



Technische Universität München



Z I E L
Zentralinstitut für
Ernährungs- und
Lebensmittelforschung

Abteilung Mikrobiologie

Schlussbericht zum Forschungsprojekt B 95

Schnellidentifizierung von Kulturhefen sowie bier- und ge- tränkeschädlichen Hefen mittels FTIR-Spektroskopie

Zeitraum 01.02.2007 – 31.01.2009

Durchführende Forschungsstellen:

Dr. Mareike Wenning
Abteilung Mikrobiologie
Zentralinstitut für Ernährungs- und
Lebensmittelforschung (Z I E L)
Technische Universität München
Weihenstephaner Berg 3
85350 Freising
Tel.: 08161-71 2257
E-Mail: mareike.wenning@wzw.tum.de

Dr.-Ing. Mathias Hutzler*
Lehrstuhl für Technologie der Brauerei II
Wissenschaftszentrum Weihenstephan
Technische Universität München
Alte Akademie 3
85350 Freising

* aktuelle Adresse:
Forschungszentrum Weihenstephan für
Brau- und Lebensmittelqualität (BLQ)
Alte Akademie 3
85354 Freising

Tel.: 08161-71 3100
E-Mail: hutzler@wzw.tum.de

Inhaltsverzeichnis

1.	Zusammenfassung	3
2.	Einleitung und Zielsetzung	4
3.	Methodisches Vorgehen	7
3.1	Projektablauf und Koordination	7
3.2	FTIR-Spektroskopie	7
4.	Ergebnisse und Diskussion	9
4.1	Fremdhefen	9
4.1.1	<i>Identifizierung von Nicht-Saccharomyces-Fremdhefen</i>	9
4.1.1.1	OPUS-Datenbank	9
4.1.1.2	Identifizierung der Gattungen <i>Pichia</i> , <i>Issatchenkia</i> , <i>Kregervanrija</i> und <i>Williopsis</i> mit künstlichen neuronalen Netzen (KNN)	12
4.1.1.3	Identifizierungspotential der erweiterten Datenbanken für brauereirelevante Stämme	16
4.1.2	<i>Identifizierung von Saccharomyces-Fremdhefen</i>	17
4.2	Identifizierung von Kulturhefestämmen der Weihenstephaner Hefebank	22
	Exkurs: Veränderung von Kulturhefen über das Weiterführen in der Brauerei	25
4.3	Einfluss selektiver Inkubationsbedingungen auf die Identifizierung	27
5.	Fazit	29
	Literaturverzeichnis	30
	Anhang	31

1. Zusammenfassung

Die Identifizierung von Fremdhefen im Brauereibereich ist komplex und problematisch, da eine sehr große Artenvielfalt herrscht und zudem kaum Techniken existieren, mit denen sich Hefen universell, einfach und mit geringem Arbeits- und Kostenaufwand identifizieren lassen. Ziel des Projektes B95 war es daher, ein praxisnahes System zur Schnellidentifizierung von Kulturhefen und bier- und getränkeschädlichen Hefen (Nicht-*Saccharomyces*- und *Saccharomyces*-Fremdhefen) mittels FTIR-Spektroskopie zu entwickeln. Hierfür sollten Schadhefen und die Begleitflora aus Brauereien mit PCR-basierten Referenzmethoden identifiziert, ihre FTIR-Spektren gemessen und mit diesen die Systeme für die Identifizierung brauereirelevanter Hefen aufgebaut werden.

Da für die FTIR-Spektroskopie bereits einige umfangreiche Datenbanken existierten und diese sehr erfolgreich in der Diagnostik von Hefen aus dem Molkerei- und Futtermittelbereich eingesetzt werden, musste nur eine Datenbank neu konzipiert werden, drei weitere konnten auf bestehenden Systemen aufbauen. Aus dem Projekt resultierten eine universell einsetzbare Datenbank zur Identifizierung von Hefen, die nun 250 Arten aus 60 Gattungen umfasst, eine spezifische Datenbank zur Identifizierung von Hefen der Gattungen *Issatchenkia*, *Pichia*, *Williopsis* und *Kregervanrija*, eine Datenbank zur Identifizierung von Hefen des *Saccharomyces* sensu stricto / sensu lato-Komplexes und eine Datenbank zur genaueren Bestimmung von Kulturhefen. Insgesamt wurden im Verlauf des Projektes 298 Isolate gesammelt und identifiziert, wovon nach Aussortierung von Duplikaten (gleiche Stämme aus der gleichen Untersuchung) 201 Stämme für die Erstellung und Erweiterung der FTIR-Datenbanken genutzt wurden. 159 waren Nicht-*Saccharomyces*-Fremdhefen, 42 gehörten zur Gattung *Saccharomyces*. Insgesamt waren 58 unterschiedliche Arten aus 22 Gattungen im Datensatz enthalten. 16 der 58 Arten waren noch nicht Bestandteil einer Datenbank und zwei dieser Arten sind sogar bislang noch unbekannte Spezies. Hier zeigt sich deutlich, dass eine Adaption der Datenbanken an den Getränkektor notwendig und sinnvoll war, da die in diesem Habitat vorkommende Mikroflora in der Datenbank noch nicht ausreichend abgedeckt war, obwohl diese sehr stark an den Lebensmittelsektor angepasst war.

Die zwei spezifischen Datenbanken zur Identifizierung von Arten der Gattungen *Issatchenkia*, *Pichia*, *Williopsis* und *Kregervanrija* sowie des *Saccharomyces* sensu stricto / sensu lato-Komplexes hatten in ihrer ursprünglichen Fassung nur eine moderate Leistungsfähigkeit, wie durch die Identifizierung der im Rahmen dieses Projektes gesammelten Isolate festgestellt wurde. Nach der Erweiterung hingegen überzeugten Sie mit einer sehr hohen Zuverlässigkeit von jeweils ca. 93 % und sind damit für den Einsatz im Brauereisektor sehr gut geeignet. Etwas mehr Schwierigkeiten bereitete das System zur Identifizierung von Kulturhefen. So war die Entwicklung einer Differenzierung von *Saccharomyces*-Kulturhefen von Fremdhefen der Arten *S. cerevisiae* und *S. pastorianus* nicht möglich, da hierfür eine Stammtypisierung der während des Projektes gesammelten Isolate hätte erfolgen müssen, was vom Arbeitsaufwand her allerdings nicht geleistet werden konnte. So sollte eine Datenbank zur Identifizierung von Kulturhefestämmen entwickelt werden, was jedoch aufgrund zu geringer Robustheit der Methode nicht vollständig gelang. Die entstandene Referenzdatenbank leistet aber eine sehr zuverlässige Differenzierung der obergärigen und untergärigen Stämme sowie eine Unterteilung der obergärigen Hefen in die drei Klassen Weißbier, Alt/Kölsch und Wein/Sekt/Brennerei.

Ungeachtet der Schwierigkeiten bei den Kulturhefen wurde im vorliegenden Projekt die FTIR-Spektroskopie für die Identifizierung von Fremdhefen im Brauereibereich optimiert und kann nun mit großer Zuverlässigkeit eingesetzt werden.

2. Einleitung und Zielsetzung

In der Getränkemikrobiologie stellen bierschädliche Bakterien und Fremdhefen die Hauptproblemkeime während und nach den Produktionsprozessen dar. Das Segment der bierschädlichen Bakterien wird in der mikrobiellen Analytik sowohl mit Kulturanreicherungen als auch mit modernen molekularbiologischen Schnellnachweismethoden sehr gut und praxisnah abgedeckt. Für die Fremdhefe Problematik ist die Problemstellung komplexer, da die Artenvielfalt an bier- und getränkeschädlichen Hefen deutlich größer als die der bierschädlichen Bakterien ist und zudem durch neue Produktentwicklungen der Brauereien, wie z. B. Biermischgetränke, leichte, alkoholfreie Biere und alkoholfreie Getränke auf Malzbasis, andere mikrobielle Anfälligkeiten und Schadkeimflora als in traditionellem Bier entstehen. Neuartige Getränke mit höheren Zuckerkonzentrationen bieten Schadhefen besonders gute Wachstumsbedingungen und sind aufgrund dessen anfälliger für hefebedingten Produktverderb. Des Weiteren erschwert die nahe Verwandtschaft der *Saccharomyces*-Fremdhefen mit den Brauereikulturen deren Identifizierung und ist bisher nur durch arbeits- und kostenintensive molekularbiologische Methoden zu bewerkstelligen.

Die FTIR-Spektroskopie ist eine etablierte, kostengünstige und schnelle Methode zur artgenauen Identifizierung von Mikroorganismen (Wenning et al., 2008) und kann z. T. auch zur Typisierung unterhalb der Artgrenze genutzt werden. Die Absorption der Zellbestandteile Proteine, Fettsäuren, Polysaccharide, etc. im mittleren Bereich des infraroten Lichts ergibt charakteristische, reproduzierbare Spektren, die unter Zuhilfenahme von Referenzdatenbanken für eine Identifizierung genutzt werden (Kümmerle et al., 1998, Naumann et al., 1991, Oberreuter et al., 2002, Wenning et al., 2002). Die Sensitivität der Methode ist dabei so hoch, dass auch auf Subspezies- oder Stammebene differenziert werden kann (Büchl et al., 2010, Sandt et al., 2003, Timmins et al., 1998, Wenning et al., 2002). FTIR-Spektroskopie bietet neben einer schnellen und äußerst kostengünstigen Identifizierung zusätzlich die Möglichkeit, die Datenbanken der Referenzspektren individuell an das lebensmittelverarbeitende Unternehmen anzupassen und damit über eine Identifizierung auf Stammebene möglicherweise auch eine Zuordnung der isolierten Organismen zu einem bestimmten Ursprung vorzunehmen (Kontaminationsrouten-Analyse).

Für die Zusammenstellung der Spektren zu Datenbanken und die Identifizierung unbekannter Spektren steht die zum Spektrometer gehörende Software OPUS zur Verfügung. Sie identifiziert unbekannte Spektren anhand berechneter Distanzwerte (Helm et al., 1991) zu den in der Referenzdatenbank enthaltenen Referenzspektren und gibt das Ergebnis als Hitliste aus (Abb. 1).

Hit No.	Hit Quality	Sample Name	File Name	Threshold
1	0.266401	Pichia anomala G-/ze010823 309-1	PicAnoZe153091.1	0.000000
2	0.284668	Pichia anomala BLQ95-21 9 A5	PicAnoB95-21_9.50	0.000000
3	0.405984	Pichia anomala G-/on020208b 6-1	PicAnoOn20208b61.1	0.000000
4	0.437609	Pichia anomala W39/G10/CB55759	PicAnoC5759.101	0.000000
5	0.444110	Pichia anomala G-/st020412 1-1	PicAnoSt2041211.1	0.000000
6	0.450343	Pichia anomala W295/G1651/DSM70130	PicAnoD70130.51	0.000000
7	0.458330	Pichia anomala G-/dr020130a 1-1	PicAnoDr20130a11.1	0.000000
8	0.460164	Pichia anomala G2111/sl990215 18-1	PicAnoG2111.100	0.000000
9	0.469090	Pichia anomala G2935/Mietke(-)	PicAnoG2935.55	0.000000
10	0.515806	Pichia anomala G1438/ze960902a	PicAnoG1438.1	0.000000
11	0.516889	Pichia anomala ! G2042/AHL4763	PicAnoG2042!5	0.000000
12	0.523282	Pichia anomala G-/ba020513a 11-1	PicAnoBa20513111.1	0.000000

Abbildung 1: Identifizierungsreport (Hitliste) für ein Hefeisolat, identifiziert mit der OPUS-Datenbank.

Die Hitliste stellt eine Auflistung der dem unbekanntem Spektrum ähnlichsten Referenzspektren dar und wird im Wesentlichen anhand der Reihenfolge der Artnamen beurteilt. Als zweites Kriterium dienen die berechneten Distanzwerte, in der Hitliste als Hitquality angegeben, die für eine akzeptable Identifizierung unter 1,5 liegen sollten. Das in Abbildung 1 dargestellte Beispiel zeigt die Identifizierung eines Hefeisolats, das anhand der Hitliste eindeutig als *Pichia anomala* identifiziert werden kann. Die OPUS-Datenbanken sind einfach zu handhaben und zu erweitern, weil sie keine Kalibrierung enthalten, die bei jeder Erweiterung aktualisiert werden müsste. Sie können somit auch von Nutzern bedient werden, die keine Erfahrung mit statistischer Datenauswertung haben, und sind aufgrund dessen auch gut für den Einsatz in Industrielaboren geeignet. Allerdings stoßen die OPUS-Datenbanken bei der Differenzierung eng verwandter Arten oder Stämme an die Grenzen ihrer Leistungsfähigkeit, da die Berechnung eines einzigen Distanzwertes hier nicht diskriminativ genug ist. Verdeutlicht wird dies in Abbildung 2, die ein nicht eindeutig zu interpretierendes Identifizierungsergebnis für ein weiteres Hefeisolat darstellt.

Hit No.	Hit Quality	Sample Name	File Name
1	0.249930	<i>Pichia galeiformis</i> ! G2939/Mietke(-)	PicGalG2939!.511
2	0.254807	<i>Issatchenkia scutulata</i> v. <i>exigua</i> W543/G2791/CB58068	IssScu-ExiC8068.12
3	0.292603	<i>Pichia deserticola</i> / <i>Candida ethanolica</i> ! G2611/Mietke(feed)	PicDesCanEthG2611!.5
4	0.305945	<i>Pichia kluyveri</i> G2908/Mietke(-)	PicKluG2908.59
5	0.314422	<i>Pichia deserticola</i> / <i>Candida ethanolica</i> ! G2612/Mietke(feed)	PicDesCanEthG2612!.52
6	0.317872	<i>Pichia galeiformis</i> ! G3019/BfEL5688	PicGalG3019!.511
7	0.326158	<i>Pichia deserticola</i> / <i>Candida ethanolica</i> ! G2615/Mietke(feed)	PicDesCanEthG2615!.13
8	0.339501	<i>Pichia membranifaciens</i> BLQ95-31 9 G7	PicMemB95-31_9.50
9	0.339622	<i>Pichia deserticola</i> / <i>Candida ethanolica</i> ! G2610/Mietke(feed)	PicDesCanEthG2610!.57
10	0.343072	<i>Pichia cactophila</i> CBS620	PicCacC620.52
11	0.349485	<i>Pichia deserticola</i> / <i>Candida ethanolica</i> ! G2611/Mietke(feed)	PicDesCanEthG2611!.57
12	0.353158	<i>Pichia deserticola</i> / <i>Candida ethanolica</i> ! G2609/Mietke(feed)	PicDesCanEthG2609!.52

Abbildung 2: Identifizierungsreport (Hitliste) für ein Hefeisolat identifiziert, mit der OPUS-Datenbank.

Es fällt auf, dass die Hitliste eine Mischung unterschiedlicher Arten zeigt, die noch dazu alle sehr ähnliche Hit Qualities haben. Eine zuverlässige Identifizierung ist hier also nicht möglich.

Um dennoch auch zwischen eng verwandten Arten differenzieren zu können, kann alternativ auf andere statistische Verfahren zurückgegriffen werden, von denen sich für besonders komplexe Klassifizierungsprobleme vor allem künstliche neuronale Netze (KNN) eignen (Maquelin et al., 2003, Udelhoven et al., 2000, Wenning et al., 2010). KNN gehören zu den Systemen der künstlichen Intelligenz und können im Rahmen eines Trainingsprozesses (Kalibrierung) anhand eines Referenzdatensatzes die Zuordnung von Daten zu vorgegebenen Klassen (z.B. Gattungen oder Arten) erlernen. Während des Trainings wird die Verarbeitung der Eingabedaten (Absorptionswerte) so optimiert, dass eine eindeutige und korrekte Klassenzuweisung erfolgt. Zur Lösung komplexer Probleme, wie der Differenzierung einer großen Zahl an Klassen, kann ein hierarchisch strukturierter Identifizierungsbaum erstellt werden, bei dem zunächst eine Differenzierung in größere Gruppen und in der nächsten Ebene eine feinere Aufschlüsselung jeder Gruppe erfolgt (siehe Abb. 3). Für die Erarbeitung von KNN zur Identifizierung von FTIR-Spektren existiert eine spezialisierte Software (NeuroDeveloper[®]) der Firma Synthon in Heidelberg (Udelhoven et al., 2003), die alle notwendigen Schritte und Methoden abdeckt und auch im Rahmen dieses Projektes Verwendung fand.

