

Aus dem Institut für Immunologie  
des Fachbereichs Veterinärmedizin  
der Freien Universität Berlin

und

aus der Arbeitsgruppe Bioinformatik und Molekularbiologie  
der Charité-Universitätsmedizin Berlin

**Die Piggeldat-Datenbank als Fundament zur Untersuchung der  
Auswirkungen von Nahrungsfaktoren bei Ferkeln**

Inaugural-Dissertation  
zur Erlangung des akademischen Doktorgrades  
*Doctor of Philosophy (Ph.D.)*  
in *Biomedical Sciences*  
an der  
Freien Universität Berlin

vorgelegt von  
**Sven Olaf Twardziok, M. Sc.**  
Bioinformatiker  
aus Wyk auf Föhr

Berlin 2014

Journal-Nr.: 3786

Gedruckt mit Genehmigung des Fachbereichs Veterinärmedizin  
der Freien Universität Berlin

Dekan: Univ.-Prof. Dr. Jürgen Zentek  
Erster Gutachter: Univ.-Prof. Dr. Michael F. G. Schmidt  
Zweiter Gutachter: Prof. Dr. Paul Wrede  
Dritter Gutachter: Prof. Dr. Knut Reinert

*Deskriptoren (nach CAB-Thesaurus):*

pigs, probiotics, pig feeding, digestive tract, microbial flora, immune system, multivariate analysis, databases

Tag der Promotion: 06. Juli 2015

*Bibliografische Information der Deutschen Nationalbibliothek*

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.ddb.de> abrufbar.

ISBN: 978-3-86387-620-3

**Zugl.: Berlin, Freie Univ., Diss., 2014**

Dissertation, Freie Universität Berlin

**D 188**

Dieses Werk ist urheberrechtlich geschützt.

Alle Rechte, auch die der Übersetzung, des Nachdruckes und der Vervielfältigung des Buches, oder Teilen daraus, vorbehalten. Kein Teil des Werkes darf ohne schriftliche Genehmigung des Verlages in irgendeiner Form reproduziert oder unter Verwendung elektronischer Systeme verarbeitet, vervielfältigt oder verbreitet werden.

Die Wiedergabe von Gebrauchsnamen, Warenbezeichnungen, usw. in diesem Werk berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürfen.

This document is protected by copyright law.

No part of this document may be reproduced in any form by any means without prior written authorization of the publisher.

Alle Rechte vorbehalten | all rights reserved

© Mensch und Buch Verlag 2015 Choriner Str. 85 - 10119 Berlin  
verlag@menschundbuch.de – www.menschundbuch.de

*Für meine Mutter  
Doris Isabella Twardziok  
(1956 – 2011)*



# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>11</b>
1.1. Fütterungsversuche mit Schweinen	11
1.2. Wechselwirkungen zwischen Nahrungsfaktoren, gastrointestinaler Mikrobiota und Wirt	12
1.2.1. Intestinale Mikrobiota	12
1.2.2. Wechselwirkungen	13
1.3. Systembiologie und -omics-Technologien	14
1.3.1. <i>Metagenomics</i>	15
1.3.2. <i>Metatranscriptomics</i>	15
1.3.3. <i>Metaproteomics</i>	15
1.3.4. <i>Metabolomics</i>	16
1.3.5. Integration von <i>Multi-Omics</i> Messdaten	16
1.4. Verwaltung von Forschungsdaten	17
1.4.1. Datenverwaltung in Forschungsprojekten	17
<b>2. Zielsetzung</b>	<b>19</b>
<b>3. Datenbankanwendung Catria</b>	<b>21</b>
3.1. Erhebung der Anforderungen an die Datenbankanwendung	21
3.1.1. Funktionelle Anforderungen	21
3.1.2. Umfang der Datenbank	22
3.2. Entwurf von Catria	23
3.2.1. Das Catria-Datenschema	23
3.2.2. Das Catria-Datenbankschema	27
3.3. Programmierung von Catria	27
3.3.1. Definition der Datenbanktabellen	28
3.3.2. Deklaration der Datenobjekttypen	29
3.4. Verwendung von Catria	30
3.4.1. Aufbau einer Catria-Website	30
3.4.2. Datenintegration mit Catria	30
3.4.3. Datenverwaltung mit Catria	33
3.4.4. Datenexport mit Catria	34
<b>4. Die Piggeldat-Datenbank</b>	<b>37</b>
4.1. Installation der Datenbankanwendung	38
4.2. Integration der Daten	38
4.2.1. Integration der Fütterungsversuche	38
4.2.2. Integration der Messdaten	38

4.3.	Inhalt der Piggeldat-Datenbank . . . . .	39
4.3.1.	Der Fütterungsversuch <i>EFFT2010</i> . . . . .	39
4.3.2.	Der Fütterungsversuch <i>EFST2010</i> . . . . .	41
4.3.3.	Der Fütterungsversuch <i>ZNFT2011</i> . . . . .	42
4.3.4.	Der Fütterungsversuch <i>ZNST2011</i> . . . . .	43
4.3.5.	Der Fütterungsversuch <i>ZNTG2012</i> . . . . .	43
4.3.6.	Der Fütterungsversuch <i>ZNPR2012</i> . . . . .	44
4.3.7.	Der Fütterungsversuch <i>EZIV2013</i> . . . . .	44
<b>5.</b>	<b>Analyse der Auswirkung des Probiotikums <i>E. faecium</i> NCIMB 10415</b>	<b>47</b>
5.1.	Beschreibung der integrierten Datensätze . . . . .	47
5.2.	Datenverarbeitung . . . . .	50
5.3.	Hauptkomponentenanalyse . . . . .	52
5.4.	<i>Orthogonal Projection to Latent Structures</i> Analyse . . . . .	53
5.4.1.	Berechnung der optimalen Vorhersagemodelle . . . . .	55
5.4.2.	Permutationstest . . . . .	56
5.4.3.	Bedeutung der Merkmale . . . . .	56
5.4.4.	Ergebnisse der OPLS-Analyse . . . . .	57
5.5.	Vergleich von unterschiedlichen Zeitpunkten . . . . .	59
5.6.	Analyse der Merkmalsgruppen . . . . .	59
5.6.1.	Berechnung der <i>Enrichment Analysis</i> . . . . .	60
5.6.2.	Ergebnisse der <i>Enrichment Analysis</i> . . . . .	60
5.7.	Netzwerkanalyse . . . . .	62
5.7.1.	Berechnung der Graphen und Vergleich der Merkmalsgruppen . . . . .	62
5.7.2.	Ergebnisse der Netzwerkanalyse . . . . .	63
<b>6.</b>	<b>Diskussion</b>	<b>69</b>
6.1.	Die Datenbankanwendung Catria . . . . .	69
6.1.1.	Vorteile von <i>Open-Access</i> bei Fütterungsversuchen . . . . .	69
6.1.2.	Verwendung eines Open-Source-Content-Management-Systems . . . . .	70
6.1.3.	Drupal zur Verwaltung von Forschungsdaten . . . . .	70
6.1.4.	Das Catria-Datenmodell . . . . .	71
6.1.5.	Vergleich von Catria mit <i>large-scale</i> -Systemen . . . . .	71
6.2.	Die Piggeldat-Datenbank . . . . .	72
6.2.1.	Die Piggeldat-Datenbank als Fundament zur Datenanalyse . . . . .	72
6.2.2.	Die Piggeldat-Datenbank als Datenarchiv . . . . .	73
6.2.3.	Verwendung der Piggeldat-Datenbank im SFB 852 . . . . .	74
6.3.	Auswirkung der Fütterung des Probiotikums <i>E. faecium</i> NCIMB 10415 beim Schwein	74
6.3.1.	Durchführung der Datenverarbeitung . . . . .	75
6.3.2.	Untersuchung der Varianz des Datensatzes . . . . .	75
6.3.3.	Vorhersage der Behandlungsgruppen und Berechnung der Auswirkung auf einzelne Merkmale . . . . .	76
6.3.4.	Vergleich der Auswirkungen zwischen verschiedenen Zeitpunkten . . . . .	76
6.3.5.	Auswirkung auf Merkmalsgruppen . . . . .	77
6.3.6.	Auswirkung auf die Assoziation zwischen Merkmalsgruppen . . . . .	77
<b>7.</b>	<b>Zusammenfassung</b>	<b>79</b>

