM motivated from statistical learning theory, SNN and QC, a one dimensional

sorting algorithm using the p-values at certain peak positions for classification (Bruker Daltonik, 2007). For this purpose, the 15 strains were divided into six groups to create the basis for the modeling. Each additional strain of the validation dataset could now be assigned with all 30 measured single spectra to one of the six groups. Two detrimental points were that the choice of the used algorithm affected the results distinctly, already shown in chapter 3.2.2 and furthermore, strains got assigned mandatorily to the most similar group even if they were clearly different, so no new groups could be delineated.

According to these considerations, DB comparison was chosen as the favorable approach. A cluster analysis of all 33 strains might give an overview of all the data but was not meaningful enough due to the fact that the cluster data vary heavily even if small changes in the dataset exist. Additionally, the comparison between two strains is clearly visible but the next level of two separate clusters is either related to a hypothetical mixed calculation of the original data or corresponds to the more dominant arm. Both variations do not reliably illustrate the reality.

4.6.2 Grouping

We generally consider that the genomic approach by delta-PCR can provide an overview of the genetic background to elucidate the strains relationship among each other whereas MALDI-TOF MS analysis focuses at the proteome level, assuming a trivial link between protein expression and metabolism, which is in the end responsible for different application areas in wine fermentation. For sure this cannot be adopted one-to-one, demonstrated previously for certain sulphur pathways by Lafaye et al. (2005), where proteome and metabolic data could be correlated either positively or negatively depending on the growth conditions. Nevertheless, it emanates that similar fermentation properties of different strains can entail similarities in the peptide fingerprint. Consequently, we compared the known fermentation behavior of the used strains with their cluster results of the MALDI-TOF MS analysis.

Lalvin S6U displayed at genome and proteome levels conspicuous attributes that enabled a clear discrimination of this strain. It belongs genetically to *S. uvarum*, which is basically used for white and rosé wine and is known as a strong phenolic off-flavor (POF) producer. Volatile phenols such as 4-vinylphenol and 4-vinylguaiacol are formed by the action of cinnamate decarboxylase, which decarboxylates *p*-coumaric and ferulic acid, respectively (Chatonnet et al., 1993). The flavor is evaluated as medicinal, clove or leather (Shinohara et al., 2000). Compared with *S. c.* it differs in a number of properties, e.g. it is cryo-tolerant (Zambonelli et al., 1997), thus it occurs mainly in colder areas

(Demuyter et al., 2004; Naumov et al., 2001). It produces smaller amounts of undesirable acetic acid, which can cause a pungent, vinegar like taste, less amyl alcohols and higher amounts of glycerol, where sweetness and good mouth feeling are the main contribution to sensory characteristics (Noble and Bursick, 1984; Remize et al., 1999). Due to the increased glycerol formation, the alcohol yield is slightly reduced. An increased production of succinic acid (Shimazu and Watanabe, 1981), malic acid, isobutyl alcohol, isoamyl alcohol and numerous secondary compounds (Sipiczki, 2008) has also been reported as well as a higher aromatic intensity (Coloretti et al., 2006).

The *S. boulardii* strain Levucell SB20 appeared in a single arm of the delta-PCR cluster. MALDI-TOF MS analysis allowed identification at strain level as well, depending on the used software (cf. Tab. 3.5). Levucell SB20 is known as one of the rarely implemented eukaryotic probiotics, which is used as a preventive and therapeutic agent for diarrhea because of its temperature tolerance up to 37 °C and its resistance to acidic conditions (Czerucka et al., 2007), which clearly distinguishes this strain from the others used in this study.

Lalvin 71B and Siha levactif 6 arise in the delta-PCR cluster (cf. Fig. 3.24) as two distinct arms with less than 50 % similarity to the next clusters, whereas MALDI-TOF MS analysis groups Lalvin 71B and Siha levactif 6 in one cluster. Both are *S. c.* strains and literature describes Lalvin 71B as a red wine strain used for Beaujolais, originally isolated by the INRA (National Agricultural Research Institute) in Narbonne, France (Richter et al., 2013) and Siha levactif 6 as a distillery yeast. It has been reported previously that commercial wine yeasts are genotypically distinct from industrial yeast strains used for sake, beer or bread (Camarasa et al., 2011; Dunn et al., 2012; Legras et al., 2007).

All four *S. c. x S. kudriavzevii* hybrid strains can be organized in one cluster related to their MALDI-TOF MS fingerprint (cf. Fig. 3.26). In contrast, the delta-PCR clustering (cf. Fig. 3.24) divides them in three separate clusters A (Lalvin W46, Lalvin W27), B (Anchor VIN7), and C (Levuline ALS), which are less than 40 % similar. All of them are used for cold white wine fermentation such as Riesling (González et al., 2006) (Bradbury et al., 2006; González et al., 2007), due to their cryo-tolerance. Additionally, it has been found that several hybrid strains, including Anchor VIN7 (Borneman et al., 2012), have increased pectinase activity, regulated by the *PGU1* gene product, which consequently yields to a higher methanol level by the degradation of pectin (Richter et al., 2013), whereas most *S. c.* wine strains lack pectinase activity. In wine making this property is desirable since liquefaction of the grape must improves filterability and release aroma

precursors as well as the availability of carbon and nitrogen sources (Eschstruth and Divol, 2011).

Uvaferm HPS is known as a strain that produces a high amount of mannoproteins during alcoholic fermentation. It is used for red wine as well as Lalvin ICV D254, from which Uvaferm HPS was a subsequent selection. They form a homogenous group in the MDS according to their MALDI-TOF mass spectra (cf. Fig. 3.27). Lalld19, used for red wines, should also cluster in the red wine region, but it appeared in cluster M together with Lalvin ICV D47 (cf. Fig. 3.24), which is recommended originally for Chardonnay white wine isolated from Côtes du Rhône. This was unexpected since these pure *S. c.* strains are not known to be related since their recommended application and origin differ (Richter et al., 2013). Lalld2 and Lalvin ICV OKAY® are related by selective pressure from Lalvin V1116 that explains the closeness to each other in Fig. 3.27 according to the MALDI-TOF MS results and cluster N of the delta-PCR pattern in Fig. 3.24. In addition to this, WAM, described as a typical Sauvignon Blanc strain, got mixed with Chardonnay strain according to its delta-PCR pattern, whereas it is assigned to the Sauvignon Blanc region by MALDI-TOF MS analysis.

Clusters P to T (cf. Fig. 3.24) include the *S. c.* and *S. c. x S. bayanus* strains, which are characterized by a high ability to convert odourless aroma precursors present in the grapes into volatile thiols, thus they are good for varietal grapes such as Sauvignon Blanc. In this group CrossEvolution, Vitilevure Elixir and VIN13 are related by selective pressures and are all *S. c. x S. bayanus*, while Vitilevure C and Zymaflore X5 are pure *S. c.* strains, but are all used for Sauvignon Blanc. These five strains group perfectly together by MALDI-TOF MS. Furthermore, F15 is arranged next to the Sauvignon Blanc group in Fig. 3.27, which may be a helpful hint for its further applications.

The Chardonnay strains are well characterized as high producers of fermentative aroma like esters and higher alcohols. The MALDI-TOF MS cluster overlaps completely with Sauvignon Blanc, the Lalld19 strain, and a small corner of Champagne strains are involved as well.

Considering that the Champagne strains are all *S. bayanus* hybrids (*S. c. x S. uvarum)*, it seems obvious that they all are grouped in one region of the MDS (cf. Fig. 3.27). This is in accordance with the delta-PCR result, which displays a homogenous cluster U (cf. Fig. 3.24). $Affinity_{ECA5}$ and Lalld6 are selective pressures from Lalvin EC1118®, which complicates the differentiation at strain level. However, more accurate genetic techniques such as microsatellite analysis and PFGE enable to differentiate all strains from the Champagne group (Lallemand, personal communication). It has been

demonstrated that the Champagne strains possess characteristics of their parents in new and interesting combinations (Coloretti et al., 2006). They are cryo-tolerant, owing this property from the cold-fermenter *S. uvarum*, and produce higher quantities of glycerol, succinic acid, malic acid, 2-phenylethanol and lower amounts of acetic acid (Castellari et al., 1994). They display a superior fermentative capacity and ferment at low and high temperatures (Zambonelli et al., 1997). The *S. c.* genomic part is responsible for the flocculation behavior, which is important for sparkling wine that undergoes second fermentation in bottles. When traditional methods take place, the bottles are riddled until the yeast settles in the neck and gets removed. Furthermore the strains have to be pressure-resistant, because the bubbles remain in the bottle.

4.6.3 Grouping using DAPC scatterplots

The DAPC methodology provides for the problem that groups lacking prior runs are sometimes unknown or uncertain. DAPC uses sequential *k*-means to infer clusters of related individuals. In addition, the differences between groups were taken into account with a higher consideration to structure the abundance of data (Jombart et al., 2010). The chosen approach displayed in Fig. 3.28 and Fig. 3.29 can be interpreted as “zooming in”. At the beginning (cf. Fig. 3.28-A), the majority of all data is overlapping in the lower left corner except G1 and G5. If certain groups were excluded from the calculation, it is visible that the large grouping can be equalized, and single groups appear which are distinct enough to define them as one cluster. We want to confirm the compiled clustering by MDS with this approach.

The two techniques for the analysis arrive at a notable similar conclusion. Both build groups with *S. uvarum,* cold white wine strains *S. kudriavzevii,* red wine strains *S. c.,* bread strains from *S. c.* and separate cluster for Lalvin W15 cold white wine. They sort the strain Lalld19 to the Chardonnay group (cf. Fig. 3.29-G), the mix of Chardonnay and Sauvignon Blanc strains stays constant (cf. Fig. 3.28-F) and all Champagne strains can be isolated (cf. Fig. 3.29-I, -J). The difference was the addition of the probiotics Levucell SB20 to the Chardonnay group by DAPC, which might expectable when comparing the results with Fig. 3.26 where the SB20 strain is located next to the Chardonnay group. Nevertheless, the grouping by DAPC seems to be more objective then the manual analysis.

4.7 Identification of biomarkers

Proteins were extracted from the genome sequenced *S. cerevisiae* strain S288C cultivated aerobically and anaerobically followed by gel electrophoresis, LC-ESI-MS/MS sequencing, and a BLAST analysis (NCBI database) was conducted to identify peaks, which account for the observed possibility to precisely differentiate between cells grown in the presence or absence of oxygen.

The detected masses, aligned to proteins, are listed in Tab. 3.7. They summarize the results achieved by MALDI-TOF MS confirmed by GeLC-MS/MS. Gnt1, the signal peptide of glucose-N-acetyltransferase is known to be capable to modify N-linked glycans in the Golgi apparatus (Yoko-o et al., 2003). Cytochrome b-c1 complex subunit 7 (Qcr7) is directly linked to the availability of oxygen, expressed under aerobic conditions and known to be part of the electron transfer from ubiquinol to cytochrome c by a protonmotive Q cycle mechanism (Mitchell, 1976) in which electron transfer is linked to proton translocation across the inner mitochondrial membrane (Hunte et al., 2003). This is in accordance with de Groot et al. (2007), who published a relative protein $ratio_{anaerobic:aerobic}$ of 0.22 for Qcr7. The other associated proteins of the oxidative phosphorylation pathway (cf. Fig. 3.32) as (20) NADH-dehydrogenase (ubiquinone) subunit (accession no. P32463), (22) cytochrome c oxidase subunits, (23) F-type H^+-transporting ATPase subunit alpha, beta, d, g, h and j also appear under anaerobic conditions. This might be unexpected, but it has been reported that the mitochondrial proteome is kept in a steady-state even upon large metabolic changes. Thus, the majority of proteins involved in the respiratory metabolism are present also during fermentative growth (Ohlmeier et al., 2004). Tab. 4.1 illustrates how the datasets which are directly associated with the respiratory chain overlap according to different techniques and conditions.

Tab. 4.1: Peptides/proteins detected under different conditions and with different techniques (computed on the basis of the amino acid sequence), associated with oxidative phosphorylation (with a – aerobic, aa – anaerobic)

Protein	Description	MALDI a	GeLC a	MALDI aa	GeLC aa
P00128	Cytochrome b-c1 complex subunit 7	x	x		
P22289	Cytochrome b-c1 complex subunit 9			x	
P37299	Cytochrome b-c1 complex subunit 10				x
P00401	Cytochrome c oxidase polypeptide 1		x		
P00424	Cytochrome c oxidase polypeptide 5A		x		x
P00425	Cytochrome c oxidase polypeptide 5B		x		x
P32799	Cytochrome c oxidase polypeptide 6A		x		

Protein	Description	MALDIa	GeLCa	MALDIaa	GeLCaa
Q01519	Cytochrome c oxidase polypeptide 6B	x		x	
P07255	Cytochrome c oxidase polypeptide 7A		x		x
P04039	Cytochrome c oxidase polypeptide 8			x	
P39103	Cytochrome c oxidase assembly protein Cox14	x		x	
P47081	Cytochrome c oxidase assembly protein Cox16		x		

Only the 60S (accession no. P46990) and 40S (accession no. P39516) ribosomal proteins of Tab. 3.7 can be mapped on the pathway (cf. Fig. S6).

Protein phosphatase 2C homologe also known as pyruvate dehydrogenase – acetyl transferring – phosphatase 1 catalyzes the dephosphorylation and concomitant reactivation of the E1 alpha subunit (Pda1) of the pyruvate dehydrogenase complex (http://www.ncbi.nlm.nih.gov/protein/Q12511) (Krause-Buchholz et al., 2006). This means that the multi enzyme complex PDH, responsible for the irreversible conversion of pyruvate into acetyl-CoA (cf. Fig. 1.2), which comprises of three enzymes: pyruvate dehydrogenase (E1), dihydrolipoamide acetyltransferase (E2), and dihydrolipoamide dehydrogenase (E3), is regulated by phosphorylation by pyruvate dehydrogenase kinase at the α subunit of E1 (James et al., 1995). The reactivation by dephosphorylation is carried out by the found pyruvate dehydrogenase phosphatase (Uhlinger et al., 1986) so that pyruvate can be converted and finally acetyl-CoA can be channeled to the TCA cycle, typically for glucose-limited, aerobic growth (cf. 1.1.1.3 Pyruvate metabolisms). Mitochondrial proteins are mostly synthesized in the cytosol and have to be imported into mitochondria for biogenesis (Morgan and Lu, 2008) (Koehler, 2004). The mitochondrial import inner membrane translocase subunit Tim9 is discussed to be an essential protein of the mitochondrial inter membrane space (IMS), forming a hexameric complex with Tim10 and provides chaperone-like functions probably to prevent aggregation of hydrophobic membrane proteins in the IMS (Morgan and Lu, 2008; Spiller et al., 2015). A detailed overview is given by Neupert and Herrmann (2007) or by Pfanner and Wiedemann (2002). Early studies assumed that during anaerobiosis mitochondria completely disappear (Chapman and Bartley, 1968) (Wallace and Linnane, 1964) that was mainly attributable to incomplete staining and could later be rebutted by several groups (Damsky et al., 1969; Jenkins et al., 1984). Nowadays, it can be assumed that mitochondrial proteins decrease under anaerobic conditions, shown in comparative proteome studies conducted by Bruckmann et al. (2009). A series of studies have focused on 12 kDa heat shock protein Hsp12, which is described as a weakly expressed

protein during exponential phase and gets induced several 100-fold when entering the stationary growth phase and under stress (Praekelt and Meacock, 1990). It is upregulated upon a variety of stress, such as high concentrations of alcohol, oxidative and osmotic stress (Varela et al., 1995) or even aging (Welker et al., 2010). Thus, the appearance of this protein during aerobic growth is simply explicable and can be confirmed by de Groot et al. (2007).

The dual-specificity protein phosphatase Sdp1 (accession no. P40479), a stress-inducible mitogen-activated protein kinase (MAPK) phosphatase, mediates dephosphorylation of MAPK substrates such as Slt2, acquiring enhanced catalytic activity under oxidative conditions (Fox et al., 2007).

Under anaerobic conditions glucosidase 2 subunit beta (accession no. Q04924) could be identified, which is responsible for sequential cleavage of the two innermost α-1,3-linked glucose residues from the oligosaccharide precursor $Glc_2Man_9GlcNAc_2$ of immature glycoproteins localized at the endoplasmic reticulum (ER) (Moremen and Molinari, 2006; Wilkinson et al., 2006). Furthermore, the signaling mucin Msb2 (accession no. P32334), a plasma membrane protein, was found that promotes activation of the MAPK for the filamentous growth pathway (Cullen et al., 2004) and works as an osmosensing component of the high osmolarity glycerol response (HOG) pathway (O'Rourke and Herskowitz, 2002) and is stimulated by nutrient limitation-dependent induction of YPS1 expression (Vadaie et al., 2008). Alpha-1,3-mannosyltransferase Mnt2 (accession no. P53059) is involved in adding the 4^{th} and 5^{th} mannose residues of O-linked glycans (Romero et al., 1999) and plasma membrane ATPase proteolipid 2 (accession no. P40975) is a subunit associated with the 100 kDa subunit of the plasma membrane H^+-ATPase located at the cell membrane and regulates its activity (Navarre et al., 1994). Also a vacuolar basic amino acid transporter 3 Vba3 (accession no. P25594) could be detected, which is a permease required for vacuolar uptake of histidine and lysine (Shimazu et al., 2005). The telomere replication protein Est3 (accession no. Q03096) is a component of the telomerase complex involved in telomere replication (Diede and Gottschling, 1999) and finally the protein Svp26 (accession no. P38869) is an integral membrane protein and plays a role in retention of a subset of membrane proteins in the early Golgi compartments (Inadome et al., 2005).