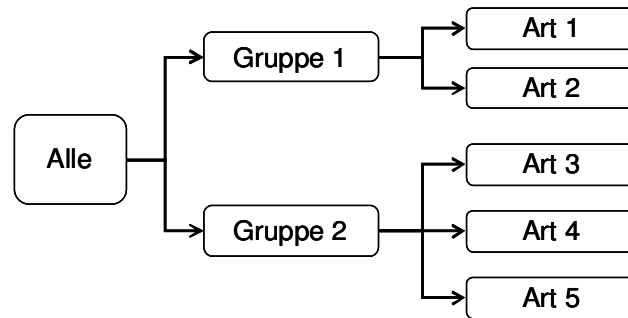


Abbildung 3: Schematische Darstellung eines hierarchisch strukturierten KNN.

Die Ergebnisse werden hier nicht mehr als Trefferliste ausgegeben, die dann noch interpretiert werden muss, sondern das als Ergebnis erscheint die Klasse, zu der das unbekannte Spektrum aufgrund der Kalibrierung zugeordnet wird (Abb. 4).

Names	Class	Classification	P...
● PicGalB95-186_9_B7.0	2.2.3.3	Pichia galeiformis	C:...

Abbildung 4: Identifizierungsreport für ein Hefeisolat identifiziert mit einem KNN.

Für die Identifizierung von Hefen bestand bereits vor Beginn des B95-Projektes eine sehr umfangreiche Referenzdatenbank basierend auf der OPUS-Software, die von der Abteilung Mikrobiologie des ZIEL über die vergangenen Jahre entwickelt wurde (Büchl et al., 2008, Kümmerle et al., 1998, Wenning et al., 2002) und die zu Beginn des B95-Projektes Spektren von 178 Arten aus 44 Gattungen enthielt. Auch ein KNN zur Identifizierung eng verwandter Arten der drei Gattungen *Pichia*, *Issatchenkia* und *Williopsis* war bereits im Rahmen eines weiteren Projektes entstanden (Büchl et al., 2008) und diente als Ausgangsbasis für die Erweiterung und Anpassung auf die Hefeflora der Brauereiindustrie. Als drittes wurde bereits im Vorfeld dieses Forschungsvorhabens in Kooperation mit einer Brauerei mit der Erstellung eines KNNs für die Identifizierung von Hefen des *Saccharomyces sensu stricto* / *sensu lato*-Komplexes begonnen.

Ziel des Projektes B95 war es, ein praxisnahes System zur Schnellidentifizierung von Kulturhefen und bier- und getränkeschädlichen Hefen (Nicht-*Saccharomyces*- und *Saccharomyces*-Fremdhefen) mittels FTIR-Spektroskopie zu entwickeln. Hierfür sollten Schadhefen und die Begleitflora aus Brauereien mit PCR-basierten Referenzmethoden identifiziert, ihre FTIR-Spektren gemessen und mit diesen die bestehenden Identifizierungssysteme an die Identifizierung brauereirelevanter Hefen angepasst werden. Zudem sollte die Differenzierung von Kulturhefestämmen der Weihenstephaner Hefebank auf Basis von KNN etabliert werden. Dieser Teil der Zielvorgabe stellte die höchste Herausforderung dar, da Kulturhefen noch näher untereinander verwandt sind als die Arten der *Saccharomyces*-Gruppe und oft voneinander abstammen. Als Teilaspekt sollte weiterhin betrachtet werden, ob die Kultivierung von Hefen auf gängigen Selektivmedien der Brauereiindustrie (Lysinagar, Kristallviolettagar, etc.) die Identifizierung über FTIR-Spektroskopie beeinträchtigt.

3. Methodisches Vorgehen

3.1 Projektablauf und -koordination

Abbildung 5 skizziert die allgemeine methodische Vorgehensweise. Fremdhefen aus verschiedenen Brauereien und verschiedenen Produktionsbereichen wurden vom Forschungszentrum für Brau und Lebensmittelqualität (BLQ), Weihenstephan zur Verfügung gestellt. Diese Sammlung wurde kontinuierlich mit aktuellen Praxisisolaten erweitert. Um sicherzustellen, dass die Hefestämme für die PCR-Referenzidentifizierung und die FTIR-Messung als Reinkulturen vorliegen, wurden diese vereinzelt, katalogisiert und in die Cryobank-Stammsammlung des Lst. BT II und des Forschungszentrums Weihenstephan integriert.

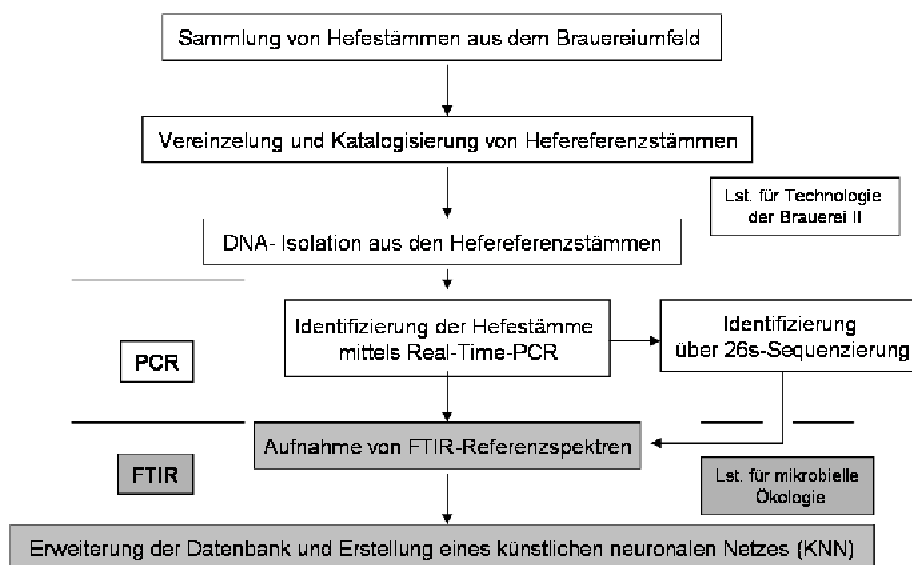


Abbildung 5: Allgemeines Schema zur Sammlung, Identifizierung, Referenzspektrenaufnahme und Datenbankgestaltung von brauereirelevanten Hefen.

Anschließend wurde von den Reinkulturen die DNA isoliert und die Hefen mit PCR-basierten Methoden (Real-Time-PCR, DNA-Sequenzierung) analysiert bzw. identifiziert. Die identifizierten Hefestämme wurden anschließend mit FTIR-Spektroskopie gemessen, wobei für die Erstellung der KNNs 10-30 unabhängige Spektren für jedes Isolat erzeugt wurden. Die Spektren aller Arten wurden in die bestehende OPUS-Datenbank integriert, die die umfassendste Datenbank ist und alle Arten enthält. Fremdhefen, die den Gattungen *Pichia*, *Issatchenkia*, *Williopsis* und *Kregervanrija* angehörten oder zum *Saccharomyces*-Komplex (inkl. der Gattungen *Kazachstania*, *Naumovia* und *Lachancea*) zählten, wurden zudem in das jeweilige KNN-System integriert.

3.2 FTIR-Spektroskopie

Die FTIR-Spektroskopie ist eine Technik, mit der Mikroorganismen über ihre biochemischen Fingerabdrücke identifiziert werden, wobei die Zellen der Organismen für die Messung nicht zerstört werden, sondern in intaktem Zustand präpariert werden (siehe Abbildung 6). Ausgehend von einem frischen Drei-Ösen-Ausstrich wird eine halbe Impföse (1 mm Durchmesser) Zellmaterial abgenommen und mit Hilfe eines Drigalski-Spatels auf YGC-Agar (Merck) aus-

plattiert. Die Inkubation erfolgt für 24 h bei 27 °C. Für die Messung wird mit einer kleinen Impfföse etwas Material aus der Mitte des entstandenen konfluenten Zellrasens abgenommen und in 100 µl autoklaviertem Wasser suspendiert. Von dieser Suspension werden 25 µl je Feld auf einen ZnSe-Probenträger aufgetragen und anschließend etwa 45 Minuten bei 40 °C getrocknet. Die Messung erfolgt direkt anschließend mit dem Tensor 27 Spektrometer gekoppelt an das HTS-XT-Modul für hohen Probendurchsatz (beide von Bruker Optik, Ettlingen) und dauert etwa eine Minute pro Probe.

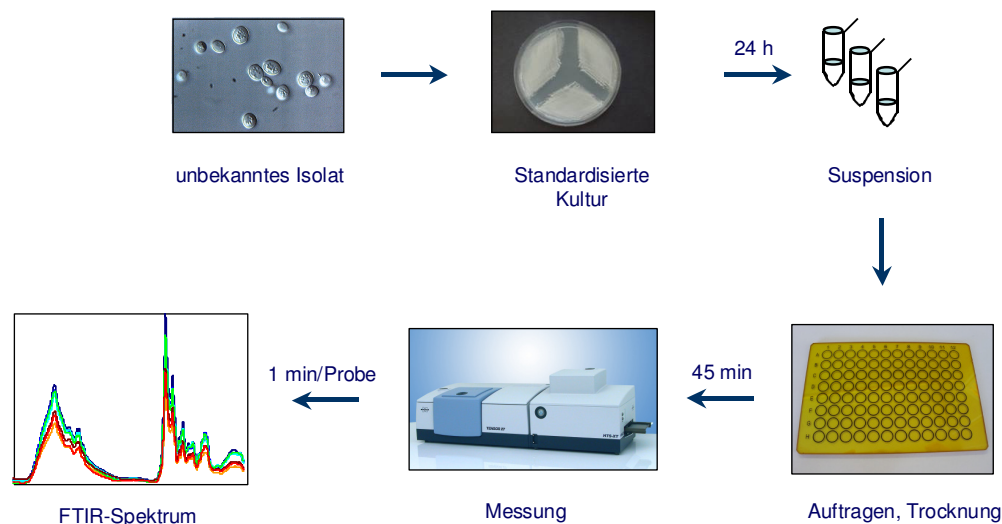


Abbildung 6: Schema für die Probenpräparation zur FTIR-Messung von Hefen.

Während der Messung im Transmissionsmodus durchströmt infrarotes Licht die getrockneten Zellen und alle in der Zelle vorhandenen Moleküle werden über ihre spezifische Absorption des IR-Lichts quantitativ erfasst. Das Resultat sind hochspezifische Fingerprints, die sich nicht nur für eine Artidentifizierung eignen, sondern auch auf Stammebene typisiert werden können. Allerdings geht die sehr hohe Spezifität mit einer erhöhten Sensitivität gegenüber Änderungen am Präparationsprotokoll einher, da Abweichungen in den Wachstumsbedingungen (z.B. Art des Mediums, Inkubationstemperatur oder -dauer) zu einer veränderten Zellzusammensetzung der Organismen führen, die sich in der Folge auch in veränderten Spektren niederschlägt und unter Umständen die Identifizierung beeinträchtigen kann. Dies bedeutet, dass die FTIR-Spektroskopie eine niedrigere Robustheit als beispielsweise genetische Methoden hat, da ein veränderter physiologischer Zustand der Organismen zu Abweichungen in den Ergebnissen führen kann. Bei strikt standardisiertem Protokoll sind jedoch sehr zuverlässige Ergebnisse zu erzielen und die Methode hat sich seit Jahren im Betrieb des mikrobiologischen Labors der Abt. Mikrobiologie des ZIEL bewährt.

4. Ergebnisse und Diskussion

Insgesamt wurden während der Projektlaufzeit 298 Isolate aus laufenden Untersuchungen des BLQ gesammelt und identifiziert. Nach Aussortierung von Duplikaten (gleiche Stämme aus der gleichen Untersuchung) wurden 201 Stämme für die Erstellung und Erweiterung der FTIR-Datenbanken genutzt (siehe Tabellen 1, 2 und 7). 159 waren Nicht-*Saccharomyces*-Fremdhefen, 42 gehörten zur Gattung *Saccharomyces*. Insgesamt waren 58 unterschiedliche Arten aus 22 Gattungen im Datensatz enthalten, was bereits deutlich macht, dass eine sehr breite Vielfalt unterschiedlicher Hefen im Umfeld und den Produkten der Brau- und Getränkeindustrie zu finden ist und zur Identifizierung ein System benötigt wird, das auf viele Arten gleichzeitig angewendet werden kann.

4.1 Fremdhefen

4.1.1 Identifizierung von Nicht-*Saccharomyces*-Fremdhefen

159 Isolate von Nicht-*Saccharomyces*-Hefen (Tabellen 1 und 2) wurden gesammelt und mit real-time PCR oder DNA-Sequenzierung identifiziert. Anschließend wurde getestet, ob sie mit den zur Verfügung stehenden FTIR-Referenzdatenbanken identifiziert werden konnten. Zunächst wurden alle Isolate mit der umfangreichsten Datenbank basierend auf der OPUS-Software des Spektrometerherstellers identifiziert. Stämme, die zu den Gattungen *Pichia*, *Issatchenkia* und *Williopsis* gehörten, wurden danach mit dem dafür existierenden KNN klassifiziert. Die Ergebnisse sollten Aufschluss darüber geben, wie gut die bereits vorhandenen Datenbanken auf Stämme aus dem Brauereiumfeld anzuwenden sind und wo Erweiterungsbedarf besteht.

4.1.1.1 OPUS-Datenbank

Die OPUS-Datenbank umfasste insgesamt 178 Arten aus 44 Gattungen und bestand zum wesentlichen Teil aus Referenzspektren von Feldisolaten, die im Rahmen von Lebensmitteluntersuchungen gewonnen wurden. Tabellen 1 und 2 geben einen Überblick über die im Verlauf des vorliegenden Projektes gesammelten Stämme aufgeschlüsselt nach Arten, die in der Referenzdatenbank bereits enthaltenen Referenzstämme für diese Arten und den Anteil der Isolate, die mit dieser Datenbank auf Artebene korrekt identifiziert werden konnten. Identifiziert wurden nur solche Arten, für die mindestens zwei Referenzstämme in der Datenbank vorlagen, da nur dann ein aussagekräftiges Ergebnis zu erwarten war.

Tabelle 1: Gesammelte Stämme an Nicht-*Saccharomyces*-Fremdhefen (ohne die Gattungen *Pichia*, *Issatchenkia*, *Williopsis* und *Kregervanrija*) sowie die zugehörigen Referenzstämme in der Datenbank und der Anteil der korrekt identifizierten Isolate. sp.: Spezies, sp. nov: neue, bislang unbekannte Art.