<b>8. Summary</b>	<b>81</b>
<b>Anhang</b>	<b>95</b>
<b>A. Das Catria-Datenmodul</b>	<b>97</b>
<b>B. Matrix-Notationen</b>	<b>99</b>
<b>C. OPLS-Berechnung und Vorhersage der Gruppenzugehörigkeit</b>	<b>101</b>
<b>D. Piggeldat-Datenbank</b>	<b>103</b>
<b>E. Merkmale des <i>EFFT2010</i> Fütterungsversuch</b>	<b>105</b>





# Abkürzungsverzeichnis

<b>Abkürzung</b>	<b>Bedeutung</b>
ANOVA	Varianzanalyse (engl. <i>Analysis of variance</i> )
CMS	Content-Management-System
CSV	<i>Comma-separated values</i>
CV	Kreuzvalidierung (engl. <i>cross-validation</i> )
DBMS	Datenbankmanagementsystem
DELSA	<i>Data-Enabled Life Sciences Alliance</i>
ER	<i>Entity-Relationship</i>
ES	<i>Enrichment score</i>
GALT	<i>Gut-associated lymphoid tissue</i>
GNU GPL	<i>GNU General Public License</i>
GMOD	<i>Generic Model Organism Database Project</i>
GUI	Grafische Benutzeroberfläche (engl. <i>Graphical User Interface</i> )
HTML	<i>HyperText Markup Language</i>
IEL	Intraepitheliale Lymphozyten
LIMS	<i>Laboratory information management system</i>
LTS	<i>Long-term support</i>
MIBBI	<i>Minimum Information for Biological and Biomedical Investigations</i>
NIPALS	<i>Non-linear iterative partial least squares</i>
OPLS	<i>Orthogonal projection to latent structures</i>
OTU	<i>Operational taxonomic unit</i>
PCA	Hauptkomponentenanalyse (engl. <i>Principle component analysis</i> )
PCR	<i>Polymerase chain reaction</i>
PHP	<i>PHP: hypertext preprocessor</i>
PRRSV	<i>Porcine reproductive and respiratory syndrome virus</i>
RNA-seq	RNA Sequenzierung (engl. <i>RNA Sequencing</i> )
SCFA	<i>Short-chain fatty acid</i>
SFB	Sonderforschungsbereich
SwIV	Schweineinfluenza (engl. <i>swine influenza virus</i> )
TGEV	<i>Transmissible gastroenteritis coronavirus</i>
URL	<i>Uniform resource locator</i>
VIP	<i>Variable importance of the projection</i>



# 1. Einleitung

## 1.1. Fütterungsversuche mit Schweinen

Das Hausschwein, *Sus scrofa domestica*, findet sowohl als Nutztier als auch als Modellorganismus eine breite Verwendung in der landwirtschaftlichen und in der biomedizinischen Forschung. Die Verwendung des Schweins als Modellorganismus profitiert hierbei immer mehr von den neuesten Fortschritten der Genom- und Proteom-Forschung an Schweinen (Bendixen et al. 2010<sup>[10]</sup>). Hierbei eignet sich das Schwein insbesondere zur Modellierung des menschlichen Verdauungstrakts (Heinritz et al. 2013<sup>[41]</sup>). Ein Schwerpunkt bei der Verwendung des Schweins als Versuchstier ist die Untersuchung der Auswirkung von Nahrungsfaktoren. Die Durchführung eines Fütterungsversuchs mit Schweinen kann dabei verschiedene gesellschaftliche, wissenschaftliche und ökonomische Ziele verfolgen:

- **Behandlung von chronischen Erkrankungen**

Das Schwein eignet sich zur Untersuchung von chronischen Erkrankungen des Verdauungstrakts. Der Einsatz von verschiedenen Nahrungsfaktoren kann die Behandlung von chronischen Darmentzündungen unterstützen (Bassaganya-Riera und Hontecillas 2006<sup>[6]</sup> und Young et al. 2012<sup>[126]</sup>).

- **Leistungssteigerung von Masttieren**

Die Verwendung von bestimmten Nahrungszusätzen kann die Gewichtszunahme von Mastschweinen erhöhen und die hohe Sterblichkeitsrate von Ferkeln verringern. Verschiedene Probiotika (Zeyner und Boldt 2006<sup>[127]</sup>, Choi et al. 2011<sup>[20]</sup> und Wang et al. 2014<sup>[115]</sup>) und Antibiotika (Cromwell 2002<sup>[23]</sup>) können die Gewichtszunahme von Mastschweinen erhöhen oder auch das Vorkommen von Durchfallerkrankungen verringern, wobei in einigen Fällen die genaue Wirkungsweise der Nahrungszusätze noch unbekannt ist.