It is eye-catching that the detected amino acid sequences and the out coming proteins isolated from cultures grown aerobically and harvested at the time point of the diauxic shift are predominantly associated to the mitochondria whereas the anaerobic grown culture displays unique proteins with various functions where none of it is associated to

mitochondria. It is observed by van den Brink et al. (2008) that most genes induced by anaerobiosis are related to cell wall and plasma membrane. Transcription analyses displayed significantly differing expression levels of up to approx. 500 genes during the comparison of aerobic and anaerobic growth conditions (Bruckmann et al., 2009; DeRisi et al., 1997; ter Linde et al., 1999), whereas the comparison of proteomic data yielded different results, suggesting a post-transcriptional regulation (de Groot et al., 2007). Nevertheless, round about 75 proteins change their levels more than two-fold when oxygen availability switches (Bruckmann et al., 2009). This is in accordance with these results and the observation that oxygen availability is 100 % reflected in the MALDI-TOF mass spectrum.

All other proteins, which are not significant according to growth conditions and detectable via MALDI-TOF MS or GeLC-MS/MS are presented in Fig. 3.32 and should be considered as an indication. Probably due to a higher resolution and split analyses, where the mass range from 5 to 12 kDa was separated in 5 different samples, the outcome by GeLC-MS/MS is higher than this of MALDI-TOF MS. Many enzymes involved in amino acid (cf. Fig. 3.32: 16, 17, 19) and purine nucleotide (cf. Fig. 3.32: 25, 30) metabolism were expressed at higher levels in the anaerobic culture, which is in accordance with previous studies by de Groot et al. (2007). This might be due to the fact that several of these proteins are localized in the mitochondria, which are strongly affected in function and morphology when oxygen is absent (Visser et al., 1994). Furthermore, the onset of ADH during anaerobiosis (cf. Fig. 3.32: 1, 13) is not surprising, as it catalyzes the conversion of acetaldehyde to EtOH during fermentation. Further enzymes of the glycolysis (cf. Fig. 3.32: 8, 11, 12) appear randomly. Tendencies can be observed that during aerobic conditions the biosynthesis pathways are preferred (cf. Fig. 3.32: 3, 4, 14, 26) whereas metabolism dominates the anaerobic cultivation (cf. Fig. 3.32: 5, 10, 17-19, 25, 30).

Folate biosynthesis is highly dependent on cultivation medium, growth rate and stage of growth. Repressible alkaline phosphatase catalyzes the reaction of 7,8-Dihydroneopterin 3'-triphosphate with water to 2-Amino-4-hydroxy-6-(D-erythro-1,2,3-trihydroxypropyl)-7,8-dihydropteridine and 3 o-phosphates, a precursor of H_4-folate (Hjortmo et al., 2008) (for details see http://www.genome.jp/128pp-bin/show_pathway?org_name=sce&mapno=00790&mapscale=&show_description=hide). It is reported that during the first hours in the respire-fermentative phase of growth, 5-CH_3-H_4-folate dominates with 100 µg/g dry matter, whereas H_4-folate reaches values of 20 µg/g dry matter, which is a much higher folate content compared with respiration

or stationary phases. A rapid drop of folate level was observed when glucose was depleted and the production certainly depends on the available amino acids. (Hjortmo et al., 2008)

In Fig. S6 the detected proteins belonging to the ribosomes are highlighted. It might seem curious that parts of the ribosomes depend on the culture conditions, even though they provide the basis for protein biosynthesis and therefore cell growth. Nonetheless, in rapidly growing cells, ribosomal protein genes account for 50 % of all polymerase II transcription-initiation events, this means taken a generation time of 100 minutes as a basis, the cell produces 2000 ribosomes per minute (Warner, 1999). The response to environmental stress conditions (Causton et al., 2001; Gasch et al., 2000; Ju and Warner, 1994) is coordinately regulated by adjustment of ribosome number and capacity regulations (Schawalder et al., 2004; Wade et al., 2004). Furthermore, 90 % of mRNA splicing is devoted to ribosomal proteins, which is also reflected in the protein mass spectrum, where proteins of the spliceosome are involved (Warner, 1999). This is in accordance with previous studies, where ribosomal proteins are published as the origin of detected peaks in MALDI-TOF MS technique (Wieser et al., 2011).

4.8 Comparison to other studies

Most of the published MALDI-TOF MS data focus on identification of bacteria (Seng et al., 2009), a smaller proportion is dealing with yeasts, probably due to their miner relevance in clinical microbiology (Wieser et al., 2011). For the sector food microbiology the application is currently rising but still shows clear gaps. Groups who work with yeasts report that elaborate extraction methods are mandatory to obtain appropriate spectra, which might be due to the thicker cell wall of yeasts compared with bacteria (Moothoo-Padayachie et al., 2013; Qian et al., 2008; van Veen et al., 2010; Wieme et al., 2014; Wieser et al., 2011), which is in accordance with the results of this study.

Most MALDI-TOF MS validation studies content oneself to generate hundreds of spectra from isolates previously identified by reference methods, compare them with the commercially available DB which results in a percentage rate of successful identifications to give an appraisal of the validity of the method (Agustini et al., 2014). Several problematic aspects should be highlighted: in many cases the tested isolates are arbitrarily put together, mostly dependent on the source of supply, which reflects on the one hand the frequency of single species according to this specialized field very well but on the other hand probable species specific difficulties were not taken into consideration

so that they might be over- or underrepresented and falsify the results. Furthermore, the success of a DB comparison is critically dependent on the quality of the DB entry (Pinto et al., 2011). If there are different extraction methods used which inevitably provoke small differences in the spectra, which are than compared with the identical DB entry, probably generated out of an alternative method, the significance of the result is debatable. Therefore, the approach to define the quality of a spectrum through the amount of generated peaks which includes an increased level of information seems to be more objective and is already used with bacterial samples (Mantini et al., 2007; Williams et al., 2003). Additionally, the far more complex metabolism of yeasts compared with bacteria complicates the accuracy of the method. For sure, most detectable peaks belong to ribosomal proteins (Amiri-Eliasi and Fenselau, 2001), which are less influenced by cultivation conditions (Stevenson et al., 2010; Wieser et al., 2011). Nevertheless the environmental or growth conditions result in volatility of the spectra so that reproducible measurements can only be achieved by a strict cultivation protocol. For instance, the same strain of *S. c.* grown for 7 h under equal conditions as the DB entry, cultivated for 24 h, cannot be identified correctly even at species level. These findings are also reported by other groups (Valentine et al., 2005). Even at highly reproducible growth conditions, measurements of the identical strain at different days yield to small differences – for *Saccharomyces* partly ranging in levels of strain differences. The results can be improved if DB entries are also generated out of measurements obtained on different days so that the variance of the measurement could be balanced somehow. Moothoo-Padayachie et al. (2013) applied this procedure to separate *Saccharomyces* strains successfully according to their application in food industry. Additionally, the success of the method depends on the applied bioinformatics. While the commercially available DB tool targets on reliable species identification and therefore levels the small nuances of the spectrum and furthermore, mainly due to the fact that MALDI-TOF MS technique is on the advance in routine microbiological laboratories where the exertion of influence of the operator on the mass spectra is increasingly undesired so that the options are more and more limited with every update, many groups employ alternative software tools. Wieme et al. (2014) used *BioNumerics* (Applied Maths) to differentiate spoilage and starter cultures in breweries in comparison to sequence analysis of D1/D2 regions of the 26 rRNA gene. Moothoo-Padayachie et al. (2013) used ClinProTools and PCA analysis to group strains along their application. Finally, but none the less relevant is the assigned task: the heterogeneity of the dataset plays an important role, while indeed, alternative techniques struggle with the same problem (Tornai-Lehoczki et al., 1996). Blättel et al. (2013) employed SAPD-PCR (specifically amplified polymorphic

DNA), based on the principal of RAPD-PCR, and MALDI-TOF MS to differentiate the genus *Saccharomyces*, focusing on the problem that a characterization of strains is hampered due to the potential of interspecific hybridization within the single members. In conclusion, both methods are powerful tools and able to differentiate most strains. Hutzler et al. (2010) employed PCR-DHPLC to differentiate *Saccharomyces* applied in brewing technology, facing with the challenge of retention time shifts within measurements of identical samples up to 0.5 min. Nonetheless, the differentiation works well and also mixtures of strains can be detected.

Wenning et al. (2002) published the rapid, inexpensive method of FT-IR (Fourier transform infrared) spectroscopy to identify yeasts, based on the comparison against a fingerprint DB. FT-IR spectra reflect all cell components according to different wavelength ranges, e.g. proteins and DNA/RNA absorb at 1800-1500 cm^{-1}. The complexity of the overlapping bands is subsequently simplified by the generation of the first or second derivative of the original spectrum. The success of the method is highly dependent on the quality of the DB and is furthermore influenced by cultivation conditions, which have to be strictly standardized – comparable to MALDI-TOF MS technique (Wenning, 2004). As there are micro colonies used for the analysis, the incubation time can be reduced to approx. 24 h.

Multiple approaches to differentiate industrially used brewing yeasts or yeasts relevant in beverage industry because of their spoilage potential are published nowadays, but the results are hardly comparable lacking the identical strain pool.

The gel electrophoresis patterns of the IGS2_314 rDNA amplificates published by Hutzler (2009) show for wheat beer strains TMW 3.250, 3.251, 3.255 and for ale yeasts TMW 3.260 and 3.261 equal amplificate sizes of 150-190 bp, in average shorter than those of the BF strains. The band width for the TF strains, and therefore the variability of the tandem repeats is strain-dependent. Altbier and Kölsch yeasts TMW 3.252 and 3.256 can be organized in the region of the BF strains. Furthermore, the BLQ Weihenstephan provides short strain characteristics, available on their homepage, including a cluster analysis depicting the relationship of a selection of strains based on interdelta PCR (δ12/δ21) and IGS2_314 rDNA fingerprints (http://www.blq-weihenstephan.de/fileadmin/user_upload/PDF/Mikroorganismen/TUM_68_PCR.pdf).

The differentiation results of the BF dataset analyzed by Hutzler et al. (2010) via PCR-DHPLC comprise a couple of strains used in our studies, amongst others TMW 3.275 - 3.286. Hutzler found very similar profiles for TMW 3.275 and 3.276, keeping in mind that TMW 3.275 is the ancestor of TMW 3.276, isolated from the same brewery

eight years before. With minor differences in fluorescence intensities, the profiles of TMW 3.279 and 3.280 were very similar to TMW 3.275 and 3.276. Interestingly, the DHPLC profiles of TMW 3.285 and 3.286 – both non-flocculating strains – show high conformity with the flocculation strain TMW 3.276 as well. Furthermore, the chromatograms of TMW 3.282 – 3.284 are almost identical, all derived from the same brewery, TMW 3.282 as the pure culture strain, TMW 3.283 as an isolate after the tank disinfection and TMW 3.284 as the production strain.

SUMMARY

Growth of spoilage yeasts in beverages results in reduced quality accompanied by economic and image losses. Therefore, biochemical and DNA-based identification methods have been developed, which are mostly time-consuming and laborious. Matrix-Assisted-Laser-Desorption/Ionization–Time-Of-Flight Mass Spectrometry (MALDI–TOF MS) could deliver discriminative peptide mass fingerprints within minutes. This makes it a rapid and reliable tool for identification and differentiation of microorganisms. However, at the beginning of this study, routine analysis of yeasts by MALDI–TOF MS was impaired by low reproducibility, limits in the level of differentiation, and effects of different physiological states of organisms on the reliability of the identification method were controversial.

In this study, the potential is demonstrated of MALDI-TOF MS to differentiate yeasts at a previously not achievable level. As a consequence, yeasts can be assessed for their spoilage potential or predicted for their potential application in specific brewing and wine making processes, which has not been possible with any other previously used (molecular) techniques with so little effort. This achievement is based on the optimization of sample preparation, evaluation of the impact of culturing conditions, enlargement of the proteome database, and development of new bioinformatics algorithms for data analyses and biomarker definition.

Firstly, sample preparation and measurement parameterization were optimized using three spoilage yeasts, i.e. *Saccharomyces cerevisiae* var. *diastaticus, Wickerhamomyces anomalus*, and *Debaryomyces hansenii*. Results demonstrate that MALDI-TOF MS presents an efficient, reliable tool for routine identification of spoilage yeasts, which appears to be only limited by the availability of suitable databases (DBs). For sample preparation, ethanol extraction is recommended, since it is easy to handle and enables a rapid determination with the highest number of reproducible peaks as evaluated by LIMPIC analysis. Other possible factors such as increasing the cell concentration prepared onto MALDI-TOF MS target slides above a minimum threshold level of $4.5*10^4$ cells/µL did not distinctly affect spectra, neither qualitatively nor quantitatively. Generally, an initial laser energy ranging from 50 to 65 % is necessary to ensure a successful measurement.

Furthermore, the influence of environmental parameters prevalent during the growth of yeast cells and their physiological state on the outcome of MALDI-TOF MS measurements was investigated. Parameters tested include nutrient availability, growth

phase as well as oxygen availability. Although core mass peaks remained constant under all tested conditions enabling reliable identification at species level, variations in the latter two parameters resulted in characteristic and reproducible differences in mass fingerprints. Accordingly, yeasts grown in the presence or absence of oxygen could be precisely differentiated along the specific mass fingerprints and a crude classification of the growth phase was possible. In contrast, the availability of the different nutrients tested result in small but not 100 % reproducible differences in mass spectra. Whereas software tools such as the Bruker Biotyper (BT) are suitable for a reliable identification of unknown samples at species level, the use of advanced modeling algorithms (e.g. as described in this study) can significantly increase the precision and reproducibility by which differences in mass spectra reflecting the history of yeast cells are detected. The fact that the sensitivity of the methods established in this study (sample extraction, target preparation, software analysis) makes it possible to retrace growth conditions, demonstrates the possibility that MALDI-TOF MS could be used for the detection of contamination sources, e.g. in brewing facilities.

Furthermore, it was tested whether the highly sensitive method established permits the differentiation of yeasts at sub-species level. For this purpose, different approaches were established to discriminate brewing strains along their biotype, i.e. fermentation behavior, industrial application, and flocculation potential. BT and in-house software were employed and the way reference database entries were created was varied, i.e. based on multiple measurements of either one single extraction (as recommended by Bruker) or different biological replicates (accounting for technical and biological variations). MASCAP analysis based on biological replicates appears to deliver best results for biotype discrimination and strain identification of top-fermenting strains. MDS and DAPC approaches are valuable tools to subtype or group strains along specific properties as demonstrated with the hardly differentiable bottom-fermenting strain pool.

In addition to brewing yeasts, the potential of MALDI-TOF MS to characterize and discriminate wine yeast strains was investigated and compared with delta-PCR, PFGE, and microsatellite approaches. Whereas delta-PCR patterns reflecting differences in the genotype are (mostly) suitable to differentiate wine yeasts at strain level, PFGE can provide useful information on chromosomic variations, and microsatellites can be used to detect small differences between strains, which are generated in response to an adaption to environmental changes. In comparison with these techniques, MALDI-TOF MS analysis indirectly also reflects the genetic background but detects differences at the proteome level, which is more suitable for a characterization of metabolic diversity and

makes it possible to differentiate wine yeast strains according to their industrial application.

Thus, the MALDI-TOF MS technique provides a powerful tool for a rapid and meaningful prediction of various industrial application fields of yeast strains. Although test fermentation processes can be probably not completely replaced by MALDI-TOF MS analysis, it could be used to reduce the number of initial exhaustive, time and money consuming test fermentation procedures with metabolite analysis for the selection of new candidate strains.

Finally, protein extraction from the genome sequenced *S. cerevisiae* strain S288C cultivated aerobically and anaerobically followed by gel electrophoresis, LC-ESI-MS/MS sequencing, and a BLAST analysis (NCBI database) was conducted to identify protein peaks, which account for the observed possibility to precisely differentiate between cells grown in the presence or absence of oxygen. Protein masses that were detected using both techniques belonged to proteins already described as being influenced by oxygen availability. This allowed for a deeper insight into factors influencing MALDI-TOF MS spectra and a more detailed characterization of the oxygen response of the tested organism. Furthermore, this demonstrates the potential of MALDI-TOF MS to be used as initial rapid screening tool to detect metabolic differences, which then can be further characterized by the described approach.

For future applications, it would be interesting to explore an approach to identify species out of mixtures without separation and enrichment of the culture, which would eliminate one of the major hurdles impeding the use of MALDI-TOF MS for the detection of contamination sources in food processing facilities. Furthermore, enhanced precision and reliability in the identification at strain level would make MALDI-TOF MS an excellent alternative to be used in quality control of fermentation processes and could replace or complement routine strain identification procedures in microbiology laboratories.

ZUSAMMENFASSUNG

Die Getränkeindustrie sieht sich derzeit mehr denn je mit der Herausforderung konfrontiert, ein innovatives Produktportfolio anzubieten, das sich durch Vielseitigkeit auszeichnet. Neben Trends wie z. B. physiologisch förderlichen Eigenschaften oder nachhaltiger Rohstoffgewinnung, sollten die Produkte konservierungsmittelfrei wenn nicht gar unbehandelt sein und ungekühlt möglichst lange lagerfähig. Diese Anforderung an die mikrobiologische Stabilität kann nur durch gute Prozesshygiene und qualitativ hochwertige Rohstoffe gewährleistet werden. Trotz der hohen Standards diesbezüglich, verursachen getränkeverderbende Hefen jährlich einen wirtschaftlichen Schaden in Millionenhöhe. Zur Vermeidung dieses mikrobiologischen Verderbs ist die Identifizierung des Schadorganismus unerlässlich um die Quelle der Kontamination einzugrenzen und die notwendigen Maßnahmen einleiten zu können. Hierzu wird momentan auf physiologische oder molekularbiologische Methoden zurückgegriffen, deren Durchführung jedoch kostspielig und zeitaufwändig ist, so dass sich die Entwicklung alternativer Methoden lohnt. MALDI-TOF Massenspektrometrie wird bereits in der klinischen Mikrobiologie zur schwerpunktmäßigen Identifizierung von Bakterien eingesetzt, ist jedoch in ihren Anwendungen für die Lebensmitteltechnologie im Allgemeinen und die Differenzierung lebensmittelrelevanter Hefen eingeschränkt.