Art	Stämme		
	neu aus B95	Referenzen in der Datenbank	Anteil korrekt identifizierter Isolate
<i>Aureobasidium pullulans</i>	5	11	60 %
<i>Candida bituminiphila</i>	1	0	-
<i>Candida boidinii</i>	4	8	75 %
<i>Candida cretensis</i>	1	0	-
<i>Candida inconspicua</i>	2	5	0 %
<i>Candida intermedia</i>	6	23	33 %
<i>Candida mesenterica</i>	1	0	-
<i>Candida norvegica</i>	1	0	-

Art	Stämme		
	neu aus B95	Referenzen in der Datenbank	Anteil korrekt identifizierter Isolate
<i>Candida oleophila</i>	1	4	100 %
<i>Candida parapsilosis</i>	4	39	100 %
<i>Candida pararugosa</i>	3	9	66 %
<i>Candida picinguabensis</i>	2	9	100 %
<i>Candida pseudolambica</i>	1	0	-
<i>Candida sake</i>	8	6	25 %
<i>Candida sophiae-reginae</i>	1	0	-
<i>Candida sorbophila</i>	1	7	100 %
<i>Candida</i> sp. nov.	3	-	-
<i>Candida vartiovaarae</i>	1	2	0 %
<i>Clavispora lusitaniae</i>	3	24	100 %
<i>Cryptococcus adeliensis</i>	1	1	-
<i>Cryptococcus albidosimilis</i>	1	0	-
<i>Cryptococcus albidus</i>	2	13	0 %
<i>Cryptococcus diffluens</i>	1	3	0 %
<i>Cryptococcus heveanensis</i>	2	0	-
<i>Cryptococcus magnus</i>	5	8	40 %
<i>Cryptococcus saitoi</i>	2	0	-
<i>Debaryomyces hansenii</i>	4	77	100 %
<i>Dekkera bruxellensis</i>	2	10	0 %
<i>Hanseniaspora uvarum</i>	2	18	100 %
<i>Kluyveromyces marxianus</i>	1	48	0 %
<i>Lodderomyces elongisporus</i>	1	0	-
<i>Pseudozyma</i> sp.	1	1	-
<i>Rhodosporidium fluviale /azoricum</i>	1	0	-
<i>Rhodotorula laryngis</i>	2	0	-
<i>Rhodotorula mucilaginosa</i>	6	9	0 %
<i>Rhodotorula slooffiae</i>	1	0	-
<i>Schizosaccharomyces pombe</i>	1	4	100 %
<i>Sporidiobolus salmonicolor</i>	2	1	-
<i>Torulaspora delbrueckii</i>	4	33	100 %
<i>Trichosporon coremiiforme</i>	1	0	-
<i>Yarrowia lipolytica</i>	2	20	100 %
Gesamt	94	393	54 %

Aus Tabelle 1 ist ersichtlich, dass die detektierte Diversität sehr hoch ist und es nur wenige Arten gibt, die mit einer großen Anzahl an Isolaten gefunden wurden (z.B. *A. pullulans*, *C. intermedia*, *C. sake*, *Cr. magnus* oder *R. mucilaginosa*). Von sehr vielen Arten hingegen wurden nur ein oder zwei Isolate gewonnen. Drei Isolate konnten keiner Art zugeordnet werden, da sie zu zwei *Candida*-Arten gehören, die bislang noch nicht bekannt waren und deshalb auch noch keinen Namen tragen (bezeichnet mit sp. nov., lateinisch für neue Art). Auch dies zeigt, dass die Biodiversität der im Brauereibereich vorkommenden Hefen noch nicht zur Genüge erforscht ist.

Auffällig ist auch, dass neben diesen zwei noch nicht beschriebenen *Candida*-Arten 14 weitere Spezies gefunden wurden, die in der Referenzdatenbank bislang noch nicht abgelegt waren, obwohl diese stark auf den Lebensmittelsektor ausgelegt war. Für die Arten, die bereits mit mindestens zwei Referenzen Bestandteil der Bibliothek waren, betrug die Identifizierungsrate auf Artebene 54 %. Rechnet man die neuen Arten und Spezies, die nicht im Referenzdatensatz enthalten sind, in die Quote der identifizierbaren Stämme ein, können statt 54 % nur 40 % korrekt identifiziert werden. Hier zeigt sich deutlich, dass eine Adaption der Datenbanken an den Getränkektor notwendig ist, da die in diesem Habitat vorkommende Mikroflora in der Datenbank noch nicht ausreichend abgedeckt ist. Doch auch Arten, für deren Identifizierung eine ausreichende Zahl an Referenzstämmen zur Verfügung stand (beispielsweise *C. intermedia*, *Cr. albidus*, *D. bruxellensis* oder *K. marxianus*), konnten nicht immer zufrieden stellend identifiziert werden, was daran liegt, dass die Zellzusammensetzungen dieser Isolate und somit auch die aufgenommenen FTIR-Spektren Unterschiede zu denen der Referenzen aufweisen, was die Identifizierbarkeit einschränkt. Das Habitat Brauereindustrie zeichnet sich gegenüber dem Molkereisektor also nicht nur durch ein anderes Artenspektrum aus, sondern auch die Spezies, die in beiden Bereichen vorkommen, sind an die durch sie besiedelte Nische so angepasst, dass es sich in ihren FTIR-Fingerprints niederschlägt. Eine Integration der im Rahmen des vorliegenden Projektes isolierten Stämme in die Datenbank wird hier eine sehr sinnvolle und signifikante Ergänzung sein und nicht nur die Breite der Datenbank (Anzahl der abgelegten Arten), sondern auch deren Tiefe (Zahl der Stämme pro Art) für viele Arten verbessern.

Doch trotz der Schwierigkeiten bei der exakten Artidentifizierung lieferte die Datenbank auch in nicht angepasstem Zustand durchaus gut verwertbare Ergebnisse, da in fast allen Fällen eine Identifizierung auf Gattungsebene möglich war und die Identifizierung oftmals auf zwei oder drei Arten der gleichen Gattung eingeschränkt werden konnte. Dieses sind für die Praxis bereits nützliche Resultate, da sehr ähnliche Arten auch ähnliche Eigenschaften haben und sich auch anhand solcher Identifizierungen sinnvolle Schlussfolgerungen, Bewertungen und ggf. zu ergreifende Maßnahmen ableiten lassen. Gerade die Tatsache, dass in der Datenbank viele Arten abgelegt sind, die mit anderen Identifizierungsmethoden nicht erfasst werden, ist hier ein entscheidender Vorteil.

Ein ähnliches Bild für die Identifizierbarkeit mit der OPUS-Datenbank ergibt sich auch für Isolate aus den Gattungen *Pichia*, *Issatchenkia*, *Williopsis* und *Kregervanrija* (Tabelle 2). Die Gattung *Pichia* ist sicherlich eine der wichtigsten im Lebensmittelsektor und stellt auch in der hier vorliegenden Studie mit *P. anomala*, *P. guilliermondii* und *P. membranifaciens* die drei am häufigsten isolierten Arten. Die Rate der identifizierbaren Stämme ist jedoch auch hier mit 56 % nicht zufrieden stellend. Auffällig ist, dass *P. guilliermondii* als eine der wichtigsten Arten zu 100 % identifiziert werden konnte, wohingegen *P. anomala* und insbesondere *P. membranifaciens* Schwierigkeiten bereiteten. Allerdings ist für diese Gruppe von Hefen eine Identifizierung auf Artebene auch besonders schwierig, da die einzelnen Arten sehr eng miteinander verwandt sind und die auf OPUS basierende Datenbank dieses Artgefüge nicht mit der nötigen Schärfe auftrennen kann. Aus diesem Grund kamen für eine verbesserte Identifizierung dieser Gruppe künstliche neuronale Netze zum Einsatz (siehe Kap. 4.1.1.2).

Tabelle 2: Gesammelte Stämme an Nicht-*Saccharomyces*-Fremdhefen der Gattungen *Pichia*, *Issatchenkia*, *Williopsis* und *Kregervanrija* sowie die zugehörigen Referenzstämme in der Datenbank und der Anteil der korrekt identifizierten Isolate.

Art	Stämme		
	neu aus B95	Referenzen in der Datenbank	Anteil korrekt identifizierter Isolate
<i>Issatchenkia occidentalis</i>	1	14	100 %
<i>Issatchenkia orientalis</i>	2	139	100 %
<i>Kregervanrija delftensis</i>	3	0	-
<i>Kregervanrija fluxuum</i>	1	4	0 %
<i>Pichia anomala</i>	20	36	50 %
<i>Pichia fabianii</i>	1	7	0 %
<i>Pichia fermentans</i>	1	29	100 %
<i>Pichia galeiformis</i>	3	6	33 %
<i>Pichia guilliermondii</i>	13	39	100 %
<i>Pichia membranifaciens</i>	17	25	35 %
<i>Pichia norvegensis</i>	1	10	0 %
<i>Williopsis californica</i>	2	11	50 %
Gesamt	65	320	56 %

Zur Erweiterung der Datenbank wurden alle 159 Isolate der Nicht-*Saccharomyces*-Hefen in den Referenzdatensatz integriert und die Datenbank somit signifikant an den Brauereisektor angepasst.

Zusatzbemerkung: Innerhalb der Gattungen *Pichia* und *Williopsis* wurden in den letzten Jahren einige Umbenennungen vorgenommen (Kurtzman et al., 2008), die jedoch noch nicht in die Datenbanken eingearbeitet wurden. In Anhang 3 findet sich eine Übersicht, welche Arten neu benannt wurden und wie die aktuelle Bezeichnung lautet.

4.1.1.2 Identifizierung der Gattungen *Pichia*, *Issatchenkia*, *Kregervanrija* und *Williopsis* mit künstlichen neuronalen Netzen (KNN)

Für die Identifizierung von Nicht-*Saccharomyces*-Schadhefen der Gattungen *Pichia*, *Issatchenkia* und *Williopsis* wurde bereits vor Beginn des B95-Projektes an der Erstellung eines KNN gearbeitet, das vor allem für die Bereiche Milchprodukte und tierische Futtermittel konzipiert war. Da sich beide Projekte zeitlich überschneiden, wurden zu Beginn des B95-Projektes 15 *Pichia*- und *Issatchenkia*-Stämme als Referenzen in das zu trainierende KNN integriert (siehe Tabelle 4, 1. Datensatz). Das resultierende KNN enthielt 245 Stämme der in Tabelle 3 gelisteten Arten und bestand aus einem Identifizierungsbaum mit 22 Klassen (für jede Art eine) verteilt auf vier hierarchisch organisierte Ebenen (Abb. 6).

Tabelle 3: Zur Erstellung des KNN zur Identifizierung von Hefen der Gattungen *Pichia*, *Issatchenkia* und *Williopsis* verwendete Referenzstämme.

Art	Stämme	Art	Stämme
<i>Issatchenkia hanoiensis</i>	3	<i>Pichia fermentans</i>	23
<i>Issatchenkia occidentalis</i>	11	<i>Pichia galeiformis</i>	6
<i>Issatchenkia orientalis</i>	36	<i>Pichia guilliermondii</i>	20
<i>Issatchenkia scutulata</i>	6	<i>Pichia jadinii</i>	11
<i>Issatchenkia terricola</i>	8	<i>Pichia kluyveri</i>	14

Art	Stämme	Art	Stämme
<i>Pichia angusta</i>	9	<i>Pichia membranifaciens</i>	21
<i>Pichia anomala</i>	18	<i>Pichia nakasei</i>	2
<i>Pichia cactophila</i>	7	<i>Pichia norvegensis</i>	8
<i>Pichia deserticola</i>	14	<i>Pichia pseudocactophila</i>	3
<i>Pichia fabianii</i>	5	<i>Pichia triangularis</i>	9
<i>Pichia farinosa</i>	4	<i>Williopsis californica</i>	7

Für Stämme aus dem Molkerei- und Futtermittelumfeld lieferte das Netz ausgezeichnete Ergebnisse, da in einer Validierung mit 1639 Spektren von 179 unbekanntem Stämmen 98,6 % der Spektren korrekt klassifiziert wurden (Büchl et al., 2008). Derart hohe Identifizierungsquoten sind mit klassischen phänotypischen Methoden nicht zu erreichen und werden sonst nur von molekularbiologischen Techniken erzielt. Um die Anwendbarkeit dieses KNN auf brauerei- und getränketechnologisch relevante Isolate zu prüfen, wurden die im weiteren Verlauf des Projektes gesammelten 50 Stämme der Gattungen *Pichia*, *Issatchenkia* und *Williopsis* (Tabelle 4, 2. Datensatz) über dieses KNN identifiziert.

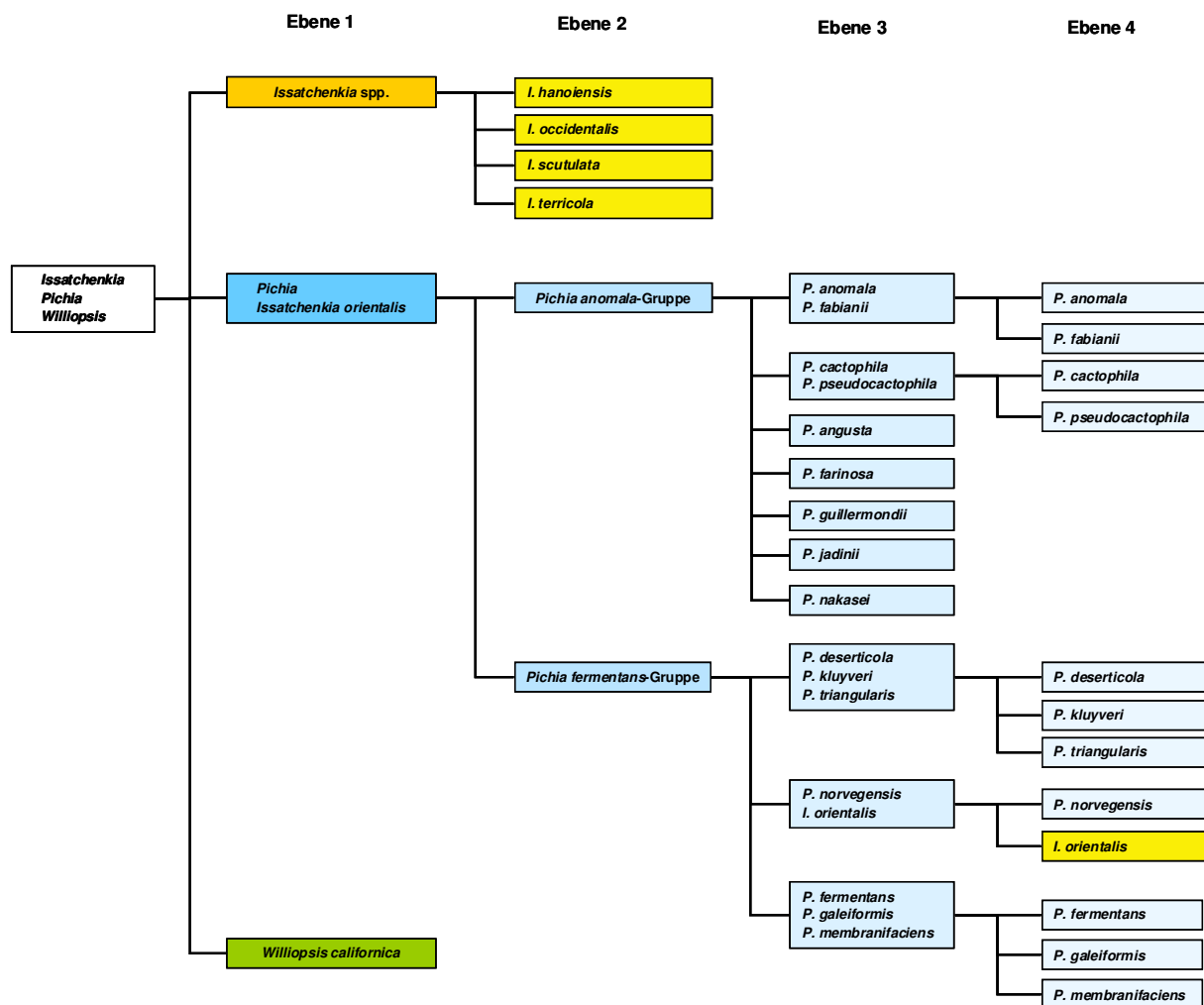


Abbildung 6: Architektur des KNN zur Identifizierung von Hefen der Gattungen *Issatchenkia* (gelb), *Pichia* (blau) und *Williopsis* (grün).