- **Qualitätssteigerung von Masttieren**

Über eine Variation der Ernährung lässt sich der Fettgehalt im Muskelgewebe des Schweins anpassen. Wichtige Einflussgrößen hierbei sind das Verhältnis des Anteils von Aminosäuren zum Anteil von Fett in der Nahrung und die Höhe der Energie in der Nahrung. Die optimale Zusammensetzung der Aminosäuren wiederum kann sowohl vom Genotyp des Tieres als auch von äußeren Faktoren abhängen (Pettigrew und Esnaola 2001<sup>[79]</sup>).

- **Reduzierung der Umweltverschmutzung**

Durch intensive Landwirtschaft mit Schweinen entsteht eine erhebliche Belastung für die Umwelt. Eine Anpassung der Fütterung von Aminosäuren an den individuellen Bedarf eines Schweins kann zum Beispiel eine Ausscheidung von Nitraten um bis zu 70 % verringern und somit die Belastung der Umwelt reduzieren (Aarnink und Verstegen 2007<sup>[1]</sup>).

Die Analyse der Auswirkung eines Nahrungsfaktors kann in der Forschung unter anderem durch einen Fütterungsversuch erfolgen. Ein Fütterungsversuch bezeichnet dabei die Durchführung eines

Experimentes, welches die Auswirkungen von mindestens einem Nahrungsfaktor bei einer Menge von Versuchstieren untersucht. Generell, wird ein Experiment zur strukturierten Erhebung von empirischen Daten im Rahmen eines empirischen Forschungsprojekts durchgeführt (Albers 2007<sup>[4]</sup> [Seite 18]). Bei einem Fütterungsversuch stellt ein einzelnes Versuchstier die Versuchseinheit dar.

Der Aufbau eines Fütterungsversuchs ergibt sich aus den experimentellen Variablen. Innerhalb von Behandlungsgruppen sind die Versuchstiere den gleichen Fütterungsbedingungen ausgesetzt. Neben den Behandlungsgruppen können noch weitere Variablen für das Experiment von Bedeutung sein. Dabei unterscheidet man man zwischen weiteren experimentellen Faktoren und Störvariablen. Multifaktorielle Experimente erlauben die Untersuchung von Wechselwirkungen zwischen unabhängigen Variablen (Albers 2007<sup>[4]</sup> [Seite 24]). Ein Beispiel für einen weiteren experimentellen Faktor ist der Einsatz von Pathogenen bei einer Teilgruppe der verwendeten Versuchstiere. Störvariablen können einen Einfluss auf die abhängigen Variablen haben. Das Alter oder das Geschlecht der Tiere sind Beispiele für mögliche Störvariablen. Durch den Aufbau des Experiments lassen sich Störvariablen kontrollieren (Albers 2007<sup>[4]</sup> [Seite 29]).

Ein Fütterungsversuch dient im Allgemeinen zur strukturierten Erhebung von sehr diversen Datensätzen aus unterschiedlichen biologischen Organisationsebenen, da die Auswirkung eines Nahrungsfaktors von einem komplexen Zusammenspiel zwischen verschiedenen äußeren und inneren Faktoren abhängen kann.

## 1.2. Wechselwirkungen zwischen Nahrungsfaktoren, gastrointestinaler Mikrobiota und Wirt

Bei Säugetieren im Allgemeinen und somit auch bei Schweinen hängt die Auswirkung einer Ernährung von einem komplexen Zusammenspiel zwischen den Bestandteilen der Nahrung, dem individuellen Immunsystem, der gastrointestinalen Mikrobiota und der Physiologie der Individuen ab (Zusammenfassungen von Kau et al. 2011<sup>[53]</sup>, Clemente et al. 2012<sup>[22]</sup>, Flint et al. 2012<sup>[32]</sup> und Tremaroli und Bäckhed 2012<sup>[106]</sup>).

### 1.2.1. Intestinale Mikrobiota

#### 1.2.1.1. Zusammensetzung der gastrointestinalen Mikrobiota

Die bakteriellen Gemeinschaften im Verdauungstrakt von Säugetieren stellen komplexe biologische Systeme dar. Beim Menschen übertrifft die gesamte Zellzahl der gastrointestinalen Mikrobiota die Zellzahl des Menschen um das Zehnfache (Xu und Gordon 2003<sup>[125]</sup>). Man schätzt, dass sich die Dichte der Bakterien von 10 Zellen pro Gramm Inhalt des Magens bis zu  $10^{12}$  Zellen pro Gramm Inhalt des Dickdarms erhöht (Sommer und Bäckhed 2013<sup>[100]</sup>). Die Stämme Firmicutes und Bacteroidetes sind die beiden häufigsten Stämme in den Fäkalien des Menschen (Qin 2010<sup>[81]</sup>, *Human Microbiome Project Consortium* 2012<sup>[46]</sup>) und auch vom Schwein (Looft et al. 2012<sup>[66]</sup>, Pajarillo et al 2014<sup>[77]</sup>). In einer Metagenom-Studie mit über 100 Personen konnten über 1000 verschiedene Spezies und über  $3 \times 10^6$  verschiedene Gene in den Fäkalien der Studienteilnehmer identifiziert werden (Qin et al. 2010<sup>[81]</sup>). Es wird angenommen, dass sich die gastrointestinale Mikrobiota vom Schwein aus einer vergleichbaren Anzahl an unterschiedlichen Spezies zusammensetzt (Leser 2002<sup>[64]</sup>).

## 1.2. Wechselwirkungen zwischen Nahrungsfaktoren, gastrointestinaler Mikrobiota und Wirt

### 1.2.1.2. Variabilität der gastrointestinalen Mikrobiota

Die Zusammensetzung der gastrointestinalen Mikrobiota bei Säugetieren ist sehr individuell und kann von verschiedenen äußeren und genetischen Faktoren abhängen (Clemente et al. 2012<sup>[22]</sup>). Bei gesunden ausgewachsenen Menschen ist die Variabilität der bakteriellen Gemeinschaft zwischen unterschiedlichen Individuen größer als innerhalb eines Individuums zu verschiedenen Zeitpunkten (Human Microbiome Project Consortium 2012<sup>[46]</sup>). Beim Schwein hängt die Zusammensetzung der gastrointestinalen Mikrobiota auch von der Rasse des Schweins ab (Pajarillo et al. 2014<sup>[77]</sup>). Die Entwicklung der Zusammensetzung der gastrointestinalen Mikrobiota bei Säugetieren lässt sich in drei Lebensabschnitte einteilen: von der Geburt bis zur Absetzung von der Muttermilch, die Umstellung der Ernährung auf feste Nahrung und eine späte Entwicklung im hohen Alter (Flint et al. 2012<sup>[32]</sup>). Der Vergleich von Metagenom-Datensätzen von gesunden Menschen hat gezeigt, dass sich die Gemeinschaften von unterschiedliche Individuen in einige wenige unterscheidbare Muster (engl. *Enterotypes*) entwickeln (Arumugam et al. 2011<sup>[5]</sup>). Diese Muster sind durch das relative Vorkommen von einigen wenigen Spezies geprägt und ihr Vorkommen hängt unter anderem von der langfristigen Ernährung der Individuen ab (Wu et al. 2011<sup>[124]</sup>). Im Gegensatz zu der hohen Variation der Zusammensetzung der gastrointestinalen Mikrobiota scheint die Verteilungen von funktionellen Gen-Gruppen im gastrointestinalen Metagenom relativ stabil zu sein (Franzosa et al. 2014<sup>[33]</sup>).