In dieser Arbeit wurde das Potenzial von MALDI-TOF MS zur Differenzierung von Hefen in einer bisher nicht erreichten Tiefe gezeigt. Folglich können Hefen hinsichtlich ihres Schadpotenzials oder möglicher Anwendungen in der Bier- und Weinherstellung eingeschätzt werden. Die Basis hierfür lieferten die Optimierung der Probenpräparation, die Evaluierung des Einflusses der Kultivierungsbedingungen, die Erweiterung der Proteom-Datenbank, die Entwicklung neuer bioinformatischer Auswertealgorithmen für die Datenanalyse und die Definition von Biomarkern.

Zunächst wurden die Geräteeinstellungen sowie das Aufarbeitungsprotokoll zur Proteinextraktion anhand dreier Modellorganismen mittels einer LIMPIC-basierten Auswertung der Massenspektren für eine objektive Vergleichbarkeit optimiert. Hierbei wird das Massenspektrum prozessiert um es um das Grundrauschen und die Basislinienverschiebung zu bereinigen, so dass auch kleine Signale erkannt und berücksichtigt werden können. Unterschiedliche Kultivierungsbedingungen wie Sauerstoffverfügbarkeit, Kulturmedien und -dauer zeigten einen deutlichen Einfluss auf die Spektren, deren Reproduzierbarkeit so gut ist, dass zum einen ein striktes Protokoll genutzt werden muss um eine akkurate Identifizierung zu erhalten, zum anderen aber auch die Möglichkeit der Zuordnung der Wachstumsbedingung der Hefe gegeben ist.

Mit dem optimierten und standardisierten Protokoll wurde eine Datenbank mit Isolaten aus der Praxis etabliert, die als Referenzmethode mit klassischer PCR Analytik beim Forschungszentrum Weihenstephan für Brau- und Lebensmittelqualität identifiziert wurden.

Um die Empfindlichkeit der Methode und somit ihre Anwendbarkeit auf die Differenzierung von Stämmen zu untersuchen, wurden gut charakterisierte Kulturhefestämme für ober- und untergäriges Bier sowie für unterschiedliche Weine und Schaumwein herangezogen und vergleichend mit verschiedenen experimentellen als auch bioinformatischen Ansätzen ausgewertet. Hierzu wurden Test-Datensätze gegen Referenzspektren verglichen, die sowohl aus einer einzigen, mehrfach gemessenen Extraktion generiert wurden, als auch Referenzspektren aus mehreren biologischen Replikaten die an unterschiedlichen Tagen gemessen wurden. Zum Einsatz kamen u. a. die vom Hersteller bereitgestellte Software MALDI Biotyper, das auf MATLAB-Basis konzipierte MASCAP sowie das *adegenet* Paket der R-Software zur Konzeption von Diskriminanzanalysen der Hauptkomponenten (DAPC).

Der Bierhefe Datensatz konnte zu 100 % in ober- und untergärige Stämme differenziert werden. Weiterhin zeichnete sich für die Obergärigen eine Typisierung analog zu deren Applikation für bestimmte Biersorten ab. Weißbier, Ale/Stout und der Verderber *S. cerevisiae* var. *diastaticus* bilden eindeutige Gruppen, Altbier und Kölsch Stämme hingegen ließen sich mit keinem Ansatz zu 100 % separieren. Die besten Ergebnisse auf Stammebene erzielte der MASCAP Ansatz mit durchschnittlich 65 % korrekt identifizierten Stämmen. Die untergärigen Stämme, die in erster Linie vom Brauer in Bruch- und Staubhefen unterteilt werden, konnten mittels DAPC zu 83 % dem richtigen Flockulationsverhalten zugeordnet werden. Eine Identifizierung auf Stammebene war mittels MASCAP in rund einem Drittel der Fälle möglich.

Der Weinhefe Datensatz war inhomogener und bestand sowohl aus reinen *S. cerevisiae* Stämmen als auch aus unterschiedlichen Hybriden mit verschiedenen Applikationsschwerpunkten. Die Gruppierung entlang dieser entsprechenden Verwendung zeigte sich als weiter reichend als es die genetische Analyse ermöglichte. Die Identifizierung auf Stammebene wurde im Testdatensatz (15 Stämme) zu Zweidrittel erreicht.

Zur Identifizierung der Massenpeaks wurde die Genom sequenzierte Hefe *S. cerevisiae* S288C nach aerober und anaerober Kultivierung einer Standard-MALDI-TOF MS Extraktion unterzogen, der Extrakt mittels Gelelektrophorese aufgetrennt, anschließend aus dem Gel isoliert und mittels LC-ESI-MS/MS sequenziert. Die generierten Sequenzen

ergaben mit Hilfe der NCBI Datenbank eine Proteinidentifizierung, die parallel mit den Massen der MALDI-TOF MS Ergebnisse verglichen wurden. Einige der übereinstimmenden Massen konnten Proteinen zugeordnet werden, deren Vorkommen sauerstoffbedingt bereits beschrieben wurden. Das untermauert das Potential der MALDI-TOF MS Technik, einen Einblick in den metabolischen Zustand der Hefe zu ermöglichen.

Zusammenfassend kann festgehalten werden, dass die MALDI-TOF MS Technik zur Speziesidentifizierung für getränkerelevante Hefen geeignet ist und auch nah verwandte Hybride sicher identifiziert werden können. Eine Stammidentifizierung liefert je nach Datensatz Übereinstimmungen von bis zu 68 %. Für Brauereien ließe sich die Technik zur Überwachung der Anstellhefe mit Kontaminationen nutzen. Darüber hinaus ermöglicht die Methode eine Applikations-Vorhersage, die bei der Etablierung unbekannter Stämme erste Hinweise liefern kann.

Für die Zukunft wäre eine Identifizierung ohne vorherigen Anreicherungsschritt, wenn möglich aus Mischungen von Organismen eine zusätzliche Ersparnis an Zeit und Ressourcen, ließ sich im Rahmen dieses Projektes jedoch noch nicht zielführend umsetzen. Weitere Unternehmungen die Differenzierung auf Stammebene zu verbessern sind notwendig.

LITERATURE

Abramova N, Sertil O, Mehta S, Lowry CV (2001) Reciprocal regulation of anaerobic and aerobic cell wall mannoprotein gene expression in *Saccharomyces cerevisiae*. *J Bacteriol* 183(9):2881-7

Aguilar-Uscanga B, Francois JM (2003) A study of the yeast cell wall composition and structure in response to growth conditions and mode of cultivation. *Lett Appl Microbiol* 37(3):268-74

Agustini BC, Silva LP, Bloch C, Jr., Bonfim TM, da Silva GA (2014) Evaluation of MALDI-TOF mass spectrometry for identification of environmental yeasts and development of supplementary database. *Appl Microbiol Biotechnol* 98(12):5645-54

Amiri-Eliasi B, Fenselau C (2001) Characterization of Protein Biomarkers Desorbed by MALDI from Whole Fungal Cells. *Anal Chem* 73:5228-31

Antonelli A, Castellari L, Zambonelli C, Carnacini A (1999) Yeast influence on volatile composition of wines. *J Agric Food Chem* 47(3):1139-44

Arias CR, Burns JK, Friedrich LM, Goodrich RM, Parish ME (2002) Yeast species associated with orange juice: evaluation of different identification methods. *Appl Environ Microbiol* 68(4):1955-61

Back W (1987) Nachweis und Identifizierung von Fremdhefen in der Brauerei. *Brauwelt*:735-7

Back W (1994) Secondary contaminations in the filling area. *Brauwelt Int* 4:326-33

Bähler J (2005) Cell-cycle control of gene expression in budding and fission yeast. *Annual review of genetics* 39:69-94

Barber M, Bordoli R, Sedgwick R, Tyler A (1981) Fast atom bombardment of solids as an ion source in mass spectrometry. *Nature* 293(5830):270-5

Barber M, Green BN, Jennings KR (1987) The analysis of small proteins in the molecular weight range 10–24 kDa by magnetic sector mass spectrometry. *Rapid Commun Mass Spectrom* 1(5):80-3

Barnett JA, Lichtenthaler FW (2001) A history of research on yeasts 3: Emil Fischer, Eduard Buchner and their contemporaries, 1880–1900. *Yeast* 18(4):363-88

Beavis R, Chaudhary T, Chait B (1992) α-Cyano-4-hydroxycinnamic acid as a matrix for matrixassisted laser desorption mass spectromtry. *Org Mass Spectrom* 27(2):156-8

Beavis RC, Bridson JN (1993) Epitaxial protein inclusion in sinapic acid crystals. *J Phys D Appl Phys* 26(3):442-7

Beavis RC, Chait BT (1989) Factors affecting the ultraviolet laser desorption of proteins. *Rapid Commun Mass Spectrom* 3(7):233-7

Beavis RC, Chait BT, Fales H (1989) Cinnamic acid derivatives as matrices for ultraviolet laser desorption mass spectrometry of proteins. *Rapid Commun Mass Spectrom* 3(12):432-5

Behr J, Israel L, Gänzle MG, Vogel RF (2007) Proteomic approach for characterization of hop-inducible proteins in *Lactobacillus brevis*. *Appl Environ Microbiol* 73(10):3300-6

Bellon JR, Eglinton JM, Siebert TE, Pollnitz AP, Rose L, de Barros Lopes M, Chambers PJ (2011) Newly generated interspecific wine yeast hybrids introduce flavour and aroma diversity to wines. *Appl Microbiol Biotechnol* 91(3):603-12

Beltran G, Torija MJ, Novo M, Ferrer N, Poblet M, Guillamon JM, Rozes N, Mas A (2002) Analysis of yeast populations during alcoholic fermentation: a six year follow-up study. *Syst Appl Microbiol* 25(2):287-93

Berkenkamp S, Menzel C, Karas M, Hillenkamp F (1997) Performance of Infrared Matrix-assisted Laser Desorption/Ionization Mass Spectrometry with lasers emitting in the 3 µm wavelength range. *Rapid Commun Mass Spectrom* 11(13):1399-406

Biemann K, Martin SA (1987) Mass spectrometric determination of the amino acid sequence of peptides and proteins. *Mass Spectrom Rev* 6(1):1-75

Bisson LF, Coons DM, Kruckeberg AL, Lewis DA (1993) Yeast sugar transporters. *Crit Rev Biochem Mol Biol* 28(4):259-308

Blättel V, Petri A, Rabenstein A, Kuever J, König H (2013) Differentiation of species of the genus Saccharomyces using biomolecular fingerprinting methods. *Appl Microbiol Biotechnol* 97(10):4597-606

Borneman AR, Desany BA, Riches D, Affourtit JP, Forgan AH, Pretorius IS, Egholm M, Chambers PJ (2012) The genome sequence of the wine yeast VIN7 reveals an allotriploid hybrid genome with *Saccharomyces cerevisiae* and *Saccharomyces kudriavzevii* origins. *FEMS Yeast Res* 12(1):88-96

Boulton CA, Box WG (2003) Formation and disappearance of diacetyl during lager fermentation, in: Brewing yeast fermentation performance. Smart KA (ed), Blackwell Science, Oxford, pp. 183-95

Bradbury JE, Richards KD, Niederer HA, Lee SA, Rod Dunbar P, Gardner RC (2006) A homozygous diploid subset of commercial wine yeast strains. *Anton Leeuw* 89(1):27-37

Brown RS, Lennon JJ (1995) Mass resolution improvement by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer. *Anal Chem* 67(13):1998-2003

Bruckmann A, Hensbergen PJ, Balog CI, Deelder AM, Brandt R, Snoek IS, Steensma HY, van Heusden GP (2009) Proteome analysis of aerobically and anaerobically grown *Saccharomyces cerevisiae* cells. *J Proteomics* 71(6):662-9

Brugnoni LI, Cubitto MA, Lozano JE (2012) *Candida krusei* development on turbulent flow regimes: Biofilm formation and efficiency of cleaning and disinfection program. *J Food Eng* 111(4):546-52

Bruker Daltonik GmbH (2007) ClinProTools user manual. Version 2.2. Bremen, Germany

Bruker Daltonik GmbH (2008) Handbook of MALDI Biotyper 2.0 Software for Microorganism Identification and Classification. Bremen, Germany

Burberg F, Zarnkow M (2009) Special Production Methods, in: Handbook of Brewing. Wiley-VCH Verlag GmbH & Co. KGaA, pp. 235-56

Cabib E, Roberts R, Bowers B (1982) Synthesis of the yeast cell wall and its regulation. *Annual review of biochemistry* 51:763-93

Camarasa C, Sanchez I, Brial P, Bigey F, Dequin S (2011) Phenotypic landscape of *Saccharomyces cerevisiae* during wine fermentation: evidence for origin-dependent metabolic traits. *PLoS ONE* 6(9):e25147

Campbell I, Msongo HS (1991) Growth of aerobic wild yeasts. *J Inst Brew* 97(4):279-82

Casaregola S, Nguyen HV, Lapathitis G, Kotyk A, Gaillardin C (2001) Analysis of the constitution of the beer yeast genome by PCR, sequencing and subtelomeric sequence hybridization. *Int J Syst Evol Microbiol* 51(4):1607-18

Cassagne C, Cella AL, Suchon P, Normand AC, Ranque S, Piarroux R (2013) Evaluation of four pretreatment procedures for MALDI-TOF MS yeast identification in the routine clinical laboratory. *Med Mycol* 51(4):371-7

Castellari L, Ferruzzi M, Magrini A, Giudici P, Passarelli P, Zambonelli C (1994) Unbalanced wine fermentation by cryotolerant vs. non-cryotolerant *Saccharomyces* strains. *Vitis* 33(1):49-52

Causton HC, Ren B, Koh SS, Harbison CT, Kanin E, Jennings EG, Lee TI, True HL, Lander ES, Young RA (2001) Remodeling of yeast genome expression in response to environmental changes. *Mol Biol Cell* 12(2):323-37

Chapman C, Bartley W (1968) The kinetics of enzyme changes in yeast under conditions that cause the loss of mitochondria. *Biochem J* 107:455-65

Chatonnet P, Dubourdie D, Boidron Jn, Pons M (1992) The origin of ethylphenols in wines. *J Sci Food Agric* 60(2):165-78

Chatonnet P, Dubourdieu D, Boidron J-N, Lavigne V (1993) Synthesis of volatile phenols by *Saccharomyces cerevisiae* in wines. *J Sci Food Agric* 62(2):191-202

Coetzee C, du Toit WJ (2012) A comprehensive review on Sauvignon blanc aroma with a focus on certain positive volatile thiols. *Food Res Int* 45(1):287-98

Coghe S, Benoot K, Delvaux F, Vanderhaegen B, Delvaux FR (2004) Ferulic acid release and 4-vinylguaiacol formation during brewing and fermentation: Indications for feruloyl esterase activity in *Saccharomyces cerevisiae*. *J Agric Food Chem* 52(3):602-8

Coghe S, Gheeraert B, Michiels A, Delvaux FR (2006) Development of Maillard reaction related characteristics during malt roasting. *J Inst Brew* 112(2):148-56

Cohen SL, Chait BT (1996) Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Anal Chem* 68(1):31-7

Colby SM, King TB, Reilly JP, Lubman D (1994) Improving the resolution of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry by exploiting the correlation between ion position and velocity. *Rapid Commun Mass Spectrom* 8(11):865-8

Coloretti F, Zambonelli C, Tini V (2006) Characterization of flocculent *Saccharomyces* interspecific hybrids for the production of sparkling wines. *Food Microbiol* 23(7):672-6

Coombes KR, Tsavachidis S, Morris JS, Baggerly KA, Hung MC, Kuerer HM (2005) Improved peak detection and quantification of mass spectrometry data acquired from surface-enhanced laser desorption and ionization by denoising spectra with the undecimated discrete wavelet transform. *Proteomics* 5(16):4107-17

Cornish TJ, Cotter RJ (1993) A curved-field reflectron for improved energy focusing of product ions in time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 7(11):1037-40

Cullen PJ, Sabbagh W, Graham E, Irick MM, van Olden EK, Neal C, Delrow J, Bardwell L, Sprague GF (2004) A signaling mucin at the head of the Cdc42- and MAPK-dependent filamentous growth pathway in yeast. *Genes & Development* 18(14):1695-708

Cullin C, Baudin-Baillieu A, Guillemet E, Ozier-Kalogeropoulos O (1996) Functional analysis of YCL09C: evidence for a role as the regulatory subunit of acetolactate synthase. *Yeast* 12(15):1511-8

Currie LA (1999) Detection and quantification limits: origins and historical overview. *Anal Chim Acta* 391:127-34

Czerucka D, Piche T, Rampal P (2007) Review article: yeast as probiotics – *Saccharomyces boulardii*. *Aliment Pharmacol Ther* 26(6):767-78