Es stellte sich heraus, dass nur 76,4 % der 657 Spektren der richtigen Art zugeordnet wurden, 15,2 % waren falsch klassifiziert und 8,4 % konnten nicht identifiziert werden, sondern wurden vom KNN als ‚failed‘ (fehlgeschlagen) abgelehnt. Eine fehlgeschlagene Identifizierung bedeutet, dass das Muster des entsprechenden Spektrums zu keiner Klasse des KNN passt und ist in diesem Fall ein Indiz dafür, dass sich Stämme aus dem Brauereiumfeld in ihrer Zellzusammensetzung von denen aus Molkereien oder Futtermitteln unterscheiden, weshalb das KNN keine Zuordnung treffen konnte. Schwierigkeiten bereitete vor allem die Identifizierung der Arten *P. anomala* und *P. membranifaciens*, auf die insgesamt 63 % der nicht korrekt identifizierten Spektren entfielen. Dies ist insofern erstaunlich, als das KNN bereits drei *P. anomala*- und fünf *P. membranifaciens*-Isolate enthielt, die als Teil des ersten Datensatzes Bestandteil des Trainings waren. Betrachtet man jedoch auch die Identifizierbarkeit dieser Stämme über die OPUS-Datenbank, so ist festzustellen, dass bis auf einen *P. membranifaciens*-Stamm alle Isolate auch mit der OPUS-Datenbank korrekt identifiziert wurden, von den 29 Stämmen des zweiten Datensatzes hingegen nur neun. Zufallsbedingt wiesen also die Stämme des ersten Datensatzes relativ große Ähnlichkeiten zu den Referenzen aus dem Molkereibereich auf und waren damit nicht ausreichend repräsentativ für den Brauereisektor. Infolgedessen war ihr Beitrag im Hinblick auf eine Anpassung des KNN an den Brauereibereich begrenzt und reichte nicht aus, die Identifizierung der übrigen Stämme aus dem B95-Projekt mit hoher Zuverlässigkeit zu bewerkstelligen.

Tabelle 4: Zur Erweiterung des KNN zur Identifizierung von Hefen der Gattungen *Pichia*, *Issatchenkia*, *Williopsis* und *Kregervanrija* verwendete Stämme aus dem B95-Projekt. Der erste Datensatz wurde bereits zur Erstellung des oben beschriebenen KNN verwendet; der zweite Datensatz stellte die endgültige Adaption an Hefen aus dem Brauereiumfeld dar.

Art	Stämme		
	1. Datensatz	2. Datensatz	gesamt
<i>Issatchenkia occidentalis</i>	1	-	1
<i>Issatchenkia orientalis</i>	-	2	2
<i>Kregervanrija delftensis</i>	-	3	3
<i>Kregervanrija fluxuum</i>	-	1	1
<i>Pichia anomala</i>	3	17	20
<i>Pichia fabianii</i>	1	-	1
<i>Pichia fermentans</i>	1	-	1
<i>Pichia galeiformis</i>	1	2	3
<i>Pichia guilliermondii</i>	3	10	13
<i>Pichia membranifaciens</i>	5	12	17
<i>Pichia norvegensis</i>	-	1	1
<i>Williopsis californica</i>	-	2	2
Gesamt	15	50	65

Das bestehende KNN war somit für die Identifizierung von Isolaten aus dem Brauereiumfeld nicht optimal ausgelegt und musste weiter an das dort vorkommende Keimspektrum angepasst werden. Zu diesem Zweck wurde es mit den übrigen 50 Stämmen der in Tabelle 4 gelisteten Arten (2. Datensatz) erweitert.

Beim Neutraining wurde das KNN zusätzlich um die Gattung *Kregervanrija* ergänzt. Die zwei Arten *Kregervanrija delftensis* und *K. fluxuum* gehörten bis vor kurzem noch der Gattung *Pichia* an, wurden jedoch auf Basis neuer phylogenetischer Studien in die neu geschaffene Gattung *Kregervanrija* ausgegliedert. Dies hatte zur Folge, dass die Architektur des KNN geändert werden musste, um diese beiden Arten sinnvoll in das System zu integrieren. Da alle vier Stämme hohe Ähnlichkeit zu *P. kluyveri* und *P. deserticola* aufwiesen, wurden

sie in diese Subgruppe integriert und auf Ebene vier als eigene Klassen etabliert. Das aus der Erweiterung resultierende KNN ist in Abbildung 7 dargestellt.

Die Leistungsfähigkeit des neuen KNN für Stämme aus der Milch- und Futtermittelindustrie wurde anhand der gleichen 179 Stämme getestet, mit denen bereits das ursprüngliche KNN validiert wurde. 97,3 % der Spektren wurden korrekt identifiziert, was nur 1,5 % schlechter als das Resultat des ursprünglichen KNN ist und immer noch eine hervorragende Treffsicherheit darstellt. Somit war es möglich, das KNN ohne nennenswerten Qualitätsverlust mit Isolat-ten aus dem Brauereisektor zu erweitern, obwohl diese zu einem nicht unwesentlichen Teil andere FTIR-Muster aufwiesen. Hier zeigt sich das große Potential von KNN, die über den Trainingsprozess erlernen, wesentliche artspezifische Muster in den FTIR-Spektren zu erkennen und stärker zu gewichten als andere Merkmale. Dies führt zu einer besseren Generalisierbarkeit der Identifizierung, die nicht mehr von stammspezifischen Charakteristika beeinträchtigt wird.

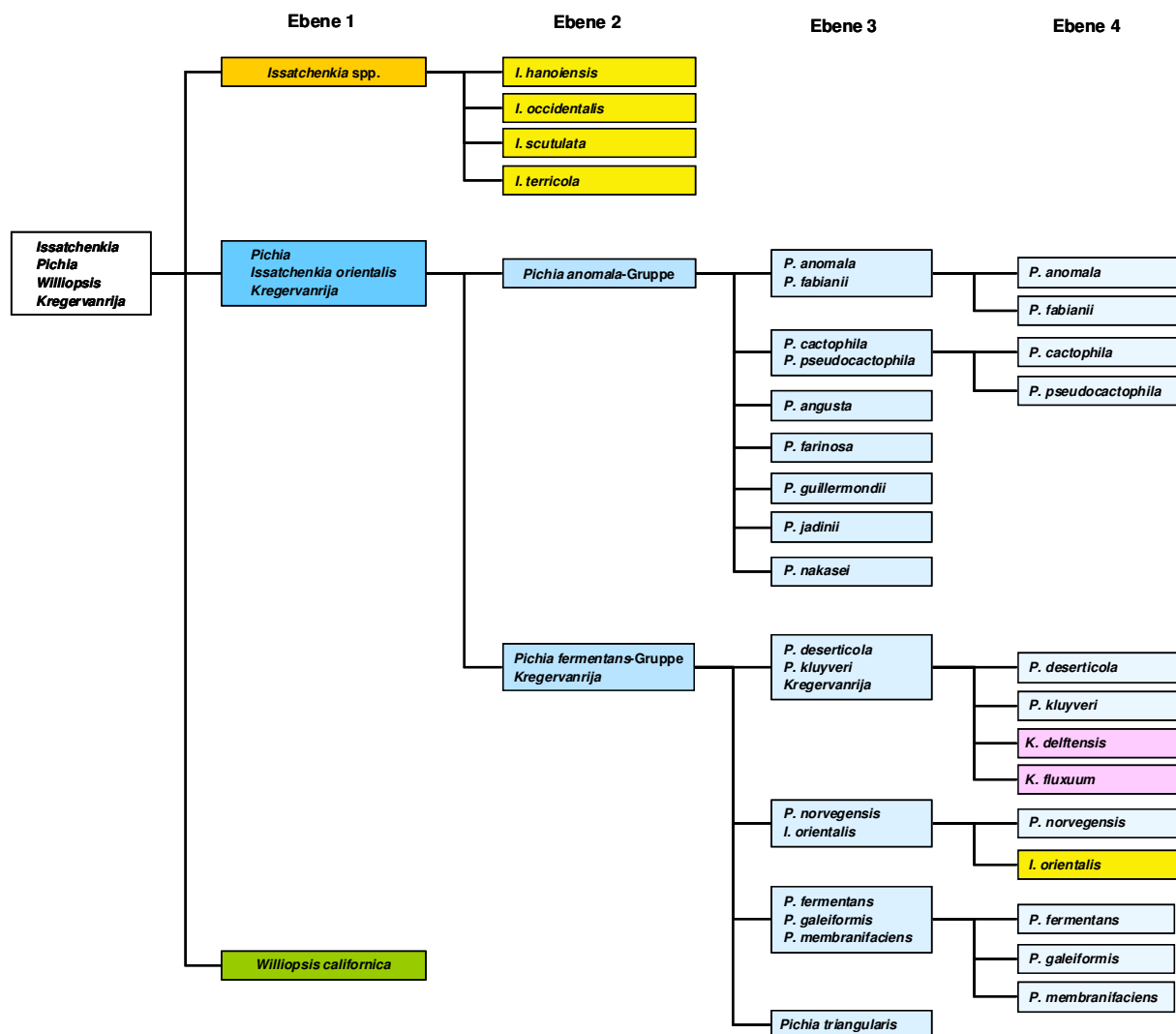


Abbildung 7: Architektur des KNN zur Identifizierung von Hefen der Gattungen *Issatchenkia* (gelb), *Pichia* (blau), *Williopsis* (grün) und *Kregervanrija* (violett).

4.1.1.3 Identifizierungspotential der erweiterten Datenbanken für brauereirelevante Stämme

Um die Leistungsfähigkeit der resultierenden Datenbank zu testen, sollte eine Validierung mit nicht im Referenzdatensatz enthaltenen Stämmen durchgeführt werden. Da aber alle im B95-Projekt gesammelten Stämme für die Erweiterung der Datenbank genutzt wurden, standen aus diesem Pool keine Stämme mehr für eine Validierung zur Verfügung. Um dennoch eine Einschätzung der Leistungsfähigkeit in Bezug auf Hefen aus dem Brauereisektor vornehmen zu können, wurden von einer Brauerei nach Ablauf der offiziellen Projektlaufzeit 34 weitere Stämme zur Verfügung gestellt, die zunächst über real-time PCR und DNA-Sequenzierung identifiziert wurden (Tabelle 5). 13 Isolate und damit knapp 40 % der Stämme gehörten zu *P. anomala*, drei zu *P. fermentans*, zwei zu *P. galeiformis*, sechs zu *P. membranifaciens* einer zu *W. californica* und acht Stämme zu *C. boidinii*. Von jedem Stamm wurden drei FTIR-Spektren aufgenommen und zunächst mit der aktualisierten OPUS-Datenbank getestet. Insgesamt konnten 60 % der Spektren auf Artebene identifiziert werden, wobei *P. anomala* mit 90 % korrekt identifizierten Spektren am besten abschnitt, *W. californica* mit keinem identifizierten Spektrum am schlechtesten (Tabelle 5). Auch die *C. boidinii*-Isolate schneiden mit lediglich 33 % identifizierten Spektren relativ schlecht ab, obwohl vier Stämme im Rahmen des B95-Projektes isoliert und in die Datenbank eingepflegt wurden. Jedoch sind insgesamt nur acht *C. boidinii*-Stämme als Referenzen hinterlegt, was für eine gute Abdeckung der Biodiversität nicht ausreichend ist. Oberreuter et al. (2002) haben ermittelt, dass für eine gute Identifizierungsrate mindestens zehn Referenzen pro Art in einer Datenbank enthalten sein sollten. Hier besteht also noch weiterer Anpassungsbedarf.

Tabelle 5: Zur Validierung der adaptierten Datenbanken verwendete Stämme und deren Identifizierbarkeit mit den jeweiligen Datenbanken auf Artebene.

Art	Stämme	Identifizierung mit		
		OPUS (neu)	KNN (alt)	KNN (neu)
<i>Candida boidinii</i>	8	33 %	-	-
<i>Pichia anomala</i>	13	90 %	44 %	100 %
<i>Pichia fermentans</i>	3	44 %	33 %	78 %
<i>Pichia galeiformis</i>	2	17 %	100 %	100 %
<i>Pichia membranifaciens</i>	6	61 %	67 %	89 %
<i>Williopsis californica</i>	1	0 %	100 %	67 %
Gesamt	33	60 % (68 %*)	55 %	93 %

* nur *Pichia*- und *Williopsis*-Arten

Die 25 Stämme der Gattungen *Pichia* und *Williopsis* wurden zudem zur Validierung des neu entstandenen KNN genutzt. Um eine mögliche Verbesserung der Identifizierungssicherheit beurteilen zu können, wurden diese zunächst mit dem alten KNN, das nur 15 Isolate aus dem Brauereiumfeld enthielt, identifiziert, wobei 55 % der Spektren der korrekten Klasse zugeordnet wurden; 23 % waren failed und 32 % falsch identifiziert. Nahezu zwei Drittel der nicht korrekt identifizierten Spektren gehörten zur Art *P. anomala*, die somit analog zu den mit den B95-Stämmen erzielten Resultaten überproportional schlechte Ergebnisse lieferte. Problematisch war hier in fast allen Fällen die Trennung zwischen *P. anomala* und *P. fabianii*, die sehr eng verwandt sind und daher über einen zusätzlichen Identifizierungsschritt erst auf Ebene vier des KNN differenziert werden. Gerade hier brachte die Erweiterung des KNN eine wesentliche Verbesserung, da nach dem Neutraining 100 % der *Pichia anomala*-Spektren korrekt klassifiziert wurden. Auch für die Spezies *P. fermentans* und *P. membranifaciens* konnte die Identifizierung wesentlich verbessert werden; nur bei *W. californica* wurde mit dem neuen KNN ein Spektrum weniger identifiziert als mit dem alten.

Insgesamt kam das neue KNN auf eine Quote von 93 % korrekt identifizierten Spektren und liegt damit fast 40 % besser als das alte KNN. Auch die Unterschiede in den Validierungsergebnissen für Stämme aus dem Molkerei- und Futtermittelumfeld und dem Brauereisektor liegt nun nur noch bei 4 %, obwohl die Referenzen aus dem Brauereiumfeld nur einen kleinen Teil des Datensatzes für die Erstellung des KNN ausmachten. Das neue KNN liefert damit auch für die Anwendung in der Brauereiindustrie sehr zuverlässige Ergebnisse. Die Ergebnisse machen auch deutlich, dass die Auswertung mit KNN einen wesentlichen Fortschritt für die Sicherheit erzielter Identifizierungsergebnisse im Vergleich zur OPUS-Datenbank darstellt, da mit KNN 25 % bessere Ergebnisse erzielt werden als mit der OPUS-Datenbank, die nur 68 % richtige Identifizierungen ergab.

4.1.2 Identifizierung von *Saccharomyces*-Fremdhefen

Im Rahmen einer Diplomarbeit in Kooperation mit einer Brauerei erfolgte die Erarbeitung eines KNN zur Identifizierung von Hefen des *Saccharomyces* sensu stricto / sensu lato-Komplexes. Obwohl diese Arbeit nicht zu dem hier vorliegenden Projekt gehört, sollen die Ergebnisse kurz dargestellt werden, da das erstellte KNN die Basis für das im Rahmen dieses Projektes zu erstellende Netz darstellte und beide Arbeiten dann nicht mehr zu differenzieren sind.

Für 174 Stämme der in Tabelle 6 aufgeführten Arten wurden jeweils 10 FTIR-Spektren aufgenommen. Von diesen wurden neun für das Training des KNN verwendet, eines diente dem Test des erhaltenen Netzes (interne Validierung).