### 1.2.1.3. Aktivität der gastrointestinalen Mikrobiota

Die Bakterien im Verdauungstrakt führen unterschiedliche Stoffwechsellvorgänge durch (Nicholson et al. 2012<sup>[75]</sup>, Tremaroli und Bäckhed 2012<sup>[106]</sup>). Zu den wichtigen Stoffwechsellvorgängen mit potentiellen Auswirkungen auf den Wirt gehören die Produktion kurzkettiger Fettsäuren (SCFA, engl. *short-chain fatty acid*) durch Abbau von pflanzlichen Polysacchariden, die Metabolisierung von Gallensäuren und die Metabolisierung von Cholin (Nicholson et al. 2012<sup>[75]</sup>, Tremaroli und Bäckhed 2012<sup>[106]</sup>). Die große Bedeutung der metabolischen Aktivität der gastrointestinalen Mikrobiota zeigt sich bei der Zusammensetzung der bakteriellen Genome (Metagenom). Ein erheblicher Anteil der Gene im gastrointestinalen Metagenom vom Schwein steht mit dem Kohlenhydratstoffwechsel, dem Aminosäurestoffwechsel und dem Proteinstoffwechsel in Verbindung (Lamendella et al. 2011<sup>[61]</sup>).

## 1.2.2. Wechselwirkungen

### 1.2.2.1. Einfluss von Nahrungsfaktoren auf die intestinale Mikrobiota

Unterschiedliche Nahrungsfaktoren können die Zusammensetzung und die Aktivität von bakteriellen Gemeinschaften beeinflussen. Hierbei hat insbesondere das Verhältnis von pflanzlichen Kohlenhydraten, Proteinen und Nahrungsfetten einen Einfluss auf die Zusammensetzung der Mikrobiota (Scott et al. 2013<sup>[94]</sup>). In einer Studie von Wu et al. 2011<sup>[124]</sup> wurde der Zusammenhang zwischen der Ernährung von menschlichen Studienteilnehmern und deren gastrointestinaler Mikrobiota untersucht. In den Fäkalien der Studienteilnehmern konnten zwei unterschiedliche Gruppen (engl. *Enterotypes*) der gastrointestinalen Mikrobiota identifiziert werden. Das Vorkommen der ersten Gruppe korrelierte mit dem Verzehr von Aminosäuren, Proteinen und Fetten. Das Vorkommen der zweiten Gruppe korrelierte mit dem Verzehr von pflanzlichen Kohlenhydraten. Die Fütterung einer fettreichen Ernährung an Mäuse kann ebenfalls die Zusammensetzung und die Aktivität der gastrointestinalen Mikrobiota verändern (Daniel et al. 2014<sup>[25]</sup>). Der Vergleich

der Fütterung von unterschiedlichen pflanzlichen Ballaststoffen an Schweine zeigte veränderte Vorkommen von einigen Spezies im Ileum und im Colon sowie eine unterschiedliche Aktivität der gastrointestinalen Mikrobiota bei der Produktion von SCFAs (Ivarsson et al. 2014<sup>[49]</sup>). Die Zugabe von Antibiotika bei Ferkeln kann die Zusammensetzung und die Aktivität der Mikrobiota verändern und das Vorkommen von Antibiotikaresistenzgenen erhöhen (Looft et al. 2012<sup>[66]</sup>).

### 1.2.2.2. Zusammenspiel zwischen intestinaler Mikrobiota und Host

Die intestinale Mikrobiota kann durch ihre metabolische Aktivität oder über Interaktionen mit Epithelzellen ihren Wirt beeinflussen (Sommer und Bäckhed 2013<sup>[100]</sup>). Die gastrointestinale Mikrobiota kann unter anderem die Bildung der Mukusschicht, die Permeabilität und die Morphologie des Darmepithels, die Entwicklung des darmassoziierten lymphatischen Gewebes (GALT, engl. *gut-associated lymphoid tissue*) und die Modulation des Immunsystems beeinflussen (Sommer und Bäckhed 2013<sup>[100]</sup>). Das gastrointestinale Immunsystem sorgt für die Eingrenzung der gastrointestinalen Mikrobiota im Lumen des Verdauungstraktes (Hooper et al. 2012<sup>[43]</sup>). Hierbei hindert die Sekretion von unspezifischen antibakteriellen Stoffen die Bakterien am Durchdringen der Darmschleimhaut (Hooper et al. 2012<sup>[43]</sup>). Des Weiteren kann das Immunsystem die Zusammensetzung der gastrointestinalen Mikrobiota durch die Sekretion von spezifischen antibakteriellen Proteinen und Peptiden verändern (Hooper et al. 2012<sup>[43]</sup>). Verschiedene Studien konnten einen Zusammenhang zwischen veränderter gastrointestinaler Mikrobiota und dem Auftreten recht unterschiedlicher Erkrankungen wie zum Beispiel Fettleibigkeit, Morbus Crohn und Autoimmunerkrankungen feststellen (Clemente et al. 2012<sup>[22]</sup>).