Damsky CH, Nelson WM, Claude A (1969) Mitochondria in anaerobically-grown, lipid-limited brewer's yeast. *The Journal of cell biology* 43(1):174-9

de Deken RH (1966) The Crabtree effect: a regulatory system in yeast. *J Gen Microbiol* 44(2):149-56

de Groot MJL, Daran-Lapujade P, van Breukelen B, Knijnenburg TA, de Hulster EAF, Reinders MJT, Pronk JT, Heck AJR, Slijper M (2007) Quantitative proteomics and transcriptomics of anaerobic and aerobic yeast cultures reveals post-transcriptional regulation of key cellular processes. *Microbiol* 153(11):3864-78

de Groot PWJ, Ram AF, Klis FM (2005) Features and functions of covalently linked proteins in fungal cell walls. *Fungal Genet Biol* 42(8):657-75

de Nobel JG, Klis FM, Munnik T, Priem J, van den Ende H (1990a) An assay of relative cell wall porosity in *Saccharomyces cerevisiae, Kluyveromyces lactis* and *Schizosaccharomyces pombe*. *Yeast* 6(6):483-90

de Nobel JG, Klis FM, Priem J, Munnik T, van den Ende H (1990b) The glucanase-soluble mannoproteins limit cell wall porosity in *Saccharomyces cerevisiae*. *Yeast* 6(6):491-9

Deak T, Beuchat LR (1996) Handbook of Food Spoilage Yeasts, 1st edn. CRC Press, London

Demuyter C, Lollier M, Legras JL, Le Jeune C (2004) Predominance of *Saccharomyces uvarum* during spontaneous alcoholic fermentation, for three consecutive years, in an Alsatian winery. *J Appl Microbiol* 97(6):1140-8

Dengis PB, Nelissen LR, Rouxhet PG (1995) Mechanisms of yeast flocculation: comparison of top- and bottom-fermenting strains. *Appl Environ Microbiol* 61(2):718-28

Dengis PB, Rouxhet PG (1997a) Flocculation mechanism of top- and bottom-fermenting brewing yeast. *J Inst Brew* 103(4):257-61

Dengis PB, Rouxhet PG (1997b) Surface properties of top- and bottom-fermenting yeast. *Yeast* 13(10):931-43

Denis CL, Ferguson J, Young E (1983) mRNA levels for the fermentative alcohol dehydrogenase of *Saccharomyces cerevisiae* decrease upon growth on a nonfermentable carbon source. *J Biol Chem* 258(2):1165-71

Denis CL, Young ET (1983) Isolation and characterization of the positive regulatory gene ADR1 from *Saccharomyces cerevisiae*. *Mol Cell Biol* 3(3):360-70

DeRisi JL, Iyer VR, Brown PO (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278(5338):680-6

Dhiman N, Hall L, Wohlfiel SL, Buckwalter SP, Wengenack NL (2011) Performance and cost analysis of matrix-assisted laser desorption ionization-time of flight mass spectrometry for routine identification of yeast. *J Clin Microbiol* 49(4):1614-6

Diamandis EP (2004) Mass spectrometry as a diagnostic and a cancer biomarker discovery tool: opportunities and potential limitations. *Mol Cell Proteomics* 3(4):367-78

Dickinson JR, Kruckeberg AL (2006) Carbohydrate metabolism, in: Yeasts in Food and Beverages. Querol A, Fleet GH (eds), Vol 2, Springer, pp. 215-42

Diede SJ, Gottschling DE (1999) Telomerase-Mediated Telomere Addition In Vivo Requires DNA Primase and DNA Polymerases α and δ. *Cell* 99(7):723-33

Dietvorst J, Londesborough J, Steensma H (2005) Maltotriose utilization in lager yeast strains: MTT1 encodes a maltotriose transporter. *Yeast* 22(10):775-88

Domin MA, Welham KJ, Ashton DS (1999) The effect of solvent and matrix combinations on the analysis of bacteria by matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 13(4):222-6

Donalies UE, Stahl U (2002) Increasing sulphite formation in *Saccharomyces cerevisiae* by overexpression of MET14 and SSU1. *Yeast* 19(6):475-84

Drebot MA, Barnes CA, Singer RA, Johnston GC (1990) Genetic assessment of stationary phase for cells of the yeast *Saccharomyces cerevisiae*. *J Bacteriol* 172(7):3584-9

Dreisewerd K (2003) The desorption process in MALDI. *Chem Rev* 103(2):395-426

Dreisewerd K, Berkenkamp S, Leisner A, Rohlfing A, Menzel C (2003) Fundamentals of matrix-assisted laser desorption/ionization mass spectrometry with pulsed infrared lasers. *Int J Mass Spectrom* 226(1):189-209

Duggleby RG, Pang SS (2000) Acetohydroxyacid synthase. *J Biochem Mol Biol* 33(1):1-36

Dunn B, Richter C, Kvitek DJ, Pugh T, Sherlock G (2012) Analysis of the *Saccharomyces cerevisiae* pan-genome reveals a pool of copy number variants distributed in diverse yeast strains from differing industrial environments. *Genome Res* 22(5):908-24

Dunn B, Sherlock G (2008) Reconstruction of the genome origins and evolution of the hybrid lager yeast *Saccharomyces pastorianus*. *Genome Res* 18(10):1610-23

Duval E, Alves Jr S, Dunn B, Sherlock G, Stambuk B (2010) Microarray karyotyping of maltose-fermenting *Saccharomyces* yeasts with differing maltotriose utilization profiles reveals copy number variation in genes involved in maltose and maltotriose utilization. *J Appl Microbiol* 109(1):248-59

Easterling ML, Colangelo CM, Scott RA, Amster IJ (1998) Monitoring protein expression in whole bacterial cells with MALDI time-of-flight mass spectrometry. *Anal Chem* 70(13):2704-9

Ehring H, Karas, M., Hillenkamp, F. (1992) Role of Photoionization and Photochemistry in Ionization Processes of Organic Molecules and Relevance for Matrix-assisted Laser Desorption Ionization Mass Spectrometry. *Org Mass Spectrom* 27:472-80

Eidtmann A, Gromus J, Bellmer H (1998) Mikrobiologische Qualitätssicherung: Der spezifische Nachweis von bierschädlichen Hefen im hefefreien Bereich. *Monatsschrift für Brauwissenschaft* 51(9-10):141-8

Eschstruth A, Divol B (2011) Comparative characterization of endo-polygalacturonase (Pgu1) from *Saccharomyces cerevisiae* and *Saccharomyces paradoxus* under winemaking conditions. *Appl Microbiol Biotechnol* 91(3):623-34

Esteve-Zarzoso B, Fernandez-Espinar MT, Querol A (2004) Authentication and identification of *Saccharomyces cerevisiae* 'flor' yeast races involved in sherry ageing. *Anton Leeuw* 85(2):151-8

Esteve-Zarzoso B, Hierro N, Mas A, Guillamón JM (2010) A new simplified AFLP method for wine yeast strain typing. *LWT - Food Sci Technol* 43(10):1480-4

Esteve-Zarzoso B, Peris-Toran MJ, Garcia-Maiquez E, Uruburu F, Querol A (2001) Yeast population dynamics during the fermentation and biological aging of sherry wines. *Appl Environ Microbiol* 67(5):2056-61

Fagerquist CK, Miller WG, Harden LA, Bates AH, Vensel WH, Wang G, Mandrell RE (2005) Genomic and proteomic identification of a DNA-binding protein used in the "fingerprinting" of *campylobacter* species and strains by MALDI-TOF-MS protein biomarker analysis. *Anal Chem* 77(15):4897-907

Feigl P, Schueler B, Hillenkamp F (1983) LAMMA 1000, a new instrument for bulk microprobe mass analysis by pulsed laser irradiation. *International Journal of Mass Spectrometry and Ion Physics* 47(0):15-8

Fenselau C, Demirev PA (2001) Characterization of intact microorganisms by MALDI mass spectrometry. *Mass Spectrom Rev* 20(4):157-71

Fernández-Espinar MT, Barrio E, Querol A (2003) Analysis of the genetic variability in the species of the *Saccharomyces* sensu stricto complex. *Yeast* 20(14):1213-26

Fiddler W, Parker WE, Wasserman AE, Doerr RC (1967) Thermal decomposition of ferulic acid. *J Agric Food Chem* 15(5):757-61

Fiehn O (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comp Funct Genomics* 2(3):155-68

Firacative C, Trilles L, Meyer W (2012) MALDI-TOF MS enables the rapid identification of the major molecular types within the *Cryptococcus neoformans/C. gattii* species complex. *PLoS One* 7(5):e37566

Fitzgerald MC, Parr GR, Smith LM (1993) Basic matrices for the matrix-assisted laser desorption/ionization mass spectrometry of proteins and oligonucleotides. *Anal Chem* 65(22):3204-11

Fleet GH (2006) The commercial and community significance of yeasts in food and beverage production in: Yeasts in food and beverages. Querol A, Fleet GH (eds), Springer Verlag, Berlin, pp. 1-12

Fleet GH (2007) Yeasts in foods and beverages: impact on product quality and safety. *Curr Opin Biotechnol* 18(2):170-5

Flikweert MT, van der Zanden L, Janssen W, Steensma HY, van Dijken JP, Pronk JT (1996) Pyruvate decarboxylase: an indispensable enzyme for growth of *Saccharomyces cerevisiae* on glucose. *Yeast* 12(3):247-57

Fox G, Shafiq M, Briggs D, Knowles P, Collister M, Didmon M, Makrantoni V, Dickinson R, Hanrahan S, Totty N (2007) Redox-mediated substrate recognition by Sdp1 defines a new group of tyrosine phosphatases. *Nature* 447(7143):487-92

Fraley C, Raftery AE (1998) How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis. *The Computer Journal* 41(8):578-88

Frank A, Pevzner P (2005) PepNovo: de novo peptide sequencing via probabilistic network modeling. *Anal Chem* 77(4):964-73

Frank A, Tanner S, Bafna V, Pevzner P (2005) Peptide sequence tags for fast database search in mass-spectrometry. *J Proteome Res* 4(4):1287-95

Frank AM (2009) A Ranking-Based Scoring Function for Peptide− Spectrum Matches. *J Proteome Res* 8(5):2241-52

Frank AM, Savitski MM, Nielsen ML, Zubarev RA, Pevzner PA (2006) De Novo Peptide Sequencing and Identification with Precision Mass Spectrometry. *J Proteome Res* 6(1):114-23

Fushiki T, Fujisawa H, Eguchi S (2006) Identification of biomarkers from mass spectrometry data using a "common" peak approach. *BMC Bioinformatics* 7:358

Gabriel E, Fagg GE, Bosilca G, Angskun T, Dongarra JJ, Squyres JM, Sahay V, Kambadur P, Barrett B, Lumsdaine A, Castain RH, Daniel DJ, Graham RL, Woodall TS Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation. In: Proceedings, 11th European PVM/MPI Users' Group Meeting, Budapest, Hungary, 2004.

Gailiusis J, Rinne RW, Benedict C (1964) Pyruvate—Oxaloacetate exchange reaction in baker's yeast. *BBA-Specialized Section on Enzymological Subjects* 92(3):595-601

Gancedo JM (1998) Yeast carbon catabolite repression. *Microbiol Mol Biol Rev* 62(2):334-61

Gasch AP, Spellman PT, Kao CM, Carmel-Harel O, Eisen MB, Storz G, Botstein D, Brown PO (2000) Genomic expression programs in the response of yeast cells to environmental changes. *Mol Biol Cell* 11(12):4241-57

Gibson B, Krogerus K, Ekberg J, Monroux A, Mattinen L, Rautio J, Vidgren V (2014) Variation in α-acetolactate production within the hybrid lager yeast group *Saccharomyces pastorianus* and affirmation of the central role of the ILV6 gene. *Yeast* 32(1):301-16

Gibson B, Liti G (2014) *Saccharomyces pastorianus*: genomic insights inspiring innovation for industry. *Yeast* 32(1):17-27

Gibson BR, Lawrence SJ, Leclaire JP, Powell CD, Smart KA (2007) Yeast responses to stresses associated with industrial brewery handling. *FEMS Microbiol Rev* 31(5):535-69

Gibson BR, Storgårds E, Krogerus K, Vidgren V (2013) Comparative physiology and fermentation performance of Saaz and Frohberg lager yeast strains and the parental species *Saccharomyces eubayanus*. *Yeast* 30(7):255-66

Giebel RA, Fredenberg W, Sandrin TR (2008) Characterization of environmental isolates of Enterococcus spp. by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Water Res* 42(4-5):931-40

Gimon M, Preston L, Solouki T, White M, Russell D (1992) Are proton transfer reactions of excited states involved in UV laser desorption ionization? *Org Mass Spectrom* 27(7):827-30

Glückmann M, Pfenninger A, Krüger R, Thierolt M, Karas M, Horneffer V, Hillenkamp F, Strupat K (2001) Mechanisms in MALDI analysis: surface interaction or incorporation of analytes? *Int J Mass Spectrom* 210:121-32

González SS, Barrio E, Gafner J, Querol A (2006) Natural hybrids from *Saccharomyces cerevisiae, Saccharomyces bayanus* and *Saccharomyces kudriavzevii* in wine fermentations. *FEMS Yeast Res* 6(8):1221-34

González SS, Gallo L, Climent MA, Barrio E, Querol A (2007) Enological characterization of natural hybrids from *Saccharomyces cerevisiae* and *S. kudriavzevii*. *Int J Food Microbiol* 116(1):11-8

Goyer M, Lucchi G, Ducoroy P, Vagner O, Bonnin A, Dalle F (2012) Optimization of the preanalytical steps of matrix-assisted laser desorption ionization-time of flight mass spectrometry identification provides a flexible and efficient tool for identification of clinical yeast isolates in medical laboratories. *J Clin Microbiol* 50(9):3066-8

Gras R, Muller M, Gasteiger E, Gay S, Binz PA, Bienvenut W, Hoogland C, Sanchez JC, Bairoch A, Hochstrasser DF, Appel RD (1999) Improving protein identification from peptide mass fingerprinting through a parameterized multi-level scoring algorithm and an optimized peak detection. *Electrophoresis* 20(18):3535-50

Haigh J, Degun A, Eydmann M, Millar M, Wilks M (2011) Improved performance of bacterium and yeast identification by a commercial matrix-assisted laser desorption ionization-time of flight mass spectrometry system in the clinical microbiology laboratory. *J Clin Microbiol* 49(9):3441

Haikara A, Helander I (2006) *Pectinatus, Megasphaera* and *Zymophilus*, in: The Prokaryotes. Vol 4, Springer Verlag, pp. 965-81

Haikara A, Penttilä L, Enari T-M, Lounatmaa K (1981) Microbiological, biochemical, and electron microscopic characterization of a Pectinatus strain. *Appl Environ Microbiol* 41(2):511-7

Han EK, Cotty F, Sottas C, Jiang H, Michels CA (1995) Characterization of AGT1 encoding a general α-glucoside transporter from *Saccharomyces*. *Mol Microbiol* 17(6):1093-107

Hardwick WA (1995) Handbook of brewing. M. Dekker, New York

Hillenkamp F, Karas M (2000) Matrix-assisted laser desorption/ionisation, an experience. *Int J Mass Spectrom* 200(1):71-7

Hiom SJ, Furr JR, Russell AD, Hann AC (1996) The possible role of yeast cell walls in modifying cellular response to chlorhexidine diacetate. *Cytobios* 86(345):123-35

Hiralal L, Olaniran AO, Pillay B (2014) Aroma-active ester profile of ale beer produced under different fermentation and nutritional conditions. *J Biosci Bioeng* 117(1):57-64

Hjortmo S, Patring J, Andlid T (2008) Growth rate and medium composition strongly affect folate content in *Saccharomyces cerevisiae*. *Int J Food Microbiol* 123(1-2):93-100

Holzer H, Goedde H (1957) Zwei Wege von Pyruvat zu Acetyl-CoenzymA in Hefe. *Biochemische Zeitschrift* 329(3):175-91

Hommes FA (1966) Mechanism of the Crabtree effect in yeast grown with different glucose concentrations. *Arch Biochem Biophys* 113(2):324-30

Horneffer V, Dreisewerd K, Lüdemann HC, Hillenkamp F, Läge M, Strupat K (1999) Is the incorporation of analytes into matrix crystals a prerequisite for matrix-assisted laser desorption/ionization mass spectrometry? A study of five positional isomers of dihydroxybenzoic acid. *Int J Mass Spectrom* 185–187(0):859-70

Hornik K, Leisch F, Zeileis A Octave. In: Zeileis A (ed) Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003) Vienna, Austria, 2003.