Tabelle 6: Zur Erstellung des KNN zur Identifizierung von Hefen des *Saccharomyces*-Komplex verwendete Arten

Art	Stämme	Art	Stämme
<i>Saccharomyces bayanus</i>	7	<i>Saccharomyces mikatae</i>	1
<i>Saccharomyces bayanus/pastorianus</i>	3	<i>Saccharomyces paradoxus</i>	7
<i>Saccharomyces pastorianus</i>	2	<i>Kazachstania exigua</i>	14
<i>Saccharomyces pastorianus</i> (untergärig)	24	<i>Kazachstania servazii</i>	3
<i>Saccharomyces cariocanus</i>	2	<i>Kazachstania unispora</i>	14
<i>Saccharomyces cerevisiae</i>	54	<i>Lachancea kluyveri</i>	3
<i>Saccharomyces cerevisiae</i> var. <i>diastaticus</i>	35	<i>Naumovia castelli</i>	1
<i>Saccharomyces kudriazevii</i>	2	<i>Naumovia dairensis</i>	2

Beim Aufbau des KNN hat sich eine hierarchische Struktur mit vier Ebenen als zielführend erwiesen. Insgesamt wurden zehn Einzelnetze trainiert und zu einem KNN zusammengefügt (Abb. 8). Auf der ersten Ebene wurden die Hefen der Gattung *Saccharomyces* als größte Klasse von den drei weiteren Gattungen *Kazachstania*, *Naumovia* und *Lachancea* getrennt und letztgenannte bereits auf der ersten Ebene in drei Untergruppen unterteilt. Auf der zweiten Ebene erfolgte die Differenzierung von *S. cerevisiae* von den übrigen *Saccharomyces*-Arten, deren weitere Auftrennung auf den Ebenen drei und vier realisiert wurde. Die größten Schwierigkeiten bereitete erwartungsgemäß die Trennung von *S. cerevisiae* und *S. cerevisiae* var. *diastaticus*, die nicht auf einer Ebene bewältigt werden konnte. Zur Lösung des Problems wurde zunächst auf Ebene drei ein Teil der *S. cerevisiae*-Stämme abgetrennt und auf der vierten Stufe die endgültige Differenzierung vorgenommen.

Nach Abschluss des Trainings wurde mit den zurückgehaltenen Spektren der Stämme, die im Training enthalten waren, eine interne Validierung durchgeführt. Diese diente der Kontrolle der gewählten Architektur des KNN und dem Test der erarbeiteten Klassifizierungsparameter. Die interne Validierung ergab mit 97,7 % korrekt klassifizierten Spektren ein gutes Ergebnis und deutete darauf hin, dass die gewählte Struktur und Hierarchie sinnvoll waren. Zum Testen des KNN unter Praxisbedingungen, wo normalerweise Stämme identifiziert werden, die dem KNN unbekannt sind, waren die Spektren von 23 Stämmen der Arten *S. pastorianus* (untergärig), *S. cerevisiae* und *S. cerevisiae* var. *diastaticus* nicht für das Training verwendet, sondern zurückgehalten worden. Eine Identifizierung dieser Spektren ergab mit 81,6 % eine deutlich niedrigere Quote korrekter Zuordnungen, die darauf hindeutete, dass das KNN noch nicht optimal funktionierte.

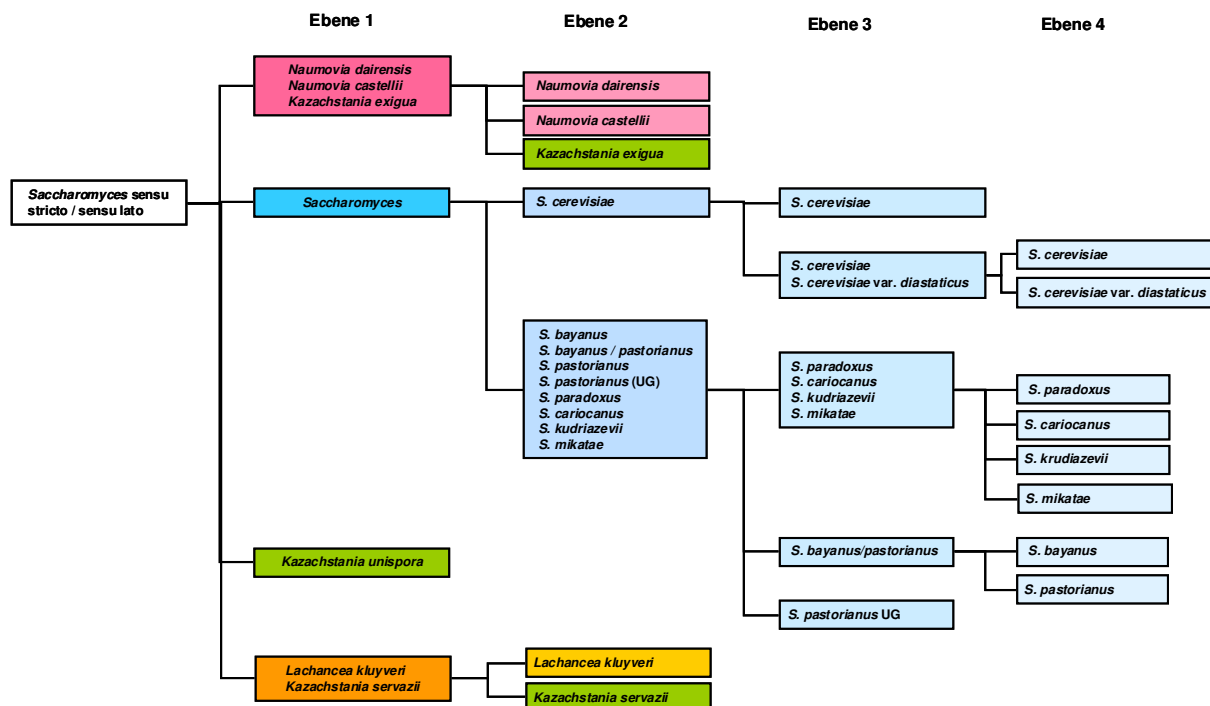


Abbildung 8: Vierschichtiges KNN zur Identifizierung von Hefen des *Saccharomyces sensu stricto /sensu lato*-Komplexes.

Zunächst ist zu bemerken, dass die Zuordnung bis zur 2. Ebene immer korrekt ist, erst ab der dritten Ebene treten Schwierigkeiten auf. Bei *S. pastorianus* konzentrierten sich die falsch oder nicht identifizierten Spektren auf zwei Stämme, was bedeutet, dass hier wohl die existierende Biodiversität dieser Art noch nicht ausreichend erfasst ist und die Datenbank mit weiteren Stämmen erweitert werden müsste. Zudem ist die Referenz-Klassifizierung von *S. pastorianus* (Nicht Brauerei-Kulturhefe) und *S. bayanus*-Stämmen anhand der verwendeten PCR-Systeme nur für den überwiegenden Teil eindeutig. Liefern die Ergebnisse der molekularen Marker identische Identifizierungsmatrizes konnte nur als Gruppe *S. bayanus / S. pastorianus* identifiziert werden. Diese Identifizierungs-Ergebnisse laufen konform mit der etablierten Klassifikation der renommierten Stammsammlung des CBS (Centralbureau voor Schimmelcultures, Utrecht, Niederlande). Dieser zufolge sind auch einige Stämme als *S. pastorianus / S. bayanus* klassifiziert. Bei *S. cerevisiae* sind die Falschklassifizierungen etwas breiter gestreut. Bemerkenswert ist, dass 70 % aller falsch identifizierten Spektren (12 % aller Spektren waren falsch identifiziert) auf die Unterscheidung zwischen *S. cerevisiae* und *S. cerevisiae*

var. *diastaticus* entfallen. Nimmt man diese Differenzierung aus dem KNN heraus, steigt die Quote der korrekt identifizierten Spektren auf 89,8 % und 98 % der *S. cerevisiae*- und *S. cerevisiae* var. *diastaticus*-Spektren werden korrekt identifiziert.

Bevor mit der Erweiterung dieses KNN mit den im Rahmen des B95-Projektes gewonnenen Isolaten begonnen wurde, sollte ermittelt werden, wie die Leistungsfähigkeit des bestehenden KNN für diese Isolate ist. Auch hier wurden ca. zehn Spektren für jeden Stamm aufgenommen und nun über das KNN identifiziert, wobei die Trennung von *S. cerevisiae* und *S. cerevisiae* var. *diastaticus* einmal berücksichtigt wurde und einmal nicht. Insgesamt standen 42 Stämme der Arten *S. bayanus / pastorianus*, *S. cerevisiae*, *S. cerevisiae* var. *diastaticus* und *Saccharomyces pastorianus* (untergärig) zur Verfügung (Tabelle 7). Die Ergebnisse entsprachen mit 59 % bzw. 71 % nicht den Erwartungen. Besonders frappierend war, dass ein großer Teil der *S. cerevisiae*-Spektren bereits auf der ersten Ebene keiner Klasse zugeordnet werden konnte, weshalb für diese Art auch ohne Berücksichtigung der Trennung von *S. cerevisiae* var. *diastaticus* nur knapp 42 % der Spektren korrekt identifiziert wurden. Dies ist ein ganz klares Indiz dafür, dass die gewählte Architektur des KNN für die Identifizierung einer großen Vielfalt an *S. cerevisiae*-Stämmen nicht geeignet ist und anders konzipiert werden muss. Auch die Trennung von *S. cerevisiae* und *S. cerevisiae* var. *diastaticus* sowie die Identifizierung der Art *S. bayanus / pastorianus* lieferte keine zufrieden stellenden Ergebnisse, weshalb das KNN grundlegend überarbeitet wurde.

Tabelle 7: Leistungsfähigkeit des ursprünglichen KNN zur Identifizierung von Hefen des *Saccharomyces*-Komplex, ermittelt mit den im B95-Projekt isolierten Stämmen.

Art	Stämme	Spektren	KNN	
			mit <i>diastaticus</i>	ohne <i>diastaticus</i>
<i>Saccharomyces bayanus / pastorianus</i>	4	41	17,1 %	17,1 %
<i>Saccharomyces cerevisiae</i>	11	110	20,0 %	41,8 %
<i>Saccharomyces cerevisiae</i> var. <i>diastaticus</i>	8	78	57,7 %	98,7 %
<i>Saccharomyces pastorianus</i> UG	19	241	85,0 %	85,0 %
Gesamt	42	470	59,4 %	71,1 %

Bevor die neue Struktur entworfen wurde, wurde eine Clusteranalyse der FTIR-Spektren durchgeführt, um einen Eindruck über Ähnlichkeiten zwischen den einzelnen *Saccharomyces*-Arten zu gewinnen (Abbildung 9). Farblich hervorgehoben sind die Arten *S. cerevisiae* (blau), *S. pastorianus* (gelb) und Kulturhefen von *S. cerevisiae* (grün).

Es stellte sich heraus, dass die Spektren in drei große Cluster zerfallen, von denen das unterste ausschließlich *S. cerevisiae* enthält und die beiden anderen *S. cerevisiae* (sowohl Kulturhefen (grün markiert) als auch Feldisolate (blau markiert)), *S. pastorianus* (gelb markiert) und andere Arten. Die Stämme von *S. cerevisiae* weisen demnach z.T. deutliche Unterschiede zueinander auf und bilden zwei große Gruppen. Die eine (unterstes Cluster) ist von allen anderen *Saccharomyces*-Arten klar getrennt, die andere mischt sich mit den übrigen Arten in den beiden oberen Clustern. Somit war klar, dass eine Trennung von *S. cerevisiae* und *S. pastorianus* in einem Schritt schwierig werden und evtl. mehrere Ebenen erfordern würde.

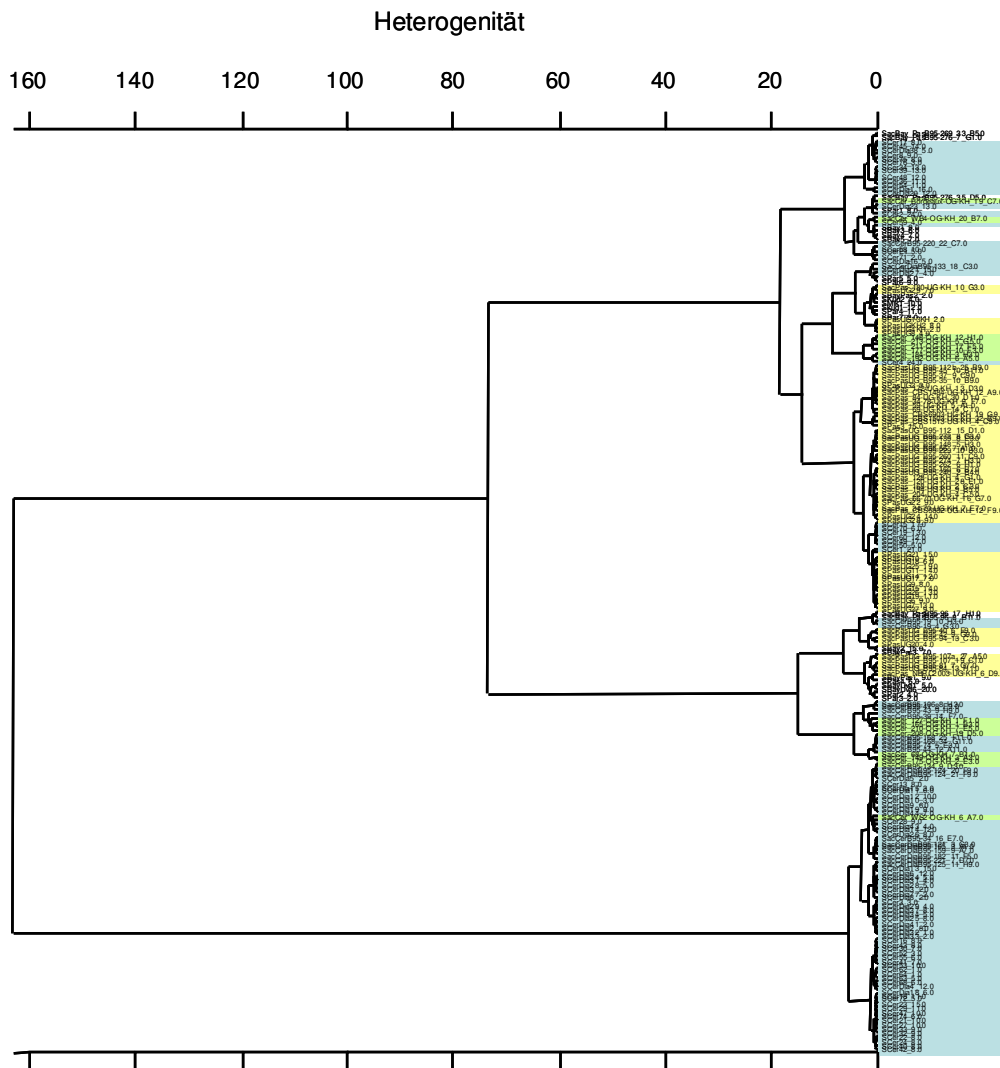


Abbildung 9: Hierarchische Clusteranalyse aller Stämme der Gattung *Saccharomyces*. Blau: *S. cerevisiae*, gelb: *S. pastorianus*, grün: *S. cerevisiae* Kulturhefen. Je kürzer die Abstände zwischen den Spektren (niedrige Heterogenität) sind, desto ähnlicher sind sie sich.