## 1.3. Systembiologie und -omics-Technologien

Im Allgemeinen lässt sich das Zusammenspiel zwischen Ernährung, gastrointestinaler Mikrobiota und Wirt als ein komplexes und sehr dynamisches System betrachten. Das Ziel der Systembiologie ist das Verständnis der Struktur und der Dynamik von biologischen Systemen (Ideker et al. 2001,<sup>[48]</sup> Kitano 2002<sup>[55]</sup>). Hierzu kombiniert die Systembiologie umfassende Informationen aus unterschiedlichen biologischen Teilgebieten und steht somit im Kontrast zu reduktionistischen Ansätzen (Ahn et al. 2006<sup>[3]</sup>, Keusch 2006<sup>[54]</sup>). Generell unterscheidet die Systembiologie zwischen *top-down* und *bottom-up*-Ansätzen (Bruggeman und Westerhoff 2006<sup>[14]</sup>). *Top-down*-Ansätze umfassen die Charakterisierung aller Elemente eines biologischen Systems und deren Interaktionen durch die Integration von Daten aus unterschiedlichen -omics-Bereichen. Die Endung -omics markiert Bereiche der Biologie, welche alle Elemente eines Systems auf einer bestimmten biologischen Ebene untersuchen (Lamendella et al. 2012<sup>[62]</sup>). Messungen mit Technologien aus den -omics-Bereichen liefern mehrdimensionale Datensätze, welche umfassende Informationen über die einzelnen Elemente der betrachteten Ebene enthalten. *Bottom-up*-Ansätze nutzen die Struktur eines Systems um das mögliche Verhalten des Systems mathematisch vorherzusagen.

Zur Untersuchung der Auswirkung von Nahrungszusätzen und insbesondere zur Analyse der Interaktionen zwischen Wirt und gastrointestinaler Mikrobiota kommen Technologien aus den Bereichen *Metagenomics*, *Metatranscriptomics*, *Metaproteomics* und *Metabolomics* zum Einsatz (Lamendella et al. 2012<sup>[62]</sup> und van Baarlen et al. 2013<sup>[111]</sup>).

### 1.3.1. Metagenomics

*Metagenomics* bezeichnet die Untersuchung der gesamten genetischen Information (Metagenom) innerhalb eines biologischen Systems (Wooley et al. 2010<sup>[122]</sup>). Bei der Untersuchung der Auswirkung von Nahrungsfaktoren dienen *Metagenomics*-Studien zur Analyse der gastrointestinalen Mikrobiota (Lamendella et al. 2012<sup>[62]</sup>). Ein Ziel hierbei ist die Bestimmung der Verteilung von taxonomischen Einheiten (OTU, *operational taxonomic unit*) im Verdauungstrakt. Üblicherweise erfolgt eine Identifizierung der OTUs durch die Sequenzierung von verschiedenen Regionen des 16S rRNA-Gens (Leser et al. 2002<sup>[64]</sup>, Arumugam et al. 2011<sup>[5]</sup>, *Human Microbiome Project Consortium* 2012<sup>[46]</sup>, Pajarillo et al. 2014<sup>[77]</sup>). Die Datensätze einiger aktueller 16S-Metagenom-Analysen enthalten je nach Umfang der Studien zwischen  $8.2 \times 10^4$  und  $3.6 \times 10^6$  Nukleotidsequenzen und beschreiben das Vorkommen von bis zu 4000 OTUs (Arumugam 2011<sup>[5]</sup>, *Human Microbiome Project Consortium* 2012<sup>[46]</sup>, Daniel 2014<sup>[25]</sup>, Pajarillo 2014<sup>[77]</sup>). Ein anderes Ziel von *Metagenomics*-Studien ist die Analyse des gesamten genetischen Potentials der Organismen. Eine solche ungerichtete Sequenzierung erlaubt beispielsweise die Untersuchung der Verteilung von funktionellen Gen-Gruppen (Lamendella et al. 2012<sup>[62]</sup>). Eine Studie von Lamendella aus dem Jahre 2011<sup>[61]</sup> charakterisierte funktionelle Gen-Gruppen des Metagenoms der gastrointestinalen Mikrobiota von Schweinen mit über  $5 \times 10^5$  Nukleotidsequenzen.

### 1.3.2. Metatranscriptomics

*Metatranscriptomics* bezeichnet die unspezifische Untersuchung der Genexpression innerhalb eines biologischen Systems (Lamendella et al. 2012<sup>[62]</sup>). Die Sequenzierung von mRNA (*RNA-seq*) ist als wichtiges Werkzeug zur Quantifizierung der Genexpression etabliert (Wang et al. 2009<sup>[117]</sup>). Bei der Untersuchung der Auswirkung von Nahrungsfaktoren dienen *Metatranscriptomics*-Studien unter anderem zur Charakterisierung der Aktivität der gastrointestinalen Mikrobiota (Lamendella et al. 2012<sup>[62]</sup>). In einer Studie zur Analyse der Genexpressionen der gastrointestinalen Mikrobiota des Menschen konnten mit dem Verfahren *RNA-Seq* über  $4.9 \times 10^5$  Sequenzen analysiert werden (Gosalbes et al. 2011<sup>[37]</sup>). Ein anderer *RNA-Seq*-Datensatz zur Untersuchung der Auswirkung von einem Muttermilchersatz bei Ferkeln enthielt über  $7.6 \times 10^5$  Sequenzen (Poroyko et al. 2010<sup>[80]</sup>).

### 1.3.3. Metaproteomics

*Metaproteomics* bezeichnet die unspezifische Untersuchung der Proteine innerhalb eines biologischen Systems (Lamendella et al. 2012<sup>[62]</sup>). Generell dienen Massenspektroskopie-Ansätze zur unspezifischen Analyse eines Proteoms (Aebersold und Mann 2003<sup>[2]</sup>). Die Untersuchung des gastrointestinalen Metaproteoms lässt sich zum Beispiel mit der sogenannten *Shotgun-Metaproteomics*-Analyse durchführen (Verberkmoes et al. 2009<sup>[113]</sup>). Bei dem Verfahren wird zuerst der gesamte Proteingehalt einer Probe isoliert, gereinigt und mit Trypsin aufgespalten. Die Peptide können dann durch die Kombination aus einem Chromatographieverfahren und Massenspektrometrie identifiziert werden. Bei einer Anwendung dieser Methode konnten etwa 2000 Proteine identifiziert werden (Verberkmoes et al. 2009<sup>[113]</sup>). In einer anderen *Metaproteomics*-Studie zur Analyse einer fettreicher Ernährung auf die intestinale Mikrobiota von Mäusen wurden das Vorkommen von 1760 Proteinen charakterisiert (Daniel et al. 2014<sup>[25]</sup>).