Hsu BH, Xie YX, Busch KL, Cooks RG (1983) Enhanced ionization of organic salts in secondary-ion mass spectrometry. *Int J Mass Spectrom Ion Physics* 51(2–3):225-33

Hunte C, Palsdottir H, Trumpower BL (2003) Protonmotive pathways and mechanisms in the cytochrome bc1 complex. *FEBS Letters* 545(1):39-46

Hutzler M (2009) Entwicklung und Optimierung von Methoden zur Identifizierung und Differenzierung von getränkerelevanten Hefen. *Dissertation* TU München, Freising-Weihenstephan

Hutzler M, Geiger E, Jacob F (2010) Use of PCR-DHPLC (Polymerase Chain Reaction - Denaturing High Performance Liquid Chromatography) for the rapid differentiation of industrial *Saccharomyces pastorianus* and *Saccharomyes cerevisiae* strains. *J Inst Brew* 116(4):464-74

Hutzler M, Wellhoener U, Tenge C, Geiger E (2008) Beer mixed beverages: dangerous spoilage yeasts, susceptible beverages. *Brauwelt International* 26:206-11

Inadome H, Noda Y, Adachi H, Yoda K (2005) Immunoisolaton of the yeast Golgi subcompartments and characterization of a novel membrane protein, Svp26, discovered in the Sed5-containing compartments. *Mol Cell Biol* 25(17):7696-710

Ingendoh A, Karas M, Hillenkamp F, Giessmann U (1994) Factors affecting the resolution in matrix-assisted laser desorption—ionization mass spectrometry. *Int J Mass Spectrom Ion Physics* 131:345-54

Jackson M (1999) Bier International, Hallwag, Bern

James AG, Cook RM, West SM, Lindsay JG (1995) The pyruvate dehydrogenase complex of *Saccharomyces cerevisiae* is regulated by phosphorylation. *FEBS Letters* 373(2):111-4

James TC, Usher J, Campbell S, Bond U (2008) Lager yeasts possess dynamic genomes that undergo rearrangements and gene amplification in response to stress. *Curr Genet* 53(3):139-52

Jay S, Anderson J (2001) Fruit juice and related products, in: Spoilage of processed foods: causes and diagnosis. Moir CJ, Andrew-Kabilafkas C, Arnold G, Cox BM, Hocking AD, Jenson I (eds), Southwest Press, Marrickville, pp. 187-98

Jeffries N (2005) Algorithms for alignment of mass spectrometry proteomic data. *Bioinformatics* 21(14):3066-73

Jenkins RO, Cartledge TG, Lloyd D (1984) Respiratory adaptation of anaerobically grown *Saccharomyces uvarum*: changes in distribution of enzymes. *J Gen Microbiol* 130(11):2809-16

Jespersen L, Cesar LB, Meaden PG, Jakobsen M (1999) Multiple α-glucoside transporter genes in brewer's yeast. *Appl Environ Microbiol* 65(2):450-6

Jespersen L, Jakobsen M (1996) Specific spoilage organisms in breweries and laboratory media for their detection. *Int J Food Microbiol* 33(1):139-55

Johnston M (1999) Feasting, fasting and fermenting. Glucose sensing in yeast and other cells. *Trends Genet* 15(1):29-33

Jombart T (2013) A tutorial for discriminant analysis of principal components (DAPC) using *adegenet* 1.4-0. *http://cranr-projectorg/web/packages/adegenet/vignettes/adegenet-dapcpdf*

Jombart T, Caitlin C, Solymos P, Ahmed I, Calboli F, Cori A (2014) adegenet: an R package for the exploratory analysis of genetic and genomic data.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC genetics* 11:94

Jones DL (2012) The Fathom Toolbox for Matlab: multivariate ecological and oceanographic data analysis. *College of Marine Science, University of South Florida, St Petersburg, FL*

Jonsson GP, Hedin AB, Hakansson PL, Sundqvist BU, Saeve BGS, Nielsen PF, Roepstorff P, Johansson KE, Kamensky I, Lindberg MS (1986) Plasma desorption mass

spectrometry of peptides and proteins adsorbed on nitrocellulose. *Anal Chem* 58(6):1084-7

Joubert R, Brignon P, Lehmann C, Monribot C, Gendre F, Boucherie H (2000) Two-dimensional gel analysis of the proteome of lager brewing yeasts. *Yeast* 16(6):511-22

Ju Q, Warner JR (1994) Ribosome synthesis during the growth cycle of *Saccharomyces cerevisiae*. *Yeast* 10(2):151-7

Juhasz P, Costello CE, Biemann K (1993) Matrix-assisted laser desorption ionization mass spectrometry with 2-(4-hydroxyphenylazo)benzoic acid matrix. *J Am Soc Mass Spectrom* 4(5):399-409

Juhasz P, Vestal ML, Martin SA (1997) On the initial velocity of ions generated by matrix-assisted laser desorption ionization and its effect on the calibration of delayed extraction time-of-flight mass spectra. *J Am Soc Mass Spectrom* 8(3):209-17

Juvonen R, Virkajärvi V, Priha O, Laitila A (2011) Microbiological spoilage and safety risks in non-beer beverages. VTT, Technical Research Centre Finland

Kaiser JF, Schafer RW (1980) On the Use of the Io -Sinh Window fax Spectrum Analysis. *IEEE Trans Acoust Speech Signal Process* 28(1):105-7

Käppeli O (1986) Regulation of carbon metabolism in *Saccharomyces cerevisiae* and related yeasts. *Adv Microb Physiol* 28:181-209

Karas M, Bachmann D, Bahr U, Hillenkamp F (1987) Matrix-assisted ultraviolet laser desorption of non-volatile compounds. *Int J Mass Spectrom* 78(0):53-68

Karas M, Bachmann D, Hillenkamp F (1985) Influence of the wavelength in high-irradiance ultraviolet laser desorption mass spectrometry of organic molecules. *Anal Chem* 57(14):2935-9

Karas M, Bahr U, Hillenkamp F (1989) UV laser matrix desorption/ionization mass spectrometry of proteins in the 100 000 dalton range. *Int J Mass Spectrom* 92(0):231-42

Karas M, Glückmann M, Schäfer J (2000) Ionization in matrix-assisted laser desorption/ionization: singly charged molecular ions are the lucky survivors. *J Mass Spectrom* 35(1):1-12

Karas M, Krüger R (2003) Ion formation in MALDI: the cluster ionization mechanism. *Chem Rev* 103(2):427-40

Kaspar von Meyenburg H (1969) Energetics of the budding cycle of *Saccharomyces cerevisiae* during glucose limited aerobic growth. *Arch Mikrobiol* 66(4):289-303

Katz JJ (1954) Anhydrous trifluoroacetic acid as a solvent for proteins. *Nature* 174(4428):509

Kawahata M, Fujii T, Iefuji H (2007) Intraspecies diversity of the industrial yeast strains *Saccharomyces cerevisiae* and *Saccharomyces pastorianus* based on analysis of the sequences of the internal transcribed spacer (ITS) regions and the D1/D2 region of 26S rDNA. *Biosci Biotechnol Biochem* 71(7):1616-20

Kern CC, Vogel RF, Behr J (2014a) Differentiation of *Lactobacillus brevis* strains using Matrix-Assisted-Laser-Desorption-Ionization-Time-of-Flight Mass Spectrometry with respect to their beer spoilage potential. *Food microbiol* 40:18-24

Kern CC, Vogel RF, Behr J (2014b) Identification and differentiation of brewery isolates of *Pectinatus sp.* by Matrix-Assisted-Laser Desorption–Ionization Time-Of-Flight Mass Spectrometry (MALDI-TOF MS). *Eur Food Res Technol* 238(5):875-80

Ketterlinus R, Hsieh SY, Teng SH, Lee H, Pusch W (2005) Fishing for biomarkers: analyzing mass spectrometry data with the new ClinProTools software. *Biotech Suppl*:37-40

King ES, Osidacz P, Curtin C, Bastian SEP, Francis IL (2011) Assessing desirable levels of sensory properties in Sauvignon Blanc wines – consumer preferences and contribution of key aroma compounds. *Aust J Grape Wine Res* 17(2):169-80

Klis FM, Boorsma A, De Groot PW (2006) Cell wall construction in *Saccharomyces cerevisiae*. *Yeast* 23(3):185-202

Knochenmuss R (2006) Ion formation mechanisms in UV-MALDI. *Analyst* 131(9):966-86

Knochenmuss R (2013) MALDI ionization mechanisms: the coupled photophysical and chemical dynamics model correctly predicts 'temperature'-selected spectra. *J Mass Spectrom* 48(9):998-1004

Kobayashi O, Hayashi N, Kuroki R, Sone H (1998) Region of FLO1 proteins responsible for sugar recognition. *J Bacteriol* 180(24):6503-10

Koehler CM (2004) New developments in mitochondrial assembly. *Annu Rev Cell Dev Biol* 20:309-35

Krause-Buchholz U, Gey U, Wünschmann J, Becker S, Rödel G (2006) YIL042c and YOR090c encode the kinase and phosphatase of the *Saccharomyces cerevisiae* pyruvate dehydrogenase complex. *FEBS Letters* 580(11):2553-60

Kresze G-B, Ronft H (1981a) Pyruvate dehydrogenase complex from baker's yeast - 2. Molecular structure, dissociation, and implications for the origin of mitochondria. *Eur J Biochem* 119(3):581-7

Kresze GB, Ronft H (1981b) Pyruvate dehydrogenase complex from baker's yeast - 1. Purification and some kinetic and regulatory properties. *Eur J Biochem* 119(3):573-9

Krogerus K, Gibson BR (2013a) 125th Anniversary Review: Diacetyl and its control during brewery fermentation. *J Inst Brew* 119(3):86-97

Krogerus K, Gibson BR (2013b) Influence of valine and other amino acids on total diacetyl and 2, 3-pentanedione levels during fermentation of brewer's wort. *Appl Microbiol Biotechnol* 97(15):6919-30

Kruckeberg AL (1996) The hexose transporter family of *Saccharomyces cerevisiae*. *Arch Microbiol* 166(5):283-92

Kumar CG, Anand S (1998) Significance of microbial biofilms in food industry: a review. *Int J Food Microbiol* 42(1):9-27

Kunicka-Styczynska A, Rajkowska K (2011) Physiological and genetic stability of hybrids of industrial wine yeasts Saccharomyces sensu stricto complex. *J Appl Microbiol* 110(6):1538-49

Kurtzman CP (2011) *Wickerhamomyces* Kurtzman, Robnett & Basehoar-Powers (2008), in: The yeast: a taxonomic study. 5th edn, Elsevier, pp. 899-917

Kurtzman CP, Robnett CJ (1998) Identification and phylogeny of ascomycetous yeasts from analysis of nuclear large subunit (26S) ribosomal DNA partial sequences. *Anton Leeuw* 73(4):331-71

Kurtzman CT (1987) Prediction of biological relatedness among yeasts from comparisons of nuclear DNA complementarity. *Reprints-US Department of Agriculture, Agricultural Research Service (USA)*

Lafaye A, Junot C, Pereira Y, Lagniel G, Tabet JC, Ezan E, Labarre J (2005) Combined proteome and metabolite-profiling analyses reveal surprising insights into yeast sulfur metabolism. *J Biol Chem* 280(26):24723-30

Lawrence DR (1988) Spoilage organisms in beer, in: Developements in Food Microbiology. Robinson RK (ed), Elsevier Applied Science, London, pp. 1-48

Lee C, Abdool A, Huang C-H (2009) PCA-based population structure inference with generic clustering algorithms. *BMC Bioinformatics* 10(Suppl I):S73

Lee S, Madee M, Jangaard N, Horiuchi E (1980) *Pectinatus*, a new genus of bacteria capable of growth in hopped beer. *J Inst Brew* 86(1):28-30

Legras J-L, Karst F (2003) Optimisation of interdelta analysis for *Saccharomyces cerevisiae* strain characterisation. *FEMS Microbiol Lett* 221(2):249-55

Legras J-L, Ruh O, Merdinoglu D, Karst F (2005) Selection of hypervariable microsatellite loci for the characterization of *Saccharomyces cerevisiae* strains. *Int J Food Microbiol* 102(1):73-83

Legras JL, Merdinoglu D, Cornuet JM, Karst F (2007) Bread, beer and wine: *Saccharomyces cerevisiae* diversity reflects human history. *Mol Ecol* 16(10):2091-102

Lenovich L, Buchanan R, Worley N, Restaino L (1988) Effect of solute type on sorbate resistance in *Zygosaccharomyces rouxii*. *J Food Sci* 53(3):914-6

Libkind D, Hittinger CT, Valerio E, Gonçalves C, Dover J, Johnston M, Gonçalves P, Sampaio JP (2011) Microbe domestication and the identification of the wild genetic stock of lager-brewing yeast. *Proc Natl Acad Sci U S A* 108(35):14539-44

Lillie SH, Pringle JR (1980) Reserve carbohydrate metabolism in *Saccharomyces cerevisiae:* responses to nutrient limitation. *J Bacteriol* 143(3):1384-94

Lilly M, Lambrechts MG, Pretorius IS (2000) Effect of increased yeast alcohol acetyltransferase activity on flavor profiles of wine and distillates. *Appl Environ Microbiol* 66(2):744-53

Lima Tribst AA, de Souza Sant'Ana A, de Massaguer PR (2009) Review: Microbiological quality and safety of fruit juices-past, present and future perspectives. *Critical reviews in microbiology* 35(4):310-39

Liti G, Peruffo A, James SA, Roberts IN, Louis EJ (2005) Inferences of evolutionary relationships from a population survey of LTR-retrotransposons and telomeric-associated sequences in the Saccharomyces sensu stricto complex. *Yeast* 22(3):177-92

Lodolo EJ, Kock JLF, Axcell BC, Brooks M (2008) The yeast *Saccharomyces cerevisiae*– the main character in beer brewing. *FEMS Yeast Res* 8(7):1018-36

Madonna AJ, Basile F, Ferrer I, Meetani MA, Rees JC, Voorhees KJ (2000) On-probe sample pretreatment for the detection of proteins above 15 KDa from whole cell bacteria by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 14(23):2220-9

Mamyrin BA, Karataev VI, Shmikk DV, Zagulin VA (1973) The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *JETP* 37(1):45-8

Mantini D, Petrucci F, Del Boccio P, Pieragostino D, Di Nicola M, Lugaresi A, Federici G, Sacchetta P, Di Ilio C, Urbani A (2008) Independent component analysis for the extraction of reliable protein signal profiles from MALDI-TOF mass spectra. *Bioinformatics* 24(1):63-70

Mantini D, Petrucci F, Pieragostino D, Del Boccio P, Di Nicola M, Di Ilio C, Federici G, Sacchetta P, Comani S, Urbani A (2007) LIMPIC: a computational method for the separation of protein MALDI-TOF-MS signals from noise. *BMC Bioinformatics* 8:101

Mantini D, Petrucci F, Pieragostino D, Del Boccio P, Sacchetta P, Candiano G, Ghiggeri GM, Lugaresi A, Federici G, Di Ilio C, Urbani A (2010) A computational platform for MALDI-TOF mass spectrometry data: application to serum and plasma samples. *J Proteomics* 73(3):562-70

Marinangeli P, Angelozzi D, Ciani M, Clementi F, Mannazzu I (2004) Minisatellites in *Saccharomyces cerevisiae* genes encoding cell wall proteins: a new way towards wine strain characterisation. *FEMS Yeast Res* 4(4–5):427-35

Marklein G, Josten M, Klanke U, Müller E, Horré R, Maier T, Wenzel T, Kostrzewa M, Bierbaum G, Hoerauf A, Sahl HG (2009) Matrix-assisted laser desorption ionization-time of flight mass spectrometry for fast and reliable identification of clinical yeast isolates. *J Clin Microbiol* 47(9):2912-7

Marvin LF, Roberts MA, Fay LB (2003) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry in clinical chemistry. *Clin Chim Acta* 337(1-2):11-21

Masneuf-Pomarède I, Le Jeune C, Durrens P, Lollier M, Aigle M, Dubourdieu D (2007) Molecular typing of wine yeast strains *Saccharomyces bayanus* var. *uvarum* using microsatellite markers. *Syst Appl Microbiol* 30(1):75-82

Mason AB, Dufour J-P (2000) Alcohol acetyltransferases and the significance of ester synthesis in yeast. *Yeast* 16(14):1287-98

Mateo JJ, Jimenez M, Huerta T, Pastor A (1992) Comparison of volatiles produced by 4 *Saccharomyces cerevisiae* strains isolated from Monastrell musts. *Am J Enol Vitic* 43(2):206-9

McDonnell G, Russell AD (1999) Antiseptics and disinfectants: activity, action, and resistance. *Clinical microbiology reviews* 12(1):147-79

McDonnell GE (2007) Alcohols, in: Antisepsis, disinfection and sterilization - Types, action, and resistance. McDonnell GE (ed), Vol 4th, ASM Press, Washington D.C., pp. 90-2

McEwen CN, Larsen BS (2014) Fifty years of desorption ionization of nonvolatile compounds. *Int J Mass Spectrom* 377:515-31

Meaden PG, Taylor NR (1991) Cloning of a yeast gene which causes phenolic off-flavours in beer. *J Inst Brew* 97(5):353-7

Meilgaard M (2001) Effects on flavour of innovations in brewery equipment and processing. *J Inst Brew* 107(5):271-86

Merico A, Sulo P, Piškur J, Compagno C (2007) Fermentative lifestyle in yeasts belonging to the *Saccharomyces* complex. *Febs Journal* 274(4):976-89

Miki BL, Poon NH, James AP, Seligy VL (1982) Possible mechanism for flocculation interactions governed by gene FLO1 in *Saccharomyces cerevisiae*. *J Bacteriol* 150(2):878-89

Mitchell P (1976) Possible molecular mechanisms of the protonmotive function of cytochrome systems. *J Theor Biol* 62(2):327-67

Moothoo-Padayachie A, Kandappa HR, Krishna SBN, Maier T, Govender P (2013) Biotyping *Saccharomyces cerevisiae* strains using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS). *Eur Food Res Technol* 236:351-64

Morales L, Dujon B (2012) Evolutionary role of interspecies hybridization and genetic exchanges in yeasts. *Microbiol Mol Biol Rev* 76(4):721-39

Moremen KW, Molinari M (2006) N-linked glycan recognition and processing: the molecular basis of endoplasmic reticulum quality control. *Curr Opin Struct Biol* 16(5):592-9

Morgan B, Lu H (2008) Oxidative folding competes with mitochondrial import of the small Tim proteins. *Biochem J* 411:115-22