Um ausreichend Stämme zum Testen der veränderten Architektur zur Verfügung zu haben, wurde das KNN zunächst nur um die *S. cerevisiae* und *S. pastorianus* Kulturhefestämme (Anhang 1) erweitert und die 23 Stämme der ursprünglichen Validierung sowie die aus dem B95-Projekt stammenden Isolate vorerst zurückgehalten. Während des Trainings stellte sich heraus, dass, wie vermutet, die Architektur des KNN zur Verbesserung der Identifizierungen deutlich verändert werden musste (Abb. 10). Auf der ersten Ebene werden nunmehr nur zwei statt vorher vier Klassen voneinander getrennt und so die Arten des *Saccharomyces* sensu stricto- von denen des sensu lato-Komplexes separiert. Der *S. sensu lato*-Komplex wurde im Folgenden in die Gattungen *Kazachstania*, *Lachancea* und *Naumovia* differenziert und die Gattung *Kazachstania* weiter in die einzelnen Arten. Aufgrund nur sehr geringer Anzahl an zur Verfügung stehenden Stämmen für *Lachancea* und *Naumovia*, wurde hier keine weitere Auftrennung vorgenommen.

Die Differenzierung der *Saccharomyces*-Arten wurde dahingehend in ihrer Struktur verändert, dass *S. cerevisiae* in zwei Gruppen unterteilt wurde, von denen eine auf Ebene zwei von den übrigen *Saccharomyces*-Arten abgetrennt, die zweite jedoch noch in die Klasse von *S. pastorianus* und den anderen Arten integriert und erst auf Ebene drei separiert wurde. Um Fehlklassifizierungen der ersten abgetrennten Gruppe von *S. cerevisiae* aufgrund großer Ähnlichkeiten zu *S. bayanus* / *pastorianus* zu vermeiden, wurde auf Ebene drei eine zusätzliche

Sicherheitsstufe zur Trennung dieser *S. cerevisiae*-Gruppe von *S. bayanus / pastorianus* einbezogen.

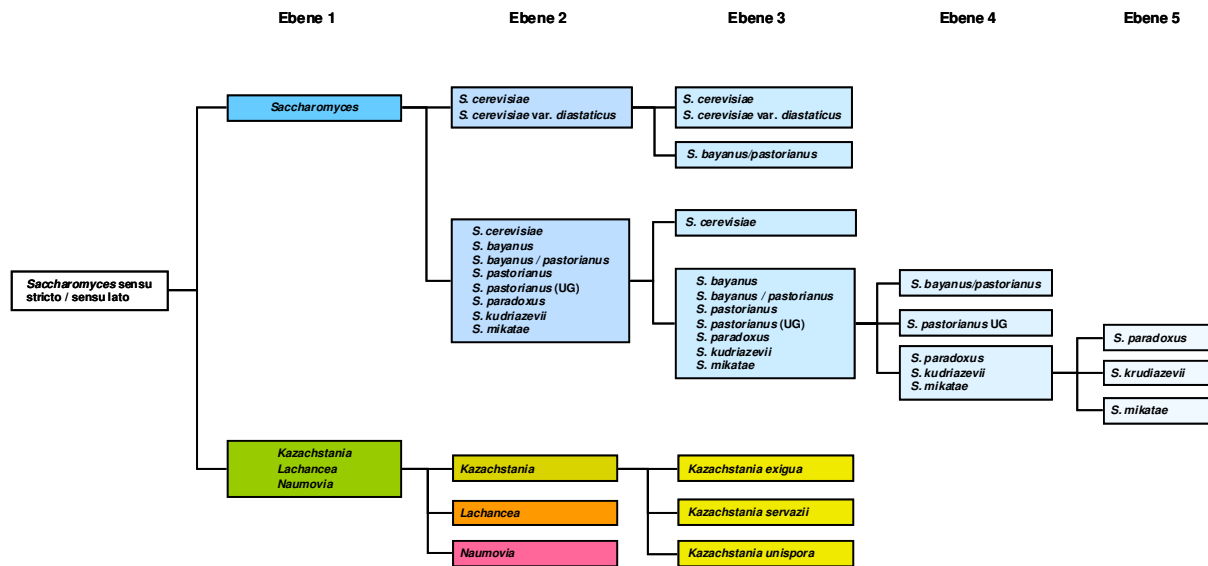


Abbildung 10: Überarbeitetes und mit Kulturhefen erweitertes KNN zur Identifizierung von Hefen des *Saccharomyces sensu stricto / sensu lato*-Komplexes.

Auf die Unterscheidung von *S. cerevisiae* und *S. cerevisiae* var. *diastaticus* musste verzichtet werden, da die Überlappungen zwischen diesen beiden Varianten so groß sind, dass ihre FTIR-Spektren nicht zuverlässig voneinander zu trennen waren. Auch die Arten *S. bayanus* und *S. pastorianus* (nicht untergärig), die im ursprünglichen KNN noch bis auf Artebene aufgesplittet wurden, wurden jetzt in einer Klasse vereint, da für einen nicht unerheblichen Teil der Referenzstämmen die konkrete Artzugehörigkeit nicht bestimmt werden konnte. Zudem ist eine Unterscheidung von *S. bayanus* und *S. pastorianus* (nicht untergärig) von untergeordneter Paraxisrelevanz. Das Identifizierungsergebnis *S. bayanus / S. pastorianus* indiziert dem Brauerei-/ Getränketechnologen die Anwesenheit einer Schadhefe, welche bei niedrigen Temperaturen wachsen kann. Diese Aussage ist im Regelfall absolut ausreichend und zweckmäßig. Das so entstandene KNN hatte nun fünf Ebenen und wurde mit den Stämmen des ursprünglichen Validierungsdatensatzes der Diplomarbeit (Datensatz 1) und den Stämmen des B95-Projektes (Datensatz 2) getestet (Tabelle 8).

Tabelle 8: Leistungsfähigkeit des mit Kulturhefen erweiterten und überarbeiteten KNN zur Identifizierung von Hefen des *Saccharomyces*-Komplex (%). In Klammern: Werte des ursprünglichen KNN.

Art	1. Datensatz	2. Datensatz
<i>Saccharomyces bayanus / pastorianus</i>	--	87,8 (17,1)
<i>Saccharomyces cerevisiae</i>	95,1 (98,4)	96,4 (41,8)
<i>Saccharomyces cerevisiae</i> var. <i>diastaticus</i>	80,0 (90,0)	92,3 (98,7)
<i>Saccharomyces pastorianus</i> UG	73,3 (74,3)	85,5 (85,0)
Gesamt	87,1 (89,8)	89,2 (71,1)

Es zeigte sich, dass die Veränderungen der Struktur und die Erweiterung mit den Kulturhefestämmen insgesamt deutliche Verbesserungen brachten. Für den ersten Datensatz wurden mit 87,1 % etwas schlechtere Ergebnisse erzielt, für die Stämme des zweiten Datensatzes jedoch mit 89,2 % ein um fast 20 % besseres Ergebnis. Vor allem die Identifizierung von *S. bayanus / pastorianus* konnte wesentlich verbessert werden, was jedoch nicht daran lag, dass die bei-

den Arten in einer Klasse zusammengefasst wurden, sondern an der Umstrukturierung des KNN. Dass diese sinnvoll und zielführend war, lässt sich auch an den Ergebnissen für *S. cerevisiae* erkennen, die von knapp 42 % auf 96 % gesteigert werden konnten. Die nun gewählte Architektur scheint somit für die Identifizierung von *Saccharomyces*-Arten gut geeignet zu sein und verlässliche Ergebnisse zu garantieren. Um die Datenbasis für die Identifizierung weiter zu verbreitern und die Identifizierungsergebnisse weiter abzusichern, wurden alle im B95-Projekt gesammelten 42 Stämme über ein neues Training in das KNN integriert, wobei die in Abbildung 10 veranschaulichte Struktur beibehalten wurde. Das so entstandene endgültige KNN wurde anschließend mit den Spektren der 23 Stämme des ersten Datensatzes validiert und erzielte mit 92,9 % ein sehr gutes und nochmals gesteigertes Ergebnis.

Einige der gesteckten Ziele für das Klassifizierungsschema konnten leider nicht realisiert werden. So ist eine Auftrennung zwischen *S. bayanus* und *S. pastorianus* mit den hier zur Verfügung stehenden Referenzstämmen nicht möglich, was jedoch auch an der geringen Stammzahl liegen kann, die für eine Extraktion der wesentlichen Merkmale aus den FTIR-Spektren nicht ausreicht. Sollten in der kommenden Zeit weitere Stämme zur Verfügung stehen, kann die Lösung dieses Problem nochmals in Angriff genommen werden. Des Weiteren war keine Differenzierung von *S. cerevisiae* und *Saccharomyces cerevisiae* var. *diastaticus* möglich, da zwischen beiden Varianten zu große Überlappungen bestehen. Hier gab es zwar keine Einschränkungen im Hinblick auf die Zahl der zur Verfügung stehenden Stämme, doch das Merkmal der übervergärenden Aktivität, die durch das Vorhandensein des Enzyms Glucoamylase determiniert wird, scheint sich in den FTIR-Spektren nicht so auszuprägen, dass es eine Differenzierung erlauben würde. Hier ist sicherlich weiterhin die real-time PCR die Methode der Wahl. Für alle weiteren Identifizierungen hingegen bietet die FTIR-Spektroskopie eine zuverlässige Alternative.

4.2 Identifizierung von Kulturhefestämmen der Weihenstephaner Hefebank

Das dritte und mit Sicherheit ehrgeizigste Ziel des B95-Projektes betraf die Differenzierung einzelner Stämme der Weihenstephaner Hefebank mit dem Ziel, eine Kontrolle von Kulturhefen in der Brauerei und eine Differenzierung der Kulturhefen von *Saccharomyces*-Fremdhefen zu ermöglichen. So sollte ein Tool entwickelt werden, mit dem beispielsweise die Betriebskultur auf Stammebene geprüft werden kann oder Reinheitskontrollen bei der Hefeführung durchgeführt werden können. Die Trennung von *Saccharomyces*-Fremdhefen der Arten *S. cerevisiae* und *S. pastorianus* von Kulturhefen konnte nicht realisiert werden, da hierfür eine Typisierung aller isolierten Stämme dieser beiden Arten hätte durchgeführt werden müssen, um sie entweder sicher einem Kulturhefestamm zuzuordnen oder als Fremdhefe deklarieren zu können. Dieses war jedoch im Rahmen dieses Projektes nicht zu leisten. So sollte ein KNN trainiert werden, das die häufigsten Kulturstämme der Arten *S. cerevisiae* und *S. pastorianus* beinhalten und möglichst auf Stammebene identifizieren sollte. Hierzu wurden insgesamt 34 Kulturhefen, wovon 29 von der Weihenstephaner Hefebank und fünf von der Universität Modena stammten, ausgewählt (Anhang 1). 16 gehörten zur Art *S. cerevisiae*, 18 zu *S. pastorianus*. Dass derartige Problemstellungen mit FTIR-Spektroskopie gelöst werden können, ist bereits belegt, da sowohl für Hefen als auch für Bakterien schon Stammtypisierungssysteme entwickelt wurden (Büchl, et al., 2010, Wenning et al., 2010). Allerdings waren beide publizierten KNN nur für die Differenzierung weniger Stämme ausgelegt und nicht wie hier geplant für über 30.

Die Schwierigkeit der Entwicklung eines KNN auf Stammebene liegt darin, feine vorhandene Unterschiede für eine Differenzierung zu nutzen und die Aufnahme der Spektren so zu gestal-

ten, dass methodenimmanente Varianzen, die z.B. über die Kultivierung der Organismen verursacht werden, nicht zu einer Beeinträchtigung der Identifizierung führen. Um unvermeidbare Varianzen beim Training des KNN herausrechnen zu können, wurden von jedem Stamm 30 unabhängige Referenzspektren aufgenommen, wobei ein kleiner Teil dieser Spektren aus dem Training herausgehalten wurde, um anschließend für den Test der Leistungsfähigkeit des entstandenen KNN verwendet zu werden. Zudem wurden einige Monate nach der Erstellung dieses ersten Datensatzes alle Stämme nochmals aus der Cryobank reaktiviert und für jeden Stamm neun weitere Spektren aufgenommen. Ziel dieses zweiten Datensatzes war, die Robustheit des KNN in Bezug auf auftretende Varianzen resultierend aus einer erneuten Reaktivierung der Stämme zu testen, da nur ein ausreichend robustes Netz die für die Praxis erforderliche Sicherheit garantieren kann.

Um einen ersten Eindruck von der Ähnlichkeit der Spektren einzelner Stämme zueinander zu gewinnen, wurde eine hierarchische Clusteranalyse mit jeweils vier repräsentativen Spektren pro Stamm durchgeführt, die in Abbildung 11 dargestellt ist.

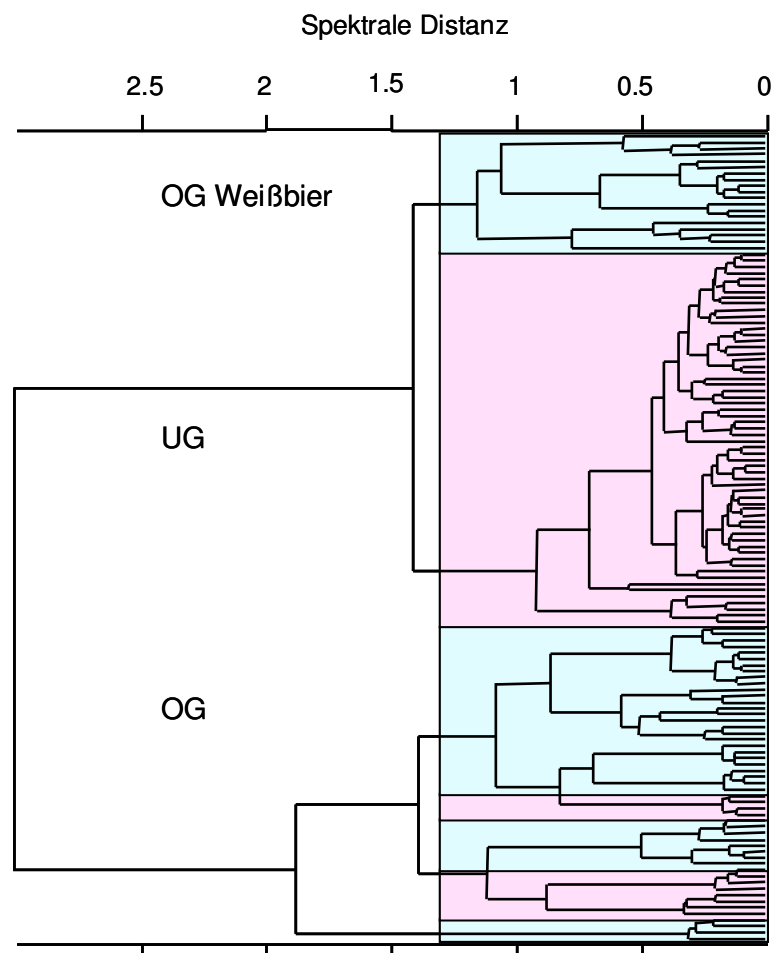


Abbildung 11: Hierarchische Clusteranalyse von 34 Kulturhefestämmen. Obergärige Stämme sind blau hervorgehoben, untergärige rot. Je kürzer die Abstände zwischen den Spektren sind (niedrige spektrale Distanz), desto ähnlicher sind sie sich. OG: obergärige Hefen, UG: untergärige Hefen.