### 1.3.4. *Metabolomics*

*Metabolomics* bezeichnet die Identifizierung und Charakterisierung von Metaboliten innerhalb eines biologischen Systems (Dettmer et al. 2007<sup>[28]</sup>). Die Analyse eines Metaboloms (*Metabolic Profiling*) erfolgt in der Regel durch die Kombination aus einem Chromatographieverfahren und Massenspektrometrie oder durch die Anwendung einer Kernspinresonanzspektroskopie (Dettmer et al. 2007<sup>[28]</sup>). Mit einer *Metabolomic*-Analyse erhält man eine Menge von Merkmalen, welche das zugehörige Metabolom beschreibt. Die Datensätze aus einer solchen Analyse können die Messwerte von über 15.000 Merkmalen enthalten (Datta und DePadilla 2006<sup>[27]</sup>). Über Datenbanken lassen sich die Merkmale einzelnen Metaboliten zuordnen. Die Anzahl der identifizierten Metabolite im menschlichen Körper ist in den letzten 10 Jahren gestiegen. Die Metabolom-Datenbank *Human Metabolome Database* enthielt im Jahre 2007 noch 2180 unterschiedliche Metabolite und im Jahre 2012 über  $4 \times 10^4$  unterschiedliche Metabolite (Wishart et al. 2012<sup>[120]</sup>). Bei der Untersuchung der Auswirkung von Nahrungsfaktoren dienen *Metabolomics*-Studien unter anderem zur Charakterisierung von bakteriellen Metaboliten im Verdauungstrakt (Lamendella et al. 2012<sup>[62]</sup>). In einer *Multi-Omics*-Studie mit Mäusen zum Vergleich einer Ernährung mit einem hohen Anteil an Kohlenhydraten und einer Ernährung mit einem hohen Anteil an Fetten konnten 2534 unterschiedliche Merkmale und damit 86 unterschiedlich vorkommende Metabolite im Verdauungstrakt der Mäuse identifiziert werden (Daniel et al. 2014<sup>[25]</sup>).

### 1.3.5. Integration von *Multi-Omics* Messdaten

Studien zur Analyse der Wechselwirkungen zwischen der Ernährung, der gastrointestinalen Mikrobiota und dem Wirt profitieren von der Integration von Messdaten aus verschiedenen *Omics*-Bereichen (van Baarlen 2013<sup>[111]</sup>). Eine Studie von Daniel 2014<sup>[25]</sup> zur Analyse der Auswirkung von fettreicher Ernährung bei Mäusen konnte durch die Integration von Metagenom-Daten, Metatranskriptom-Daten und Metabolom-Daten eine veränderte Zusammensetzung der gastrointestinalen Mikrobiota mit einem veränderten Verhalten der gastrointestinalen Mikrobiota in Verbindung bringen. Eine Studie von Franzosa et al. 2014<sup>[33]</sup> zur Analyse der gastrointestinalen Mikrobiota des Menschen konnte durch die Integration von Metagenom-Daten mit Metatranskriptom-Daten die unterschiedlichen Regulationen von spezifischen Gen-Familien zeigen.

Die Grundlage zur Integration von Messdaten ist die Verknüpfung der Messdaten mit Metadaten (Wruck et al. 2014<sup>[123]</sup>). Für einen Datensatz können Metadaten beispielsweise angeben, von welchen Tieren dieser Datensatz stammt und mit welchen Techniken die Messung durchgeführt wurde. Es existieren verschiedene Metadaten-Standards mit dem Ziel die Reproduzierbarkeit und die Integration von Messdaten zu verbessern (Taylor et al. 2008<sup>[105]</sup>, Kolker et al. 2014<sup>[58]</sup>). Das Projekt *Minimum Information for Biological and Biomedical Investigations* (MIBBI) stellt verschiedene Standards zur Beschreibung der Durchführung von verschiedenen biologischen Methoden zur Verfügung, wobei der Schwerpunkt der Standards bei der verbesserten Reproduzierbarkeit von Messungen liegt (Taylor et al. 2008<sup>[105]</sup>). Das Projekt *Data-Enabled Life Sciences Alliance* (DELSA) publizierte im Jahr 2014 die Metadaten-Kontrollliste *Multi-Omics Metadata Checklist* zur Integration von *Multi-Omics*-Daten (Kolker et al. 2014<sup>[58]</sup>).



## 1.4. Verwaltung von Forschungsdaten

Die strukturierte Verwaltung von biologischen Forschungsdaten erfolgt üblicherweise über eine Datenbank. Eine Datenbank ist eine gezielte Sammlung von zusammengehörigen Daten (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 18]). Die Erstellung und die Verwaltung einer Datenbank erfolgt dabei mit einem Datenbankmanagementsystem (DBMS) (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 19]). Ein Datenbanksystem umfasst eine Datenbank-Instanz in Verbindung mit einem DBMS (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 19]). Datenbanken sind mittlerweile eine unverzichtbare Grundlage in allen Bereichen der Forschung in den Biowissenschaften und ihre Anzahl ist in den letzten 5 Jahren stark angestiegen. Die Meta-Datenbank MetaBase (Bolser et al. 2012<sup>[12]</sup>) listet 1802 unterschiedliche biologische Datenbanken auf (Stand 13.08.2014). Die Ziele der biologischen Datenbanken können dabei variieren. Generell kann man zwischen öffentlichen Datenarchiven und Projekt-Datenbanken unterscheiden. Das Ziel von Datenarchiven ist die umfassende Speicherung von Informationen zum Aufbau, zur Struktur oder zur Interaktionen von biologischen Komponenten. Es folgen einige Beispiele für Datenbanken aus den Bereichen Biologie und Biomedizin:

- Das öffentliche Datenarchiv *GenBank* enthält die genetischen Informationen mit über  $1.7 \times 10^8$  Nukleotidsequenzen von unterschiedlichen Spezies<sup>[74]</sup>.
- Die Datenbank *Drug2Gene* enthält mögliche Interaktionen zwischen Genen und Wirkstoffen (Roider et al. 2014<sup>[84]</sup>). Die Integration der Daten erfolgt dabei auf Basis von Literatursuchen, Messungen der Bindungsaffinität oder durch die Analysen der Aktivität von Genen. Der freie Zugriff auf die Datenbank erfolgt über die Website <http://drug2gene.com/>.
- Die Datenbank *PhenomicDB* enthält mögliche Zusammenhänge zwischen Genotypen und Phänotypen bei unterschiedlichen Spezies (Groth et al. 2007<sup>[38]</sup>). Die integrierten Daten basieren auf *Knockout*-Experimenten. Der freie Zugriff auf die Datenbank erfolgt über die Website <http://www.phenomicdb.de/>.
- Die Datenbank *PHI-base* enthält Informationen über mögliche Interaktionen zwischen Pathogenen und Wirt-Organismen (Winnenburg et al. 2008<sup>[119]</sup>). Der freie Zugriff auf die Datenbank erfolgt über die Website <http://www.phi-base.org/>.