Nakao Y, Kanamori T, Itoh T, Kodama Y, Rainieri S, Nakamura N, Shimonaga T, Hattori M, Ashikari T (2009) Genome sequence of the lager brewing yeast, an interspecies hybrid. *DNA Res* 16(2):115-29

Naumov GI, James SA, Naumova ES, Louis EJ, Roberts IN (2000a) Three new species in the Saccharomyces sensu stricto complex: *Saccharomyces cariocanus*, *Saccharomyces kudriavzevii* and *Saccharomyces mikatae*. *Int J Syst Evol Microbiol* 50 Pt 5:1931-42

Naumov GI, Masneuf I, Naumova ES, Aigle M, Dubourdieu D (2000b) Association of *Saccharomyces bayanus* var. *uvarum* with some French wines: genetic analysis of yeast populations. *Res Microbiol* 151(8):683-91

Naumov GI, Naumova ES, Masneuf-Pomarede I (2010) Genetic identification of new biological species *Saccharomyces arboricolus* Wang et Bai. *Anton Leeuw* 98(1):1-7

Naumov GI, Nguyen HV, Naumova ES, Michel A, Aigle M, Gaillardin C (2001) Genetic identification of *Saccharomyces bayanus* var. *uvarum*, a cider-fermenting yeast. *Int J Food Microbiol* 65(3):163-71

Navarre C, Catty P, Leterme S, Dietrich F, Goffeau A (1994) Two distinct genes encode small isoproteolipids affecting plasma membrane H^+-ATPase activity of *Saccharomyces cerevisiae*. *J Biol Chem* 269(33):21262-8

Nazim Boutaghou M, Cole RB (2012) 9, 10-Diphenylanthracene as a matrix for MALDI-MS electron transfer secondary reactions. *J Mass Spectrom* 47(8):995-1003

Neupert W, Herrmann JM (2007) Translocation of proteins into mitochondria. *Annu Rev Biochem* 76:723-49

Nguyen H-V, Gaillardin C (2005) Evolutionary relationships between the former species *Saccharomyces uvarum* and the hybrids *Saccharomyces bayanus* and *Saccharomyces pastorianus;* reinstatement of *Saccharomyces uvarum* (Beijerinck) as a distinct species. *FEMS Yeast Res* 5(4–5):471-83

Nho S, Anderson MJ, Moore CB, Denning DW (1997) Species differentiation by internally transcribed spacer PCR and HhaI digestion of fluconazole-resistant Candida krusei, Candida inconspicua, and Candida norvegensis strains. *J Clin Microbiol* 35(4):1036-9

Nielsen J (2003) It is all about metabolic fluxes. *J Bacteriol* 185(24):7031-5

Nielsen J, Oliver S (2005) The next wave in metabolome analysis. *Trends Biotechnol* 23(11):544-6

Niessen L, Böhm-Schrami M, Vogel H, Donhauser S (1993) Deoxynivalenol in commercial beer—screening for the toxin with an indirect competitive ELISA. *Mycotoxin Research* 9(2):99-109

Noble AC, Bursick GF (1984) The contribution of glycerol to perceived viscosity and sweetness in white wine. *Am J Enol Vitic* 35(2):110-2

Nykänen L, Nykänen I (1977) Production of esters by different yeast strains in sugar fermentation. *J Inst Brew* 83(1):30-1

O'Rourke SM, Herskowitz I (2002) A third osmosensing branch in *Saccharomyces cerevisiae* requires the Msb2 protein and functions in parallel with the Sho1 branch. *Mol Cell Biol* 22(13):4739-49

Oey I, Lille M, Van Loey A, Hendrickx M (2008) Effect of high-pressure processing on colour, texture and flavour of fruit-and vegetable-based food products: a review. *Trends Food Sci Technol* 19(6):320-8

Ogata T, Izumikawa M, Kohno K, Shibata K (2008) Chromosomal location of Lg-FLO1 in bottom-fermenting yeast and the FLO5 locus of industrial yeast. *J Appl Microbiol* 105(4):1186-98

Ohlmeier S, Kastaniotis AJ, Hiltunen JK, Bergmann U (2004) The yeast mitochondrial proteome, a study of fermentative and respiratory growth. *J Biol Chem* 279(6):3956-79

Opsal RB, Owens KG, Reilly JP (1985) Resolution in the linear time-of-flight mass spectrometer. *Anal Chem* 57(9):1884-9

Özcan S, Dover J, Johnston M (1998) Glucose sensing and signaling by two glucose receptors in the yeast *Saccharomyces cerevisiae*. *EMBO J* 17(9):2566-73

Özcan S, Johnston M (1999) Function and regulation of yeast hexose transporters. *Microbiol Mol Biol Rev* 63(3):554-69

Papp A, Winnewisser W, Geiger E, Briem F (2001) Influence of (+)-Catechin and Ferulic Acid on Formation of Beer Haze and Their Removal through different Polyvinylpolypyrollidone-Types. *J Inst Brew* 107(1):55-60

Paradh A, Mitchell W, Hill A (2011) Occurrence of *Pectinatus* and *Megasphaera* in the major UK breweries. *J Inst Brew* 117(4):498-506

Peddie HAB (1990) Ester formation in brewery fermentation. *J Inst Brew* 96:327-31

Peris D, Lopes CA, Belloch C, Querol A, Barrio E (2012) Comparative genomics among *Saccharomyces cerevisiae x Saccharomyces kudriavzevii* natural hybrid strains isolated from wine and beer reveals different origins. *BMC Genomics* 13:407

Perkins DN, Pappin DJC, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20(18):3551-67

Pfanner N, Wiedemann N (2002) Mitochondrial protein import: two membranes, three translocases. *Curr Opin Cell Biol* 14(4):400-11

Pinto A, Halliday C, Zahra M, van Hal S, Olma T, Maszewska K, Iredell JR, Meyer W, Chen SC (2011) Matrix-assisted laser desorption ionization-time of flight mass spectrometry identification of yeasts is contingent on robust reference spectra. *PLoS One* 6(10):e25712

Piper P, Calderon CO, Hatzixanthis K, Mollapour M (2001) Weak acid adaptation: the stress response that confers yeasts with resistance to organic acid food preservatives. *Microbiol* 147(10):2635-42

Pires EJ, Teixeira JA, Brányik T, Vicente AA (2014) Yeast: the soul of beer's aroma—a review of flavour-active esters and higher alcohols produced by the brewing yeast. *Appl Microbiol Biotechnol* 98(5):1937-49

Piškur J, Rozpędowska E, Polakova S, Merico A, Compagno C (2006) How did *Saccharomyces* evolve to become a good brewer? *Trends Genet* 22(4):183-6

Pizarro F, Vargas FA, Agosin E (2007) A systems biology perspective of wine fermentations. *Yeast* 24(11):977-91

Plesset J, Ludwig JR, Cox BS, McLaughlin CS (1987) Effect of cell cycle position on thermotolerance in *Saccharomyces cerevisiae*. *J Bacteriol* 169(2):779-84

Pope GA, MacKenzie DA, Defernez M, Aroso MAMM, Fuller LJ, Mellon FA, Dunn WB, Brown M, Goodacre R, Kell DB, Marvin ME, Louis EJ, Roberts IN (2007) Metabolic footprinting as a tool for discriminating between brewing yeasts. *Yeast* 24(8):667-79

Posthumus M, Kistemaker P, Meuzelaar H, Ten Noever de Brauw M (1978) Laser desorption-mass spectrometry of polar nonvolatile bio-organic molecules. *Anal Chem* 50(7):985-91

Postma E, Verduyn C, Scheffers WA, Van Dijken JP (1989) Enzymic analysis of the crabtree effect in glucose-limited chemostat cultures of *Saccharomyces cerevisiae*. *Appl Environ Microbiol* 55(2):468-77

Praekelt UM, Meacock PA (1990) HSP12, a new small heat shock gene of *Saccharomyces cerevisiae*: analysis of structure, regulation and function. *Mol Gen Genet* 223(1):97-106

Preissler P (2011) Categorization of *Lactobacillus brevis* along their beer-spoiling potential. *Dissertation* TU München, Freising-Weihenstphan

Priest FG (1981) Contamination, in: An Introduction of Brewing Science and Technology, Part II. London, pp. 23-31

Priest FG, Campbell I (2003) Brewing Microbiology, 3rd edn. Plenum Publisher, New York

Pronk JT, Steensma HY, van Dijken JP (1996) Pyruvate metabolism in *Saccharomyces cerevisiae*. *Yeast* 12(16):1607-33

Pronk JT, Wenzel TJ, Luttik MA, Klaassen CC, Scheffers WA, Steensma HY, van Dijken JP (1994) Energetic aspects of glucose metabolism in a pyruvate-dehydrogenase-negative mutant of *Saccharomyces cerevisiae*. *Microbiol* 140(3):601-10

Putignani L, Del Chierico F, Onori M, Mancinelli L, Argentieri M, Bernaschi P, Coltella L, Lucignano B, Pansani L, Ranno S, Russo C, Urbani A, Federici G, Menichella D (2011) MALDI-TOF mass spectrometry proteomic phenotyping of clinically relevant fungi. *Mol Biosyst* 7(3):620-9

Qian J, Cutler JE, Cole RB, Cai Y (2008) MALDI-TOF mass signatures for differentiation of yeast species, strain grouping and monitoring of morphogenesis markers. *Anal Bioanal Chem* 392(3):439-49

Querol A, Bond U (2009) The complex and dynamic genomes of industrial yeasts. *FEMS Microbiol Lett* 293(1):1-10

Rabin D, Forget C (2014) Dictionary of Beer and Brewing. Routledge

Rainieri S, Kodama Y, Kaneko Y, Mikata K, Nakao Y, Ashikari T (2006) Pure and mixed genetic lines of *Saccharomyces bayanus* and *Saccharomyces pastorianus* and their contribution to the lager brewing strain genome. *Appl Environ Microbiol* 72(6):3968-74

Rainieri S, Zambonelli C, Kaneko Y (2003) *Saccharomyces* sensu stricto: systematics, genetic diversity and evolution. *J Biosci Bioeng* 96(1):1-9

Ramette A (2007) Multivariate analyses in microbial ecology. *FEMS Microbiol Ecol* 62(2):142-60

Ramos-Jeunehomme C, Laub R, Masschelein C Why is ester formation in brewery fermentations yeast strain dependent. In: Proc Congr Eur Brew Conv, 1991. Vol 23. p. 23

Rauhut D (1994) Yeasts - production of sulfur compounds, in: Wine microbiology and biotechnology. Fleet GH (ed), Harwood Academic, Chur, Schwitzerland, pp. 183-223

Remize F, Roustan JL, Sablayrolles JM, Barre P, Dequin S (1999) Glycerol overproduction by engineered *Saccharomyces cerevisiae* wine yeast strains leads to substantial changes in By-product formation and to a stimulation of fermentation rate in stationary phase. *Appl Environ Microbiol* 65(1):143-9

Renard A, Gómez di Marco P, Egea-Cortines M, Weiss J (2008) Application of whole genome amplification and quantitative PCR for detection and quantification of spoilage yeasts in orange juice. *Int J Food Microbiol* 126(1):195-201

Replansky T, Koufopanou V, Greig D, Bell G (2008) Saccharomyces sensu stricto as a model system for evolution and ecology. *Trends in ecology & evolution* 23(9):494-501

Richter CL, Dunn B, Sherlock G, Pugh T (2013) Comparative metabolic footprinting of a large number of commercial wine yeast strains in Chardonnay fermentations. *FEMS Yeast Res* 13(4):394-410

Rodrigues F, Ludovico P, Leão C (2006) Sugar metabolism in yeasts: an overview of aerobic and anaerobic glucose catabolism, in: Biodiversity and ecophysiology of yeasts. Springer, pp. 101-21

Rohde M, Lim F, Wallace JC (1991) Electron microscopic localization of pyruvate carboxylase in rat liver and *Saccharomyces cerevisiae* by immunogold procedures. *Arch Biochem Biophys* 290(1):197-201

Romano P, Capece A, Jespersen L (2006) Taxonomic Ecological Diversity of Food and Beverage Yeasts, in: The Yeast Handbook - Yeasts in Food and Beverages. Querol A, Fleet GH (eds), Vol 2, Springer, Berlin, pp. 13-53

Romero PA, Lussier M, Veronneau S, Sdicu A-M, Herscovics A, Bussey H (1999) Mnt2p and Mnt3p of *Saccharomyces cerevisiae* are members of the Mnn1p family of α-1,3-mannosyltransferases responsible for adding the terminal mannose residues of O-linked oligosaccharides. *Glycobiology* 9(10):1045-51

Rossell S, van der Weijden CC, Kruckeberg AL, Bakker BM, Westerhoff HV (2005) Hierarchical and metabolic regulation of glucose influx in starved *Saccharomyces cerevisiae*. *FEMS Yeast Res* 5(6–7):611-9

Ruiz-Amil M, De Torrontegui G, Palacian E, Catalina L, Losada M (1965) Properties and function of yeast pyruvate carboxylase. *J Biol Chem* 240(9):3485-92

Saerens SM, Delvaux F, Verstrepen KJ, Van Dijck P, Thevelein JM, Delvaux FR (2008a) Parameters affecting ethyl ester production by *Saccharomyces cerevisiae* during fermentation. *Appl Environ Microbiol* 74(2):454-61

Saerens SM, Verbelen PJ, Vanbeneden N, Thevelein JM, Delvaux FR (2008b) Monitoring the influence of high-gravity brewing and fermentation temperature on flavour formation by analysis of gene expression levels in brewing yeast. *Appl Microbiol Biotechnol* 80(6):1039-51

Sakamoto K, Konings WN (2003) Beer spoilage bacteria and hop resistance. *Int J Food Microbiol* 89(2):105-24

Salema-Oom M, Pinto VV, Gonçalves P, Spencer-Martins I (2005) Maltotriose utilization by industrial *Saccharomyces* strains: characterization of a new member of the α-glucoside transporter family. *Appl Environ Microbiol* 71(9):5044-9

Sarlin T, Nakari-Setälä T, Linder M, Penttilä M, Haikara A (2005) Fungal hydrophobins as predictors of the gushing activity of malt. *J Inst Brew* 111(2):105-11

Sato M, Maeba H, Watari J, Takashio M (2002) Analysis of an inactivated Lg-FLO1 gene present in bottom-fermenting yeast. *J Biosci Bioeng* 93(4):395-8

Satten GA, Datta S, Moura H, Woolfitt AR, Carvalho Mda G, Carlone GM, De BK, Pavlopoulos A, Barr JR (2004) Standardization and denoising algorithms for mass spectra to classify whole-organism bacterial specimens. *Bioinformatics* 20(17):3128-36

Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal Biochem* 166(2):368-79

Schawalder SB, Kabani M, Howald I, Choudhury U, Werner M, Shore D (2004) Growth-regulated recruitment of the essential yeast ribosomal protein gene activator Ifh1. *Nature* 432(7020):1058-61

Scherer A (2002) Entwicklung von PCR-Methoden zur Klassifizierung industriell genutzter Hefen. *Dissertation* TU München, Freising-Weihenstephan

Schönenberg S (2010) Der physiologische Zustand und der Sauerstoffbedarf von Bierhefen unter brautechnologischen Bedingungen - Optimierung der Propagations- und Anstelltechnologie in Hinblick auf die Gäreigenschaften und die Bierqualität. *Dissertation* TU München, Freising-Weihenstephan

Schuller D, Valero E, Dequin S, Casal M (2004) Survey of molecular methods for the typing of wine yeast strains. *FEMS Microbiol Lett* 231(1):19-26

Schwarz PB, Casper HH, Beattie S (1995) Fate and development of naturally occurring *Fusarium* mycotoxins during malting and brewing. *J Am Soc Brew Chem* 53:121-7

Seng P, Drancourt M, Gouriet F, La Scola B, Fournier PE, Rolain JM, Raoult D (2009) Ongoing revolution in bacteriology: routine identification of bacteria by matrix-assisted laser desorption ionization time-of-flight mass spectrometry. *Clin Infect Dis* 49(4):543-51

Shi X, Zhu X (2009) Biofilm formation and food safety in food industries. *Trends Food Sci Technol* 20(9):407-13

Shimazu M, Sekito T, Akiyama K, Ohsumi Y, Kakinuma Y (2005) A family of basic amino acid transporters of the vacuolar membrane from *Saccharomyces cerevisiae*. *J Biol Chem* 280(6):4851-7

Shimazu Y, Watanabe M (1981) Effects of yeast strains and environmental conditions on formation of organic acids in must during fermentation: studies on organic acids in grape must and wine (III). *J Ferment Technol* 59(1):27-32

Shinohara T, Kubodera S, Yanagida F (2000) Distribution of phenolic yeasts and production of phenolic off-flavors in wine fermentation. *J Biosci Bioeng* 90(1):90-7

Sipiczki M (2008) Interspecies hybridization and recombination in *Saccharomyces* wine yeasts. *FEMS Yeast Res* 8(7):996-1007

Smirnov IP, Zhu X, Taylor T, Huang Y, Ross P, Papayanopoulos IA, Martin SA, Pappin DJ (2004) Suppression of alpha-cyano-4-hydroxycinnamic acid matrix clusters and reduction of chemical noise in MALDI-TOF mass spectrometry. *Anal Chem* 76(10):2958-65

Sogawa K, Watanabe M, Sato K, Segawa S, Ishii C, Miyabe A, Murata S, Saito T, Nomura F (2011) Use of the MALDI BioTyper system with MALDI-TOF mass spectrometry for rapid identification of microorganisms. *Anal Bioanal Chem* 400(7):1905-11

Spiller MP, Guo L, Wang Q, Tran P, Lu H (2015) Mitochondrial Tim9 protects Tim10 from degradation by the protease Yme1. *Biosci Rep* 35(3):e00193

Steensels J, Snoek T, Meersman E, Picca Nicolino M, Aslankoohi E, Christiaens JF, Gemayel R, Meert W, New AM, Pougach K, Saels V, van der Zande E, Voordeckers K, Verstrepen KJ (2012) Selecting and generating superior yeasts for the brewing industry. *Cerevisia* 37(2):63-7

Steensels J, Verstrepen KJ (2014) Taming Wild Yeast: Potential of Conventional and Nonconventional Yeasts in Industrial Fermentations. *Annual review of microbiology*(0)

Stevenson LG, Drake SK, Shea YR, Zelazny AM, Murray PR (2010) Evaluation of matrix-assisted laser desorption ionization-time of flight mass spectrometry for identification of clinically important yeast species. *J Clin Microbiol* 48(10):3482-6

Stewart GG, Murray CR, Panchal CJ, Russell I, Sills AM (1984) The selection and modification of brewer's yeast strains. *Food Microbiol* 1(4):289-302

Stoll R, Röllgen FW (1979) Laser desorption mass spectrometry of thermally labile compounds using a continuous wave CO_2 laser. *Org Mass Spectrom* 14(12):642-5

Storgårds E (2000) Process hygiene control in beer production and dispensing. *Espoo, VTT Publications, Technical Research Centre Finland*:105 p.