Es ist zu erkennen, dass die Spektren drei Cluster gebildet haben, die jeweils (fast) ausschließlich eine Gruppe von Kulturhefen repräsentieren. Das oberste Cluster (blau) enthält Spektren der vier Weißbierhefen, das mittlere (rot) die Spektren der untergärigen Hefen und das unterste (blau und rot) die übrigen obergärigen Hefen sowie zwei offizielle *S. pastorianus*-Stämme

und den untergärigen Stamm WHB 180. Die Stämme von *S. cerevisiae* zerfallen also auch hier, ähnlich wie beim KNN für den *Saccharomyces sensu stricto / sensu lato*-Komplex, in zwei Gruppen. Weiterhin kann anhand der spektralen Distanzen abgelesen werden, dass die meisten untergärigen Stämme einander deutlich ähnlicher sind als die obergärigen. Im mittleren Cluster haben viele Spektren eine Distanz von 0,5 oder weniger, was für eine hohe Ähnlichkeit spricht. Bei den obergärigen Hefen besteht hingegen eine höhere phänotypische Diversität, da die Spektren einzelner Stämme weiter auseinander liegen. Anhand der Clusteranalyse lässt sich also bereits abschätzen, dass die obergärigen Hefen auf Stammebene leichter zu trennen sein werden, als die untergärigen. Es lässt sich auch ableiten, dass eine Trennung von ober- und untergärigen Hefen gut möglich sein sollte und bei den obergärigen Hefen zwei Gruppen existieren, die zum einen die Weißbierstämme und zum anderen die übrigen obergärigen Stämme enthalten. Diese Anhaltspunkte aus der Clusteranalyse wurden genutzt, um eine sinnvolle Architektur und Hierarchie für das KNN zu entwickeln, wobei zunächst mit der Differenzierung der obergärigen Hefen begonnen wurde.

Abbildung 12 zeigt die Architektur des resultierenden KNN. Die Trennung der Stämme in obergärige und untergärige Hefen konnte trotz der beobachteten Gruppenbildung bei *S. cerevisiae* bereits auf der ersten Ebene bewältigt werden und war hier besser umzusetzen als für die *S. cerevisiae*-Hefen im KNN des *Saccharomyces sensu stricto / sensu lato*-Komplexes. Auf der zweiten Ebene wurde dann die Abtrennung der vier Weißbierhefen von den übrigen obergärigen Stämmen realisiert und auf der dritten Ebene die Kulturstämme für die Bierherstellung von den drei Sekt-, Wein- und Brennereihefen sowie Stamm WHB 127 von den übrigen Weißbierhefen getrennt. Auf Ebene vier erfolgte eine weitere Differenzierung der Gruppen in Einzelstämme oder kleinere Gruppen.

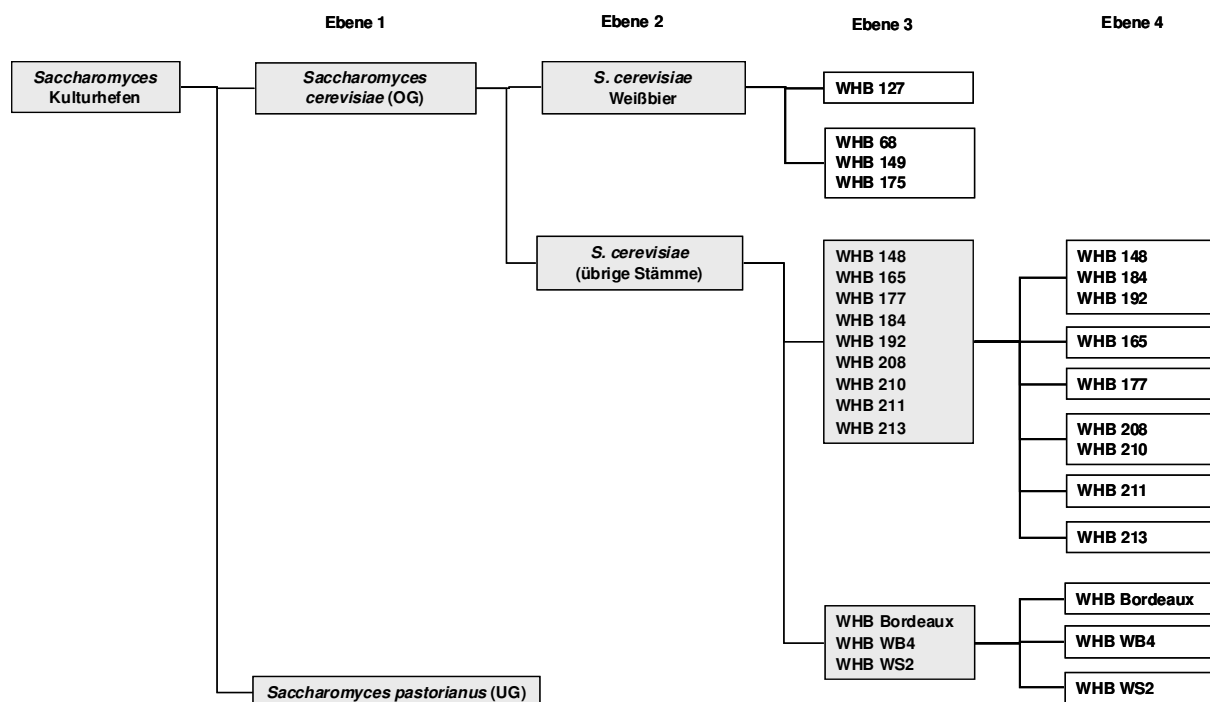


Abb. 12: Architektur des KNN zur Differenzierung von *Saccharomyces*-Kulturhefen. Der Teil des KNN mit guter Verlässlichkeit und Robustheit ist grau hinterlegt.

Eine Validierung des KNN mit den Testspektren des ersten Datensatzes ergab mit 100 % korrekten Zuordnungen ein ausgezeichnetes Ergebnis, bei Nutzung des zweiten Datensatzes an Testspektren traten jedoch einige Schwierigkeiten bei der Klassifizierung auf den Ebenen drei und vier auf. Die Trennung der Weißbierhefen auf Ebene drei (in Abb. 12 weiß hinterlegt)

gelang nur noch mit 82 % der Spektren und die Klassifizierung der übrigen Stämme auf Ebene vier rutschte unter 50 % korrekter Identifizierungen. Gute Raten von 98 % und damit sichere und robuste Identifizierungen ergaben sich nur für die in Abbildung 12 grau hinterlegten Hierarchiestufen.

Die Ergebnisse verdeutlichen, dass ein KNN mit hoher Zuverlässigkeit und guter Eignung für den Einsatz in der Praxis eines Brauereilabors für die Differenzierung einer großen Zahl an Stämmen nicht realisiert werden kann. Zwar ist eine Trennung obergäriger Hefen auf Stammebene möglich, jedoch ist das resultierende System nicht robust genug gegenüber methodenbedingten Varianzen, was jedoch eine unbedingte Voraussetzung für eine zuverlässige Nutzung im Routinelabor ist. Hier liegt ganz klar einer der Schwachpunkte der FTIR-Technologie gegenüber molekularbiologischen Methoden, die zwar deutlich arbeitsaufwendiger, jedoch unempfindlich gegenüber kultivierungsbedingten Abweichungen sind. Da die Schwierigkeiten mit unzureichender Robustheit bereits bei den obergärigen Hefen dazu führten, dass eine Differenzierung auf Stammebene nicht möglich war, wurde für die untergärigen Hefen auf eine weitere Differenzierung verzichtet, da diese noch größere Ähnlichkeit zueinander aufweisen und davon auszugehen ist, dass sich die Robustheitsprobleme hier noch stärker ausprägen.

Dennoch kann konstatiert werden, dass eine verlässliche Trennung obergäriger und untergäriger Hefen möglich ist und die obergärigen Hefen zudem in drei Untergruppen getrennt werden können. Kombiniert man die umfassende OPUS-Datenbank mit den hier entwickelten KNNs, kann mit nur einer Analyse ermittelt werden, ob es sich um eine *Saccharomyces*-Hefe handelt, wenn ja, zu welcher Art sie gehört und ob es ein unter- oder obergäriger Stamm ist. Auch wenn die endgültige Entscheidung auf Stammebene mit dem KNN nicht getroffen werden kann, ist mit relativ wenig Aufwand eine sehr enge Eingrenzung der Identität möglich.

Auch eine weitergehende Differenzierung in Kultur- oder Fremdhefe ist vorstellbar, wenn auch nicht über den Einsatz eines KNN. Alternativ wäre aber das Mitführen von Positivkontrollen in Form der eingesetzten Kulturstämmen bei der FTIR-Messung möglich. Diese würden dann im Parallelansatz mit den unbekanntem Stämmen präpariert, was die kultivierungsbedingten Varianzen für alle Isolate gleichschaltet und somit eine gute Vergleichbarkeit garantiert. Über eine Clusteranalyse der Spektren unbekannter Stämme mit den mitgeführten Kulturstämmen kann dann ermittelt werden, ob markante Unterschiede zwischen den Spektren bestehen. Kann dies bejaht werden, ist es ein starkes Indiz für das Vorliegen einer Fremdhefe. Clusteranalysen können mit der Spektrometersoftware OPUS sehr leicht berechnet werden, da diese Funktion nutzerfreundlich in die Software integriert ist. Mit nur wenigen Minuten Mehraufwand kann so auch ohne Vorhandensein einer Datenbank geprüft werden, ob es sich bei einem Isolat um die eingesetzten Kulturhefe oder eine Fremdhefe handelt.

Exkurs: Veränderung von Kulturhefen über das Weiterführen in der Brauerei

Ein weiteres Argument für das Mitführen von Kulturhefen bei den FTIR-Messungen ist die Hypothese, dass sich Stämme, die über Jahre in der Brauerei weitergeführt werden, verändern könnten und ihre FTIR-Spektren mit den Referenzen in der Datenbank schlechter in Deckung zu bringen wären, was folglich auch die Identifizierbarkeit einschränken würde. Untersucht wurde diese These anhand der Kulturhefen einer Brauerei, die jeweils eine ober- und untergärige Hefe einsetzt. Beide Hefen stammten ursprünglich von der Weihenstephaner Hefebank und werden seit Jahren in der Brauerei weitergeführt. Die obergärige Hefe hatte zum Zeitpunkt der Analyse ca. 250 Führungen hinter sich, die untergärige ca. 300. Für eine Detektion möglicher Veränderung der Hefen gegenüber der Originalkultur der Hefebank wurden sowohl die von der Brauerei eingesetzten Stämme als auch deren frisch von der Hefebank bezogenen Pendanten mit FTIR-Spektroskopie untersucht.

Für jedes Isolat wurden vier unabhängige Spektren aufgenommen, wobei die vier Stämme immer parallel bearbeitet wurden, um kultivierungsbedingte Varianzen zu minimieren. Die Spektren wurden anschließend mit hierarchischer Clusteranalyse auf Unterschiede untersucht (Abb. 13). Das Cluster in Abbildung 13 zeigt zwei klar voneinander getrennte Subcluster, eines für die obergärige (OG) und das andere für die untergärige (UG) Hefe. Das Subcluster der obergärigen Hefe kann wieder in zwei Einzelcluster aufgeteilt werden, von denen eines die Originalkultur der Hefebank, bezeichnet mit WHB, und das andere die von der Brauerei eingesetzte Hefe beinhaltet. Dieser Stamm hat sich also zwar nicht weit, aber doch nachweisbar von der Originalkultur wegentwickelt. Für die untergärige Hefe ist das Ergebnis im Trend ähnlich, jedoch deutlich schwächer ausgeprägt. Hier sind nicht zwei klar getrennte Cluster zu beobachten, sondern eher drei Subcluster, von denen zwei jeweils nur Spektren entweder der Originalkultur oder der der Brauerei enthalten. Das dritte hingegen enthält Spektren beider Kulturen.

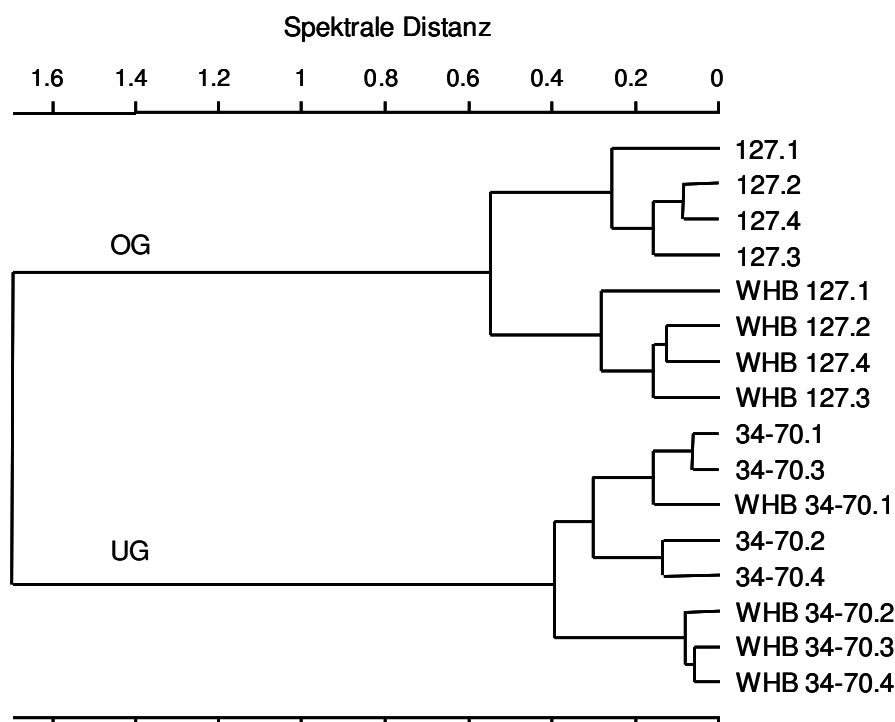


Abb. 13: Vergleich zweier Kulturhefestämme, die seit Jahren in der Brauerei weitergeführt werden mit dem jeweiligen Originalstamm der Weihenstephaner Hefebank. WHB: Weihenstephaner Hefebank.

Die Analysen belegen, dass beim Weiterführen der Kulturhefen in der Brauerei über die Jahre Veränderungen zur Originalkultur auftreten, wobei die Ausprägung bei obergärigen Hefen stärker ist, als bei untergärigen. Diese Veränderungen sind über FTIR-Spektroskopie zu detektieren und könnten unter Umständen eine Stammidentifizierung beeinträchtigen, wenn eine Datenbank verwendet würde, die nur Spektren der Originalkultur enthält. Hausintern wäre es also für Brauereien sinnvoll, die von ihnen verwendeten Kulturen als Positivkontrollen bei der Identifizierung unbekannter Isolate mitzuführen.

4.3 Einfluss selektiver Inkubationsbedingungen auf die Identifizierung

In Brauereien werden unterschiedlichste Selektivmedien zur Fremdhefediagnostik eingesetzt, von denen dann bei positiven Ergebnissen unter Umständen auch Kolonien für eine Identifizierung mit FTIR-Spektroskopie abgenommen werden sollen. Da mit FTIR-Spektroskopie nicht einzelne Merkmale von Organismen bestimmt, sondern die Zusammensetzungen der Zellen vollständig und äußerst sensitiv erfasst werden, haben Kulturmedien einen großen Einfluss auf die Spektren. Je nach Komposition des Mediums verändert sich die Zusammensetzung der Zellen, da unterschiedliche Stoffwechselwege aktiv sind oder Speicherstoffe gebildet werden können. Fraglich ist, ob diese Unterschiede im Rahmen der standardisierten Kultivierung für die FTIR-Spektroskopie, die in der Regel 24 Stunden beträgt, ausgeglichen werden. Auch müssen sich die Zellen nach dem Transfer von einem Selektivmedium auf den Standardagar YGC erst wieder an die geänderten Nährstoffbedingungen anpassen. Dies kann das Wachstum verzögern und somit Einfluss auf den Zellzustand sowie die FTIR-Spektren und deren Identifizierung nehmen. Eine Beleuchtung dieser Fragestellung ist daher von Bedeutung für die Anwendung in der Praxis, da die FTIR-Technologie möglichst gut und reibungslos, jedoch ohne Qualitätsverluste, in die bestehenden Abläufe der Diagnostik im Brauereilabor integriert werden soll.