### 1.4.1. Datenverwaltung in Forschungsprojekten

Projekt-Datenbanken dienen zur Datenverwaltung und zur Datenintegration von Forschungsdaten innerhalb von Forschungsprojekten (Schmitt und Burchinal 2011<sup>[91]</sup>). Die Schwierigkeiten hierbei liegen meistens bei der Integration sehr diverser Datensätze aus unterschiedlichen Instituten (Schmitt und Burchinal 2011<sup>[91]</sup>). Zu den wichtigen Aufgaben der zugehörigen Datenbankanwendungen gehören die Archivierung der Messdaten und die Regelung des Zugriffs auf die Daten über Benutzerschnittstelle (Wruck et al. 2014<sup>[123]</sup>). Viele Forschungsprojekte nutzen bestehende Datenbankanwendungen und passen diese an die eigenen Bedürfnisse an, da nur selten Anwendungen existieren, die alle geforderten Datentypen unterstützen und alle gewünschten Funktionalitäten implementieren (Schmitt und Burchinal 2011<sup>[91]</sup>). Hierbei erfolgt in vielen Forschungsprojekten die Verwaltung der Forschungsdaten über eine Website unter der Verwendung eines Content-Management-Systems (CMS) (Mooney und Baenziger 2008<sup>[73]</sup> und Schmitt und Burchinal 2011<sup>[91]</sup>). Dieses Vorgehen hat den Vorteil, dass der Aufruf der Datenbankanwendung unabhängig vom Standort und vom Betriebssystem der Nutzer erfolgen kann. Der

## Einleitung

Transregio/Sonderforschungsbereich 32 nutzt Beispielsweise eine angepasste Website in Kombination mit einem Dateisystem und einer MySQL-Datenbank zur Verwaltung der Messdaten und Metadaten (Curdt et al. 2012<sup>[24]</sup>). Der Transregio/Sonderforschungsbereich 57 und der Sonderforschungsbereich 832 nutzen die Datenbankanwendung *FileMaker* zur Verwaltung der Daten und zur Präsentation der Daten über eine Webschnittstelle (Huss et al. 2011<sup>[47]</sup>). Der Sonderforschungsbereich 882 nutzt die Drupal Distribution *OpenAtrium* zur Datenverwaltung (Friedhoff et al. 2013<sup>[34]</sup>).

## 2. Zielsetzung

Der Sonderforschungsbereich (SFB) 852 „Ernährung und intestinale Mikrobiota - Wirtsinteraktionen beim Schwein“ untersuchte in den Jahren 2010 bis 2013 die Auswirkungen des Probiotikums *Enterococcus faecium* NCIMB 10415 und des Fütterungszusatzes Zink auf die Gesundheit von Ferkeln. Hierfür führte der SFB 852 sieben Fütterungsversuche durch. Das Ziel der Messungen war die Untersuchung der Auswirkung der Nahrungsfaktoren auf das Zusammenspiel zwischen der intestinalen Mikrobiota, dem Immunsystem und dem Wirt durch die Integration von Messdaten aus verschiedenen biologischen Organisationsebenen.

Das Ziel dieser Arbeit besteht in der Implementation der Piggeldat-Datenbank zur Verwaltung und zur Integration von *Multi-Omics*-Messdaten aus den Fütterungsversuchen des SFB 852. Die Piggeldat-Datenbank dient dabei zur strukturierten Archivierung der Forschungsdaten des SFB 852 und als Fundament für die Durchführung von Analysen der Auswirkungen von Fütterungszusätzen auf Basis der integrierten Messdaten. Des Weiteren erleichtert das zentrale Datenarchiv den Austausch der Messdaten zwischen den Mitgliedern des Forscherverbands und es unterstützt die Online-Publikation von Messdaten.

Zur Implementation der Piggeldat-Datenbank erfolgt die Entwicklung der Datenbankanwendung Catria. Catria dient zur Verwaltung von *Multi-Omics*-Messdaten aus Fütterungsversuchen. Der Entwurf von Catria berücksichtigt die Verwendung der Anwendung innerhalb von Forschungsprojekten mit verschiedenen beteiligten Wissenschaftlern und Instituten. Die Mitglieder der Forschungsprojekte können mit Catria die eigenen Messdaten in eine Datenbank integrieren. Nach der Integration der Daten können die Nutzer den Zugriff auf die eigenen Daten kontrollieren. Die Implementation von Catria verläuft als Webanwendung. Somit können die Nutzer die Anwendung unabhängig vom eigenen Betriebssystem und vom eigenen Standort ausführen. Ebenfalls können hierdurch externe Nutzer auf freigegebene Daten zugreifen, womit sich Catria auch zur Online-Publikation von Messdaten eignet.

Die Integration der Messdaten mit der Piggeldat-Datenbank dient zur Analyse der Auswirkung der Zugabe des Probiotikums *E. faecium* NCIMB 10415 bei Ferkeln. In vorherigen Analysen mit den selben Daten konnten bereits Effekte des Probiotikums auf die Beschaffenheit des Darmepithels (Klingspor et al. 2013<sup>[56]</sup>), die Aktivierung von zytotoxischen T-Zellen (Kreuzer et al. 2012<sup>[60]</sup>) und die veränderte Genexpression von Zytokinen (Siepert et al. 2014<sup>[95]</sup>) nachgewiesen werden. Das Ziel der vorliegenden Untersuchung ist die systematische Untersuchung des gesamten Datensatzes zur *Knowledge Discovery* durch die Anwendung von multivariaten Verfahren. Hierbei wird insbesondere das Zusammenspiel zwischen der gastrointestinalen Mikrobiota, der Reaktion des Immunsystem und der Physiologie des Wirts analysiert.



## 3. Datenbankanwendung Catria

Die Datenbankanwendung Catria (engl. *Collaborative animal trials*) dient zur Integration, zur Verwaltung und zur strukturierten Archivierung von *Multi-Omics*-Messdaten aus Fütterungsversuchen. Die Entwicklung von Catria verlief in Anlehnung an die Verwendung von gängigen Phasenmodellen zur Softwareentwicklung (Goeken 2006<sup>[36]</sup> [Seite 96]) in drei Phasen:

1. Erhebung der Anforderungen an die Datenbankanwendung (funktionelle Anforderungen und Umfang der Datenbank)
2. Entwurf der Datenbankanwendung
3. Programmierung der Datenbankanwendung