Stratford M (1992) Yeast flocculation: a new perspective. *Adv Microb Physiol* 33:2-71

Strupat K, Karas M, Hillenkamp F (1991) 2,5-Dihydroxybenzoic acid: a new matrix for laser desorption-ionization mass spectrometry. *Int J Mass Spectrom* 111:89-102

Subileau M, Schneider R, Salmon JM, Degryse E (2008) New insights on 3-mercaptohexanol (3MH) biogenesis in Sauvignon Blanc wines: Cys-3MH and (E)-hexen-2-al are not the major precursors. *J Agric Food Chem* 56(19):9230-5

Suomalainen H (1981) Yeast esterases and aroma esters in alcoholic beverages. *J Inst Brew* 87(5):296-300

Suzuki K, Iijima K, Sakamoto K, Sami M, Yamashita H (2006) A review of hop resistance in beer spoilage lactic acid bacteria. *J Inst Brew* 112(2):173-91

Suzuki M, Prasad GS, Kurtzman CP (2011) *Debaryomyces* Lodder & Kreger-van Rij (1952), in: The yeast: a taxonomic study. 5th edn, Elsevier, pp. 361-72

Swiegers JH, Kievit RL, Siebert T, Lattey KA, Bramley BR, Francis IL, King ES, Pretorius IS (2009) The influence of yeast on the aroma of Sauvignon Blanc wine. *Food Microbiol* 26(2):204-11

Swiegers JH, Pretorius IS (2007) Modulation of volatile sulfur compounds by wine yeast. *Appl Microbiol Biotechnol* 74(5):954-60

Swiegers JH, Willmott R, Hill-Ling A, Capone DL, Pardon KH, Elsey GM, Howell KS, de Barros Lopes MA, Sefton MA, Lilly M, Pretorius IS (2006) Modulation of volatile thiol and ester aromas by modified wine yeast, in: Developments in Food Science. Wender LPB, Mikael Agerlin P (eds), Vol 43, Elsevier, Oxford, UK, pp. 113-6

Tchango Tchango J, Tailliez R, Eb P, Njine T, Hornez J (1997) Heat resistance of the spoilage yeasts *Candida pelliculosa* and *Kloeckera apis* and pasteurization values for some tropical fruit juices and nectars. *Food Microbiol* 14(1):93-9

Team RDC (2011) R: A language and environment for statistical computing R Foundation for Statistical Computing. Vienna, Austria

ter Kuile BH, Westerhoff HV (2001) Transcriptome meets metabolome: hierarchical and metabolic regulation of the glycolytic pathway. *FEBS Letters* 500(3):169-71

ter Linde JJ, Liang H, Davis RW, Steensma HY, van Dijken JP, Pronk JT (1999) Genome-wide transcriptional analysis of aerobic and anaerobic chemostat cultures of *Saccharomyces cerevisiae*. *J Bacteriol* 181(24):7409-13

Thibon C, Marullo P, Claisse O, Cullin C, Dubourdieu D, Tominaga T (2008) Nitrogen catabolic repression controls the release of volatile thiols by *Saccharomyces cerevisiae* during wine fermentation. *FEMS Yeast Res* 8(7):1076-86

Tholozan J-L, Membré J-M, Grivet J-P (1997) Physiology and development of *Pectinatus cerevisiiphilus* and *Pectinatus frisingensis*, two strict anaerobic beer spoilage bacteria. *Int J Food Microbiol* 35(1):29-39

Thomas D, Davenport R (1985) *Zygosaccharomyces bailii* - a profile of characteristics and spoilage activities. *Food Microbiol* 2(2):157-69

Thurston PA (1986) Detection and control of contaminants in pitching yeast, in: European Brewery Convention - Monograph-XII. E.B.C.-Symposium on brewer's yeast, Vuoranta (Helsinki), Finland, pp. 84-94

Timke M, Wang-Lieu NQ, Altendorf K, Lipski A (2008) Identity, beer spoiling and biofilm forming potential of yeasts from beer bottling plant associated biofilms. *Anton Leeuw* 93(1-2):151-61

Tong DL, Boocock DJ, Coveney C, Saif J, Gomez SG, Querol S, Rees R, Ball GR (2011) A simpler method of preprocessing MALDI-TOF MS data for differential biomarker analysis: stem cell and melanoma cancer studies. *Clin Proteomics* 8:14

Torija MJ, Beltran G, Novo M, Poblet M, Guillamon JM, Mas A, Rozes N (2003) Effects of fermentation temperature and *Saccharomyces* species on the cell fatty acid composition and presence of volatile compounds in wine. *Int J Food Microbiol* 85(1-2):127-36

Tornai-Lehoczki J, Dlauchy D (2000) Delimination of brewing yeast strains using different molecular techniques. *Int J Food Microbiol* 62(1-2):37-45

Tornai-Lehoczki J, Péter G, Dlauchy D, Deák T (1996) Some remarks on “a taxonomic key for the genus *Saccharomyces*” (Vaughan Martini and Martini 1993). *Anton Leeuw* 69(3):229-33

Uhlinger DJ, Yang CY, Reed LJ (1986) Phosphorylation-dephosphorylation of pyruvate dehydrogenase from bakers' yeast. *Biochem* 25(19):5673-7

Vadaie N, Dionne H, Akajagbor DS, Nickerson SR, Krysan DJ, Cullen PJ (2008) Cleavage of the signaling mucin Msb2 by the aspartyl protease Yps1 is required for MAPK activation in yeast. *J Cell Biol* 181(7):1073-81

Valentin E, Herrero E, Rico H, Miragall F, Sentandreu R (1987) Cell wall mannoproteins during the population growth phases in *Saccharomyces cerevisiae*. *Arch Microbiol* 148(2):88-94

Valentine N, Wunschel S, Wunschel D, Petersen C, Wahl K (2005) Effect of culture conditions on microorganism identification by matrix-assisted laser desorption ionization mass spectrometry. *Appl Environ Microbiol* 71(1):58-64

van den Brink J, Daran-Lapujade P, Pronk JT, de Winde JH (2008) New insights into the *Saccharomyces cerevisiae* fermentation switch: dynamic transcriptional response to anaerobicity and glucose-excess. *BMC Genomics* 9(1):100

van der Aa Kühle A, Jespersen L (1998) Detection and identification of wild yeasts in lager breweries. *Int J Food Microbiol* 43(3):205-13

van Dijken JP, Scheffers WA (1986) Redox balances in the metabolism of sugars by yeasts. *FEMS Microbiol Lett* 32(3–4):199-224

van Dijken JP, Weusthuis RA, Pronk JT (1993) Kinetics of growth and sugar consumption in yeasts. *Anton Leeuw* 63(3-4):343-52

van Herendael BH, Bruynseels P, Bensaid M, Boekhout T, de Baere T, Surmont I, Mertens AH (2012) Validation of a modified algorithm for the identification of yeast isolates using

matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (MALDI-TOF MS). *Eur J Clin Microbiol Infect Dis* 31(5):841-8

van Mulders SE, Ghequire M, Daenen L, Verbelen PJ, Verstrepen KJ, Delvaux FR (2010) Flocculation gene variability in industrial brewer's yeast strains. *Appl Microbiol Biotechnol* 88(6):1321-31

van Rensburg S, Cook-Mozaffari P, van Schalkwyk D, van der Watt J, Vincent T, Purchase I (1985) Hepatocellular carcinoma and dietary aflatoxin in Mozambique and Transkei. *Brit J Cancer* 51(5):713

van Urk H, Schipper D, Breedveld GJ, Mak PR, Scheffers WA, van Dijken JP (1989) Localization and kinetics of pyruvate-metabolizing enzymes in relation to aerobic alcoholic fermentation in *Saccharomyces cerevisiae* CBS 8066 and *Candida utilis* CBS 621. *Biochim Biophys Acta* 992(1):78-86

van Urk H, Voll WS, Scheffers WA, van Dijken JP (1990) Transient-state analysis of metabolic fluxes in crabtree-positive and crabtree-negative yeasts. *Appl Environ Microbiol* 56(1):281-7

van Veen SQ, Claas EC, Kuijper EJ (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(3):900-7

Vanbeneden N, Gils F, Delvaux F, Delvaux FR (2008) Formation of 4-vinyl and 4-ethyl derivatives from hydroxycinnamic acids: Occurrence of volatile phenolic flavour compounds in beer and distribution of Pad1-activity among brewing yeasts. *Food chemistry* 107(1):221-30

Varela JC, Praekelt UM, Meacock PA, Planta RJ, Mager WH (1995) The *Saccharomyces cerevisiae* HSP12 gene is activated by the high-osmolarity glycerol pathway and negatively regulated by protein kinase A. *Mol Cell Biol* 15(11):6232-45

Vargha M, Takats Z, Konopka A, Nakatsu CH (2006) Optimization of MALDI-TOF MS for strain level differentiation of Arthrobacter isolates. *J Microbiol Methods* 66(3):399-409

Vaughan-Martini A, Martini A (2011) *Saccharomyces* Meyen ex Reess (1870), in: The yeast: a taxonomic study. 5th edn, Elsevier, London, UK

Vaughan A, O'Sullivan T, Sinderen D (2005) Enhancing the microbiological stability of malt and beer—a review. *J Inst Brew* 111(4):355-71

Verbelen P, Dekoninck T, Saerens S, Van Mulders S, Thevelein J, Delvaux F (2009) Impact of pitching rate on yeast fermentation performance and beer flavour. *Appl Microbiol Biotechnol* 82(1):155-67

Verbelen PJ, Mulders S, Saison D, Laere S, Delvaux F, Delvaux FR (2008) Characteristics of high cell density fermentations with different lager yeast strains. *J Inst Brew* 114(2):127-33

Verstrepen KJ, Derdelinckx G, Dufour J-P, Winderickx J, Thevelein JM, Pretorius IS, Delvaux FR (2003a) Flavor-active esters: Adding fruitiness to beer. *J Biosci Bioeng* 96(2):110-8

Verstrepen KJ, Derdelinckx G, Verachtert H, Delvaux FR (2003b) Yeast flocculation: what brewers should know. *Appl Microbiol Biotechnol* 61(3):197-205

Verstrepen KJ, Klis FM (2006) Flocculation, adhesion and biofilm formation in yeasts. *Mol Microbiol* 60(1):5-15

Vestal M, Juhasz P (1998) Resolution and mass accuracy in matrix-assisted laser desorption ionization-time-of-flight. *J Am Soc Mass Spectrom* 9(9):892-911

Vestal M, Juhasz P, Martin S (1995) Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 9(11):1044-50

Vidgren V, Londesborough J (2011) 125th anniversary review: yeast flocculation and sedimentation in brewing. *J Inst Brew* 117(4):475-87

Visser W, Scheffers WA, Batenburg-van der Vegte WH, van Dijken JP (1990) Oxygen requirements of yeasts. *Appl Environ Microbiol* 56(12):3785-92

Visser W, van der Baan AA, Batenburg-van der Vegte W, Scheffers WA, Kramer R, van Dijken JP (1994) Involvement of mitochondria in the assimilatory metabolism of anaerobic *Saccharomyces cerevisiae* cultures. *Microbiol* 140(11):3039-46

Vorm O, Roepstorff P, Mann M (1994) Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. *Anal Chem* 66(19):3281-7

Wade JT, Hall DB, Struhl K (2004) The transcription factor Ifh1 is a key regulator of yeast ribosomal protein genes. *Nature* 432(7020):1054-8

Wallace PG, Linnane AW (1964) Oxygen-induced synthesis of yeast mitochondria. *Nature* 201:1191

Walther A, Hesselbart A, Wendland J (2014) Genome sequence of *Saccharomyces carlsbergensis*, the world's first pure culture lager yeast. *G3: Genes*| *Genomes*| *Genetics* 4(5):783-93

Wang SA, Bai FY (2008) *Saccharomyces arboricolus sp. nov*., a yeast species from tree bark. *Int J Syst Evol Microbiol* 58(2):510-4

Wang X, Zhu W, Pradhan K, Ji C, Ma Y, Semmes OJ, Glimm J, Mitchell J (2006) Feature extraction in the analysis of proteomic mass spectra. *Proteomics* 6(7):2095-100

Wareing P, Davenport RR (2005) Microbiology of soft drinks and fruit juices, in: Chemistry and technology of soft drinks and fruit juices. Ashurst PR (ed), Blackwell Publishing, London, UK, pp. 279-97

Warner JR (1999) The economics of ribosome biosynthesis in yeast. *Trends Biochem Sci* 24(11):437-40

Welker M, Moore, E. R. B. (2011) Applications of whole-cell matrix-assisted laser-desorption/ionization time-of-flight mass spectrometry in systematic microbiology. *Syst Appl Microbiol* 34:2-11

Welker S, Rudolph B, Frenzel E, Hagn F, Liebisch G, Schmitz G, Scheuring J, Kerth A, Blume A, Weinkauf S, Haslbeck M, Kessler H, Buchner J (2010) Hsp12 is an intrinsically unstructured stress protein that folds upon membrane association and modulates membrane function. *Mol Cell* 39(4):507-20

Wenning M (2004) Identifizierung von Mikroorganismen durch Fourier-transformierte Infrarot (FTIR)-Mikrospektroskopie. *Dissertation* TU München, Freising-Weihenstephan

Wenning M, Seiler H, Scherer S (2002) Fourier-transform infrared microspectroscopy, a novel and rapid tool for identification of yeasts. *Appl Environ Microbiol* 68(10):4717-21

Werner-Washburne M, Braun E, Johnston GC, Singer RA (1993) Stationary phase in the yeast *Saccharomyces cerevisiae*. *Microbiol Rev* 57(2):383-401

Westman A, Demirev P, Huth-Fehre T, Bielawski J, Sundqvist B (1994) Sample exposure effects in matrix-assisted laser desorption—ionization mass spectrometry of large biomolecules. *Int J Mass Spectrom* 130(1):107-15

Whittal RM, Li L (1995) High-resolution matrix-assisted laser desorption/ionization in a linear time-of-flight mass spectrometer. *Anal Chem* 67(13):1950-4

Wieme AD, Spitaels F, Vandamme P, Van Landschoot A (2014) Application of matrix-assisted laser desorption/ionization time-of-flight mass spectrometry as a monitoring tool for in-house brewer's yeast contamination: a proof of concept. *J Inst Brew* 120(4):438-43

Wieser A, Schneider L, Jung J, Schubert S (2011) MALDI-TOF MS in microbiological diagnostics-identification of microorganisms and beyond (mini review). *Appl Microbiol Biotechnol* 93(3):965-74

Wightman P, Quain DE, Meaden PG (1996) Analysis of production brewing strains of yeast by DNA fingerprinting. *Lett Appl Microbiol* 22(1):90-4

Wilcox LJ, Balderes DA, Wharton B, Tinkelenberg AH, Rao G, Sturley SL (2002) Transcriptional profiling identifies two members of the ATP-binding cassette transporter superfamily required for sterol uptake in yeast. *J Biol Chem* 277(36):32466-72

Wiley W, McLaren IH (1955) Time-of-flight mass spectrometer with improved resolution. *Rev Sci Instrum* 26(12):1150-7

Wilkinson BM, Purswani J, Stirling CJ (2006) Yeast GTB1 encodes a subunit of glucosidase II required for glycoprotein processing in the endoplasmic reticulum. *J Biol Chem* 281(10):6325-33

Williams TL, Andrzejewski D, Lay JO, Musser SM (2003) Experimental factors affecting the quality and reproducibility of MALDI TOF mass spectra obtained from whole bacteria cells. *J Am Soc Mass Spectrom* 14(4):342-51

Wong SF, Mak JW, Pook PC (2007) New mechanical disruption method for extraction of whole cell protein from *Candida albicans*. *Southeast Asian J Trop Med Public Health* 38(3):512-8

Wunschel DS, Hill EA, McLean JS, Jarman K, Gorby YA, Valentine N, Wahl K (2005) Effects of varied pH, growth rate and temperature using controlled fermentation and batch culture on matrix assisted laser desorption/ionization whole cell protein fingerprints. *J Microbiol Methods* 62(3):259-71

Yamagishi H, Ogata T (1999) Chromosomal structures of bottom fermenting yeasts. *Syst Appl Microbiol* 22(3):341-53