Zu diesem Zweck wurden fünf Inkubationsbedingungen auf die Identifizierung von 53 Stämmen (Anhang 2) aus dem in diesem Projekt gesammelten Pool an Isolaten getestet. Die Stämme gehörten zu 20 Arten, die sowohl unterschiedliche *Saccharomyces*- als auch Nicht-*Saccharomyces*-Hefen repräsentierten, und deckten somit einen relativ breiten Bereich der vorkommenden Hefen ab. Die für den Test verwendeten Medien waren Kupfersulfat-Agar, Lysin-Agar, Kristallviolett-Agar und Würze-Agar. Die Inkubationstemperatur wurde auf 28°C festgelegt, auf Würze-Agar wurden auch 37°C getestet. Als Positivkontrolle wurde jeder Stamm zudem auf YGC-Agar inkubiert. Betont werden muss, dass es sich bei den getesteten Konditionen nicht um die Anzucht für die FTIR-Messung handelte, die immer auf YGC-Agar bei 27°C erfolgte, sondern um eine Vorkultur, von der Impfmateriale für die FTIR-Anzucht abgenommen wurde. In Tabelle 9 sind die verwendeten Parameter sowie die erzielten Ergebnisse zusammengefasst.

Die Ergebnisse zeigen, dass Würze-Agar bei 28°C Inkubationstemperatur ohne jede Beeinträchtigung verwendet werden kann und die so erzielten Ergebnisse sich voll mit denen von YGC-Agar decken. Hier handelte es sich auch nicht um selektive Bedingungen, sondern um die Standardanzucht von Hefen im Brauereilabor. Während der standardisierten Anzucht für die FTIR-Messung müssen somit lediglich die vorhandenen Unterschiede in der Zellzusammensetzung des Inokulats ausgeglichen werden, was im Rahmen der 24-stündigen Inkubation ohne Einschränkung gelingt. Wird die Temperatur auf 37°C hochgesetzt, kann nur noch bei ca. 50 % der Stämme Wachstum verzeichnet werden, die Identifizierung dieser 50 % ist jedoch nicht beeinträchtigt.

Tabelle 9: Zusammenfassung der Bestimmung des Einflusses selektiver Inkubationsbedingungen der Vorkultur auf die Identifizierung von FTIR-Spektren.

Medium	Funktion	Wachstum	Identifizierung
YGC	Positivkontrolle, Detektion aller Arten		
Würze 28°C	Detektion aller Arten	Keine Ausfälle	Keine Beeinträchtigung
Würze 37°C	Detektion aller Arten, die bei 37°C wachsen	Wachstum nur bei 50 % der Stämme	Keine Beeinträchtigung, wenn Wachstum vorhanden
Kupfersulfat	Detektion der überwiegenden Anzahl von Nicht- <i>Saccharomyces</i> und <i>Saccharomyces</i> -Fremdhefen	Ausfälle bei <i>S. cerevisiae</i> und <i>S. pastorianus</i>	Keine Beeinträchtigung
Lysin	Detektion von Nicht- <i>Saccharomyces</i> -Hefen	Auch viele <i>Saccharomyces</i> -Hefen wachsen (in Minikolonien)	Identifizierung von <i>Saccharomyces</i> -Hefen geringfügig beeinträchtigt Nicht- <i>Saccharomyces</i> -Hefen nicht beeinträchtigt
Kristallviolett	Detektion von <i>Saccharomyces</i> -Fremdhefen	Auch 30 % der Nicht- <i>Saccharomyces</i> -Hefen wachsen	Keine Beeinträchtigung

Ähnliche Ergebnisse lassen sich auch bei den drei Selektivmedien Kupfersulfat-, Lysin- und Kristallviolett-Agar machen. Abgesehen von der Tatsache der nicht vollständigen Selektivität (siehe Tabelle 9), ist die Identifizierbarkeit der Hefen nicht herabgesetzt. Eine Ausnahme bilden einige *Saccharomyces*-Hefen auf Lysin-Agar, deren Identifizierung leicht beeinträchtigt war. Vermutlich war hier der physiologische Zustand einiger Stämme nach der Kultur auf Lysin-Agar soweit gestört, dass sie sich davon innerhalb der 24 Stunden Anzucht auf YGC-Agar nicht erholen konnten.

Zusammenfassend lässt sich feststellen, dass, wenn das Protokoll für die standardisierte Anzucht der Hefen für die FTIR-Messung eingehalten wird, die Nutzung von Selektivmedien zur Detektion von Fremdhefen nicht problematisch und die Identifizierung der resultierenden FTIR-Spektren in keiner Weise beeinträchtigt ist. Die in der Brauereiindustrie gängigen Medien sind mit der Nutzung von FTIR-Spektroskopie somit voll kompatibel und gut zu vereinbaren. Eine kleine Einschränkung muss nur für die Hefen gemacht werden, die aufgrund mangelnder Selektivität u.U. doch wachsen, aber nur kleine Kolonien bilden, hier empfiehlt sich eine zusätzliche Passage auf Würze oder YGC-Agar, um die volle Stoffwechselaktivität zu reaktivieren.

5. Fazit

Im vorliegenden Projekt B95 wurde die FTIR-Spektroskopie für den Einsatz zur Identifizierung von Fremdhefen im Brauereibereich weiterentwickelt und stellt nun eine einfache und zuverlässig anzuwendende Analytik dar. Für Kulturhefen wurde ein System entwickelt, das sehr zuverlässig obergärige und untergärige Hefen differenziert und eine Einteilung obergäriger Stämme in drei Gruppen vornimmt. Die Kombination einer umfassenden Referenzdatenbank auf Basis der Spektrometersoftware OPUS mit spezifisch konzipierten neuronalen Netzen für die Lösung spezieller Identifizierungsprobleme bietet sowohl eine breite Abdeckung unterschiedlichster Arten und Gattungen als auch exakte Identifizierungen eng verwandter Artgruppen. Als besonders vorteilhaft erweist sich hier die digitale Natur der Daten, die nach ihrer Erzeugung mit unterschiedlichsten Datenbanken ausgewertet werden können ohne weitere Analysen durchführen zu müssen. Auch eine Erweiterung der Datenbanken in der Zukunft ist möglich, wobei die auf OPUS basierende Bibliothek auf einfache Weise von den Nutzern selbst ergänzt werden kann. So kann jeder Betrieb über die Zeit eine an sein Hefenspektrum angepasste Datenbank entwickeln.

Die FTIR-Spektroskopie lässt sich als einfach anzuwendende Methode gut in Betriebsabläufe integrieren und bietet mit ihrer universellen Einsetzbarkeit ein hohes Potential für die Hefediagnostik. Mit Ausnahme der öffentlich verfügbaren DNA-Sequenzdatenbank existiert kein anderes Identifizierungssystem, das über eine so breit aufgestellte Referenzdatenbank verfügt und dennoch kaum laufende Kosten verursacht. Letztendlich ist die Entwicklung einer Datenbank aber nie abgeschlossen, da es kaum möglich ist, die gesamte vorkommende Biodiversität zu erfassen. So werden in der Praxis immer wieder Stämme auftauchen, die nicht identifiziert werden können und die eine gute Ergänzung für die Datenbank wären. Die Abteilung Mikrobiologie aktualisiert ihre Datenbanken deshalb jährlich und passt sie so den neuen Entwicklungen an.

Für die Einführung der Technologie in Brauereilabors sowie die Weiterentwicklung und weitergehende Adaption der Datenbanken bietet die Abteilung Mikrobiologie des ZIEL in Kooperation mit dem Forschungszentrum Weihenstephan für Brau- und Lebensmittelqualität (BLQ Weihenstephan) ihre Unterstützung an.

Literaturverzeichnis

- BÜCHL, N. R.; HUTZLER, M.; MIETKE-HOFMANN, H.; WENNING, M.; SCHERER, S.: Differentiation of probiotic and environmental *Saccharomyces cerevisiae* strains in animal feed. *J. Appl. Microbiol.* 109 (2010) 783-91.
- BÜCHL, N. R.; WENNING, M.; SEILER, H.; MIETKE-HOFMANN, H.; SCHERER, S.: Reliable identification of closely related *Issatchenkia* and *Pichia* species using artificial neural network analysis of Fourier-transform infrared spectra. *Yeast.* 25 (2008) 787-98.
- HELM, D.; LABISCHINSKI, H.; NAUMANN, D.: Elaboration of a procedure for identification of bacteria using Fourier-Transform IR spectral libraries: a stepwise correlation approach. *J. Microbiol. Methods.* 14 (1991) 127-142.
- KÜMMERLE, M.; SCHERER, S.; SEILER, H.: Rapid and reliable identification of food-borne yeasts by Fourier-transform infrared spectroscopy. *Appl. Environ. Microbiol.* 64 (1998) 2207-14.
- KURTZMAN, C. P.; ROBNETT, C. J.; BASEHOAR-POWERS, E.: Phylogenetic relationships among species of *Pichia*, *Issatchenkia* and *Williopsis* determined from multigene sequence analysis, and the proposal of *Barnettozyma* gen. nov., *Lindnera* gen. nov. and *Wickerhamomyces* gen. nov. *FEMS Yeast Res.* 8 (2008) 939-54.
- MAQUELIN, K.; KIRSCHNER, C.; CHOO-SMITH, L. P.; NGO-THI, N. A.; VAN VREESWIJK, T.; STAMMLER, M.; ENDTZ, H. P.; BRUINING, H. A.; NAUMANN, D.; PUPPELS, G. J.: Prospective study of the performance of vibrational spectroscopies for rapid identification of bacterial and fungal pathogens recovered from blood cultures. *J. Clin. Microbiol.* 41 (2003) 324-9.
- NAUMANN, D.; HELM, D.; LABISCHINSKI, H.: Microbiological characterizations by FT-IR spectroscopy. *Nature.* 351 (1991) 81-2.
- OBERREUTER, H.; SEILER, H.; SCHERER, S.: Identification of coryneform bacteria and related taxa by Fourier-transform infrared (FT-IR) spectroscopy. *Int. J. Syst. Evol. Microbiol.* 52 (2002) 91-100.
- SANDT, C.; SOCKALINGUM, G. D.; AUBERT, D.; LEPAN, H.; LEPOUSE, C.; JAUSSAUD, M.; LEON, A.; PINON, J. M.; MANFAIT, M.; TOUBAS, D.: Use of Fourier-transform infrared spectroscopy for typing of *Candida albicans* strains isolated in intensive care units. *J. Clin. Microbiol.* 41 (2003) 954-9.
- TIMMINS, E. M.; QUAIN, D. E.; GOODACRE, R.: Differentiation of brewing yeast strains by pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. *Yeast.* 14 (1998) 885-93.
- UDELHOVEN, T.; NAUMANN, D.; SCHMITT, J.: Development of a Hierarchical Classification System with Artificial Neural Networks and FT-IR Spectra for the Identification of Bacteria. *Appl. Spectrosc.* 54 (2000) 1471-9.
- UDELHOVEN, T.; NOVOZHILOV, M.; SCHMITT, J.: The NeuroDeveloper: a tool for modular neural classification of spectroscopic data. *Chemometrics and Intelligent laboratory systems.* 66 (2003) 219-26.
- WENNING, M.; BÜCHL, N. R.; SCHERER, S.: Species and strain identification of lactic acid bacteria using FTIR spectroscopy and artificial neural networks. *Journal of biophotonics.* 3 (2010) 493-505.
- WENNING, M.; SCHERER, S.; NAUMANN, D.: Infrared Spectroscopy in the Identification of Microorganisms. *In* M. Diem, P. R. Griffiths, and J. M. Chalmers (Hrsg.), *Vibrational Spectroscopy for Medical Diagnosis.* John Wiley & Sons, Ltd. Chichester (2008) S. 71-96.
- WENNING, M.; SEILER, H.; SCHERER, S.: Fourier-transform infrared microspectroscopy, a novel and rapid tool for identification of yeasts. *Appl. Environ. Microbiol.* 68 (2002) 4717-21.

Anhang 1

Kulturhefestämme

Art	Stammnummer	Getränketyt	Fermentation
<i>S. cerevisiae</i>	WHB 68	Weizen	obergärlg
<i>S. cerevisiae</i>	WHB 127	Weizen	obergärlg
<i>S. cerevisiae</i>	WHB 148	Alt	obergärlg
<i>S. cerevisiae</i>	WHB 149	Weizen	obergärlg
<i>S. cerevisiae</i>	WHB 165	Kölsch	obergärlg
<i>S. cerevisiae</i>	WHB 175	Weizen	obergärlg
<i>S. cerevisiae</i>	WHB 177	Kölsch	obergärlg
<i>S. cerevisiae</i>	WHB 184	Alt	obergärlg
<i>S. cerevisiae</i>	WHB 192	Alt	obergärlg
<i>S. cerevisiae</i>	WHB 208	Alt	obergärlg
<i>S. cerevisiae</i>	WHB 210	Ale	obergärlg
<i>S. cerevisiae</i>	WHB 211	Ale	obergärlg
<i>S. cerevisiae</i>	WHB 213	Ale	obergärlg
<i>S. cerevisiae</i>	Bordeaux	Wein	obergärlg
<i>S. cerevisiae</i>	WB4	Brennerei	obergärlg
<i>S. cerevisiae</i>	WS2	Sekt	obergärlg
<i>S. pastorianus</i>	WHB 34/70	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 34/78	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 59	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 66/70	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 69	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 84	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 120	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 128	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 168	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 172	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 180	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 194	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	WHB 204	Pilsner/Lager	untergärlg
<i>S. pastorianus</i>	CBS 1484	offizieller Stamm	untergärlg
<i>S. pastorianus</i>	CBS 1504	offizieller Stamm	untergärlg
<i>S. pastorianus</i>	CBS 1513	offizieller Stamm	untergärlg
<i>S. pastorianus</i>	CBS 6903	offizieller Stamm	untergärlg
<i>S. pastorianus</i>	NBRC 2003	offizieller Stamm	untergärlg

WHB: Weihenstephaner Hefebank (Freising, Deutschland), CBS: Centraalbureau voor Schimmelcultures (Utrecht, Niederlande), NBRC: NITE Biological Resource Center (Japan).

Anhang 2

Stämme, die für die Bestimmung des Einflusses selektiver Inkubationsbedingungen auf FTIR-Spektren genutzt wurden.

Art	Stämme	Art	Stämme
<i>Candida boidinii</i>	3	<i>Pichia guilliermondii</i>	5
<i>Candida parapsilosis</i>	2	<i>Pichia membranifaciens</i>	4
<i>Candida pararugosa</i>	1	<i>Rhodotorula mucilaginosa</i>	1
<i>Clavispora lusitaniae</i>	2	<i>Saccharomyces bayanus / pastorianus</i>	1
<i>Cryptococcus magnus</i>	1	<i>Saccharomyces cerevisiae</i>	10
<i>Debaryomyces hansenii</i>	1	<i>Saccharomyces pastorianus</i>	6
<i>Issatchenkia orientalis</i>	2	<i>Schizosaccharomyces pombe</i>	1
<i>Pichia anomala</i>	7	<i>Torulasporea delbrueckii</i>	1
<i>Pichia fabianii</i>	1	<i>Williopsis californica</i>	1
<i>Pichia galeiformis</i>	1	<i>Yarrowia lipolytica</i>	1

Anhang 3

Umbenennungen von Arten der Gattungen *Pichia* und *Issatchenkia* nach Kurtzman et al. (2008).

Ursprüngliche Bezeichnung	Neue Bezeichnung
<i>Pichia anomala</i>	<i>Wickerhamomyces anomalus</i>
<i>Pichia fabianii</i>	<i>Lindnera fabianii</i>
<i>Pichia jadinii</i>	<i>Lindnera jadinii</i>
<i>Williopsis californica</i>	<i>Barnettozyma californica</i>