Yamagishi H, Ohnuki S, Nogami S, Ogata T, Ohya Y (2010) Role of bottom-fermenting brewer's yeast KEX2 in high temperature resistance and poor proliferation at low temperatures. *J Gen Appl Microbiol* 56(4):297-312

Yang C, He Z, Yu W (2009) Comparison of public peak detection algorithms for MALDI mass spectrometry data analysis. *BMC Bioinformatics* 10

Yao J, Scott JR, Young MK, Wilkins CL (1998) Importance of matrix:analyte ratio for buffer tolerance using 2,5-dihydroxybenzoic acid as a matrix in matrix-assisted laser desorption/ionization-Fourier transform mass spectrometry and matrix-assisted laser desorption/ionization-time of flight. *J Am Soc Mass Spectrom* 9(8):805-13

Yasui Y, McLerran D, Adam BL, Winget M, Thornquist M, Feng Z (2003) An automated peak identification/calibration procedure for high-dimensional protein measures from mass spectrometers. *J Biomed Biotechnol* 2003(4):242-8

Yoko-o T, Wiggins CAR, Stolz J, Peak-Chew SY, Munro S (2003) An N-acetylglucosaminyltransferase of the Golgi apparatus of the yeast *Saccharomyces cerevisiae* that can modify N-linked glycans. *Glycobiology* 13(8):581-9

Yoshihashi-Suzuki S, Sato I, Awazu K (2008) Wavelength dependence of matrix-assisted laser desorption and ionization using a tunable mid-infrared laser. *Int J Mass Spectrom* 270(3):134-8

Zakett D, Schoen A, Cooks R, Hemberger P (1981) Laser-desorption mass spectrometry/mass spectrometry and the mechanism of desorption ionization. *J Am Chem Soc* 103(5):1295-7

Zambonelli C, Passarelli P, Rainieri S, Bertolini L, Giudici P, Castellari L (1997) Technological properties and temperature response of interspecific *Saccharomyces* hybrids. *J Sci Food Agric* 74(1):7-12

Zastrow C, Hollatz C, De Araujo P, Stambuk B (2001) Maltotriose fermentation by *Saccharomyces cerevisiae*. *J Ind Microbiol Biotechnol* 27(1):34-8

Zenobi R, Knochenmuss R (1998) Ion Formation in MALDI Mass Spectrometry. *Mass Spectrom Rev* 17:337-66

Zheng X, D'Amore T, Russell I, Stewart GG (1994) Transport kinetics of maltotriose in strains of *Saccharomyces*. *J Ind Microbiol* 13(3):159-66

Zlotnik H, Fernandez MP, Bowers B, Cabib E (1984) *Saccharomyces cerevisiae* mannoproteins form an external cell wall layer that determines wall porosity. *J Bacteriol* 159(3):1018-26

Zupfer JM, Churchill KE, Rasmusson DC, Fulcher RG (1998) Variation in ferulic acid concentration among diverse barley cultivars measured by HPLC and microspectrophotometry. *J Agric Food Chem* 46(4):1350-4

LIST OF PUBLICATIONS

Original Paper:

- Usbeck JC, Kern CC, Vogel RF, Behr J* (2013) Optimization of experimental and modeling parameters for the differentiation of beverage spoiling yeasts by Matrix-Assisted-Laser-Desorption/Ionization – Time-of-Flight Mass Spectrometry (MALDI-TOF MS) in response to varying growth conditions. *Food Microbiol* 36(2):379-87.
- Usbeck JC, Wilde C, Bertrand D, Behr J*, Vogel RF (2014) Wine yeast typing by MALDI-TOF MS*. Applied Microbiol Biotechnol*, 98(8):3737-52.
- Kern CC, Usbeck JC, Vogel RF, Behr J* (2013) Optimization of Matrix-Assisted-Laser-Desorption/Ionization-Time-Of-Flight Mass Spectrometry for the identification of bacterial contaminants in beverages. *J Microbiol Methods*, 93(3):185-91.

Additional scientific communications:

- Usbeck JC, Kern CC, Vogel RF, Behr J* (2012) Fast and reliable identification of beverage spoiling yeasts by MALDI-TOF MS. World Brewing Congress, Portland, USA, poster presentation.
- Usbeck JC, Behr J, Vogel RF* (2012) Differentiation of top- and bottom-fermenting brewing yeasts and insight their metabolic status by MALDI-TOF MS. World Brewing Congress, Portland, USA, poster presentation.
- Usbeck JC, Behr J, Vogel RF* (2013) Monitoring of yeast metabolism by Matrix-Assisted-Laser-Desorption/Ionization – Time-of-Flight Mass Spectrometry (MALDI-TOF MS). 26th International Conference of Yeast Genetics and Molecular Biology, Goethe University Frankfurt, Germany, poster presentation.
- Kern CC, Usbeck JC, Behr J, Vogel RF* (2012) Identification of bacterial contaminants in beverages by MALDI-TOF MS. World Brewing Congress, Portland, USA, poster presentation.
- Schnabel J, Lenz CA, Usbeck JC, Kern CC, Behr J, Vogel RF* (2012) Correlation of MALDI-TOF MS patterns and high pressure resistance of *Clostridium botulinum* type E endospores, 7th International Conference on High Pressure Bioscience and Biotechnology, Otsu, Japan, oral presentation.

INDEX OF FIGURES

INDEX OF TABLES

TABLE OF SUPPLEMENTS

APPENDIX

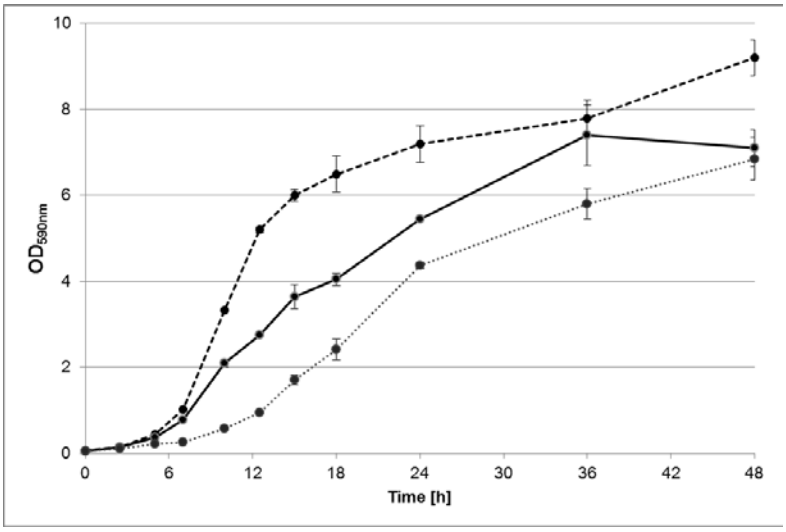

Fig. S1: The OD_{590nm} of TMW 3.236 (dashed), TMW 3.237 (solid) and TMW 3.238 (dotted) was determined simultaneously to the MALDI-TOF MS measurements, which were generated over a period of 2.5 to 48 h in W.

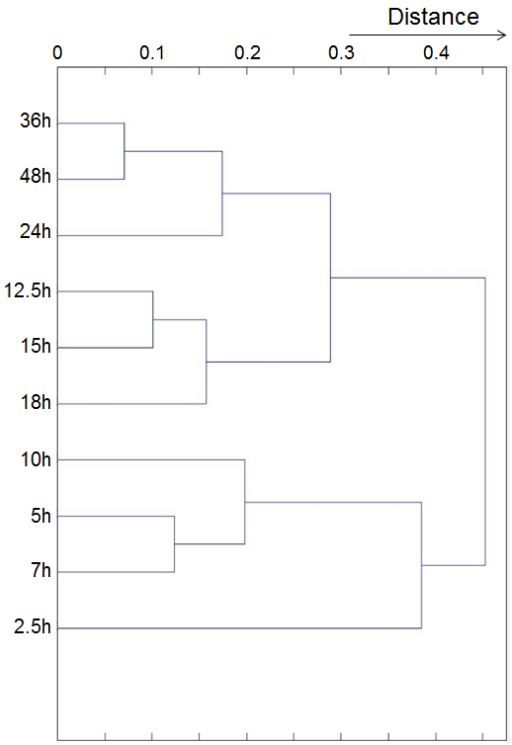

Fig. S2: Dendrogram of TMW 3.236 sum spectra at different growth phases in YM broth. Maximal distance level is 1.0 (analysis by in-house software).

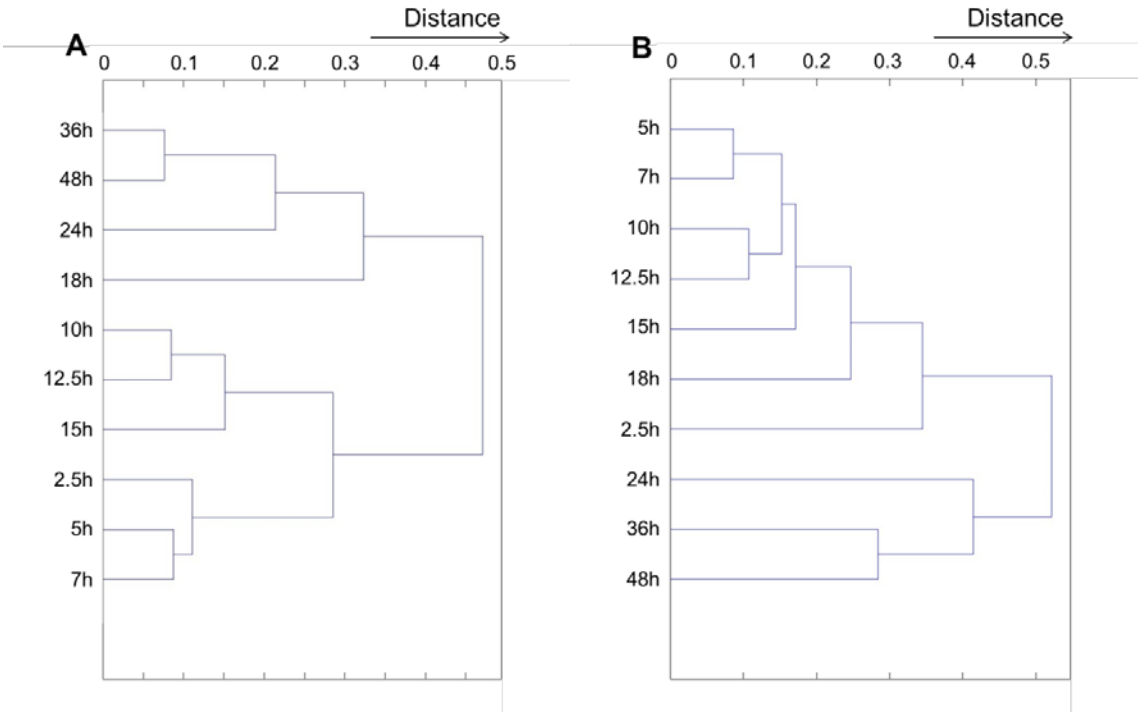

Fig. S3: Sum spectra dendrograms of (A) TMW 3.237 and (B) TMW 3.238 measured at different time points grown in W broth aerobically. Maximal distance level is 1.0 (analysis by in-house software).

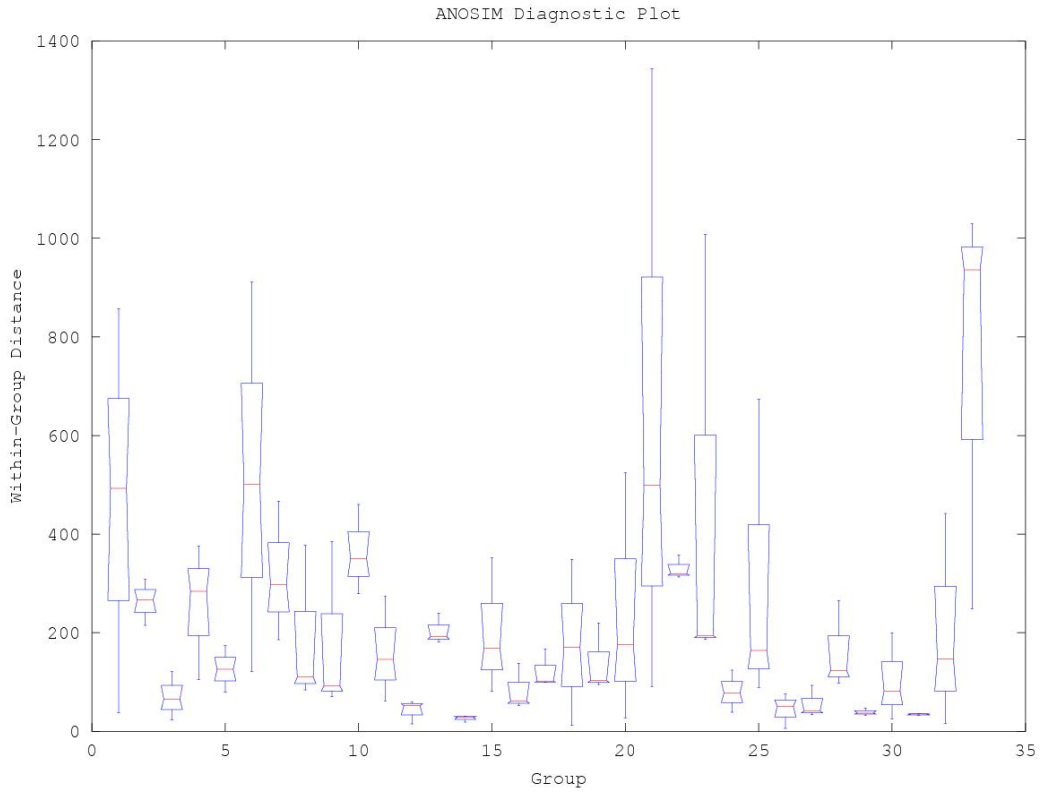

Fig. S4: ANOSIM boxplot of all 33 wine strains (x-axis). Ten single spectra were summarized employing MASCAP to one sum spectrum to illustrate the results clearly. The box corresponds to 50 % of all data. Additionally, the median is marked as the red line in between. The extending vertical lines (whiskers) indicate variability outside the upper and lower quartiles, whereas outliers may be plotted as individual points.

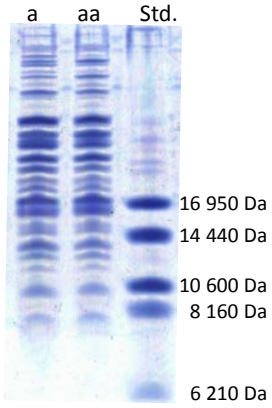

Fig. S5: SDS-PAGE of *S. cerevisiae* TMW 3.308 protein extract according to the MALDI-TOF MS protocol, cultures cultivated aerobically (a) and anaerobically (aa), Coomassie stained.

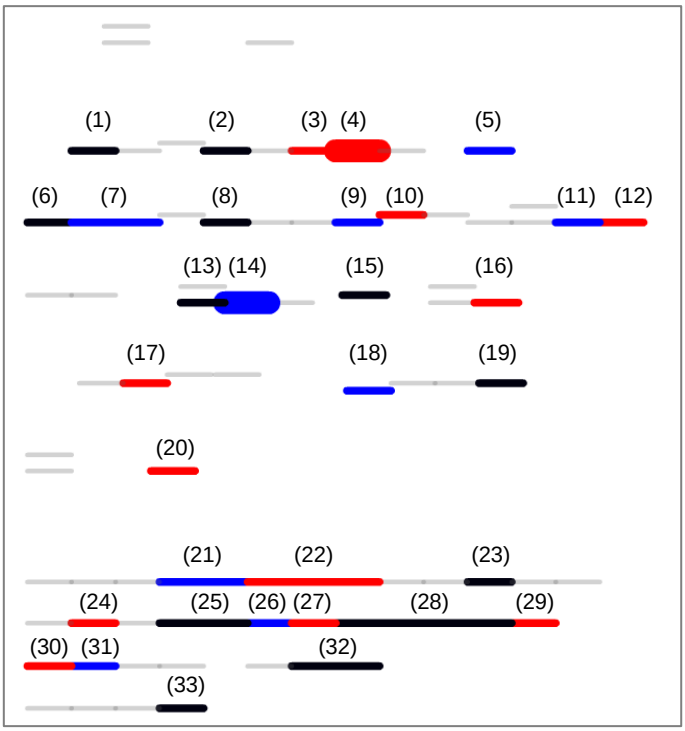

Fig. S6: Protein functions mapped on an interactive regulatory pathway of ribosomes of *S. cerevisiae* (by iPath explorer2). Displayed are the proteins or peptides belonging to masses detected via MALDI-TOF MS and GeLC-MS/MS under aerobic conditions (—), under anaerobic conditions (—), masses, which appear independently of the growth conditions (—) and the thin lines with the equal color code represent masses which could be detected with only one technique. (1) RPS20, (2) RPL32A, (3) RPS15, (4) RPL17, (5) RPL35, (6) RPS11, (7) RPL23/RPL26, (8) RPL11, (9) RPL9, (10) RPL32, (11) RPL7, (12) RPL27A, (13) RPS18, (14) RPS14, (15) RPL18, (16) RPS16, (17) RPS23, (18) RPLP1, (19) RPL12, (20) RPS13, (21) RPL21/RPL24, (22) RPL31/RPL35A/RPL37, (23) RPL40, (24) RPS6, (25) RPS17/RPS19, (26) (RPS24), (27) RPS25, (28) RPS26/RPS27/RPS27A/RPS28, (29) RPS30, (30) RPL6, (31) RPL18A), (32) RPL36/RPL38, (33) RPS21, with RP: ribosomal protein, L: large, S: small.