### 3.1. Erhebung der Anforderungen an die Datenbankanwendung

#### 3.1.1. Funktionelle Anforderungen

In der Softwareentwicklung ist die genaue Beschreibung der Aufgaben und der funktionellen Anforderungen an das Programm der erste entscheidende Schritt (Goeken 2006<sup>[36]</sup> [Seite 110]). Die Formulierung der funktionellen Anforderungen kann dabei durch die Formulierung von Anwendungsfällen (engl. *Use-Cases*) erfolgen (Goeken 2006<sup>[36]</sup> [Seite 119ff]). Die funktionellen Anforderungen von Catria basierten auf persönlichen Gesprächen mit den vorgesehenen Nutzern der Datenbankanwendung, auf eigenen Erfahrungen beim Entwurf von Datenbankanwendungen und auf dem formulierten Ziel der Datenbankanwendung. Hierdurch erfolgte die Formulierung von insgesamt vier Anwendungsfällen:

- **Anwendungsfall 1 - Nutzerverwaltung**

Nutzer können sich zur Authentifizierung bei Catria anmelden. Catria unterscheidet die drei Nutzergruppen Administratoren, Projektmitglieder und Gäste. Projektmitglieder sind die Mitglieder des Forschungsprojekts. Administratoren können Einstellungen an der Software vornehmen und die anderen Projektmitglieder verwalten. Gäste sind externe Nutzer, die freigegebene Messdaten untersuchen können. Das Nutzersystem von Catria reguliert den Zugriff aller Nutzer auf die Daten der Datenbank und auf die Funktionen der Datenbankanwendung.

- **Anwendungsfall 2 - Datenintegration**

Projektmitglieder können Messdaten mit Catria in einer Datenbank speichern. Die Speicherung umfasst insbesondere Messdaten aus den Bereichen: *Metagenomics*, *Metatranscriptomics*, *Metaproteomics* und *Metabolomics*. Zur Integration der Messdaten erfolgt die Speicherung von Metadaten. Die Speicherung der Metadaten erfüllt dabei den DELSA-Standard zur Speicherung von Daten aus *Meta-Omics*-Technologien (Kolker et al. 2014

[58]). Zur Speicherung von Messdaten und Metadaten erfolgt ein Dateitransfer mit den jeweiligen Informationen von lokalen Computern auf den Catria-Server.

- **Anwendungsfall 3 - Datenverwaltung**

Projektmitglieder können die eigenen integrierten Daten verändern und wieder aus der Datenbank löschen. Ebenfalls können Projektmitglieder den Zugriff auf die eigenen Messdaten kontrollieren. Die Zugriffskontrolle reguliert den Zugriff aller Nutzer auf die Messdaten. Die Nutzer können nur dann auf fremde Messdaten zugreifen, wenn diese freigegeben sind.

- **Anwendungsfall 4 - Datenansicht**

Alle Nutzer können Catria zur Untersuchung der Daten verwenden. Catria stellt Messdaten und Metadaten in einem übersichtlichen Format dar. Die Nutzer können Daten aus der Datenbank in gängige Datenformate exportieren. Der Export der Daten erlaubt die Anwendung von lokalen Datenanalyse-Werkzeugen durch die jeweiligen Nutzer. Die Anzeige und der Export der Daten ist von der Zugriffskontrolle abhängig.

### 3.1.2. Umfang der Datenbank

Catria nutzt eine Datenbank zur Speicherung der Forschungsdaten aus Fütterungsversuchen. Der Umfang der Datenbank ergab sich aus der allgemeinen Beschreibung eines Fütterungsversuchs (Abschnitt 1.1) und den im vorigen Abschnitt beschriebenen funktionellen Anforderungen. Die Metadaten, die sich bei einer unbestimmten Menge von Fütterungsversuchen ergeben, ließen sich in fünf Objektgruppen einteilen:

- **Objektgruppe 1 - FÜTTERUNGSVERSUCH**

Die Datenbank dient zur Speicherung von Daten aus unterschiedlichen Fütterungsversuchen. Zu jedem Fütterungsversuch speichert die Datenbank eine ausführliche Beschreibung mit Informationen zum Aufbau des Experiments, zur Durchführung des Experiments, zu den verwendeten Tieren und zur Haltung der Tiere.

- **Objektgruppe 2 - VARIABLE**

In einem Fütterungsversuch werden unterschiedliche experimentelle Faktoren und Störvariablen berücksichtigt. Jede Variable gehört dabei zu genau einem Fütterungsversuch. Zu jeder Variable speichert die Datenbank eine ausführliche Beschreibung mit Informationen zu möglichen Ausprägungen der Variable.

- **Objektgruppe 3 - TIER**

Zu jedem Tierversuch gehört eine feste Menge von Tieren. Dabei gehört jedes Tier zu genau einem Fütterungsversuch. Zu jedem Tier speichert die Datenbank eine Beschreibung mit Informationen zu den einzelnen Individuen. Mit der Beschreibung kann die Datenbank zum Beispiel Abweichungen bei der Ernährung von einzelnen Tieren oder eventuelle medizinische Behandlungen festhalten. Jedes Tier wird durch Variablen (experimentelle Faktoren und Störvariablen) beschrieben. Die Datenbank speichert für jede relevante Kombination von Tieren und Variablen eine Ausprägung.

- **Objektgruppe 4 - MERKMAL**

Die Durchführung einer Messung bestimmt die Messwerte von verschiedenen Merkmalen. Jedes Merkmal kann dabei genau einer Messung zugeordnet werden. Zu jedem Merkmal speichert die Datenbank eine ausführliche Beschreibung und die physikalische Einheit des

Merkmals. Des Weiteren können Merkmale zu Merkmalsgruppen gehören. Mit Merkmalsgruppen lassen sich Merkmale in einen Zusammenhang mit anderen Merkmalen stellen.

- **Objektgruppe 5 - MESSUNG**

An jedem Tier können mehrere Messungen vorgenommen werden. Die Durchführung einer Messung kann dabei eine Menge von Tieren betreffen. Jede Messung ist genau einem Fütterungsversuch zugeordnet. Zu jeder Messung speichert die Datenbank eine ausführliche Beschreibung mit Informationen zur Durchführung der Messung. Hierzu zählen sowohl die Methoden der Datengewinnung als auch anschließende Datenverarbeitungen. Des Weiteren speichert die Datenbank zu jeder Messung Verweise auf alle wichtigen Publikationen, welche sich auf die Durchführung oder auf die Ergebnisse der Messung beziehen. Zu jeder Messung wird ein Status gespeichert. Dieser Status gibt an, ob andere Nutzer auf die Messwerte einer Messung zugreifen können. Die Messung eines Merkmals an einem Tier zu einem bestimmten Zeitpunkt liefert einen Messwert.

Neben den Forschungsdaten speichert die Datenbank noch Informationen über die Nutzer der Datenbank. Die Nutzer stellen eine weitere Objektgruppe dar.

- **Objektgruppe 6 - NUTZER**

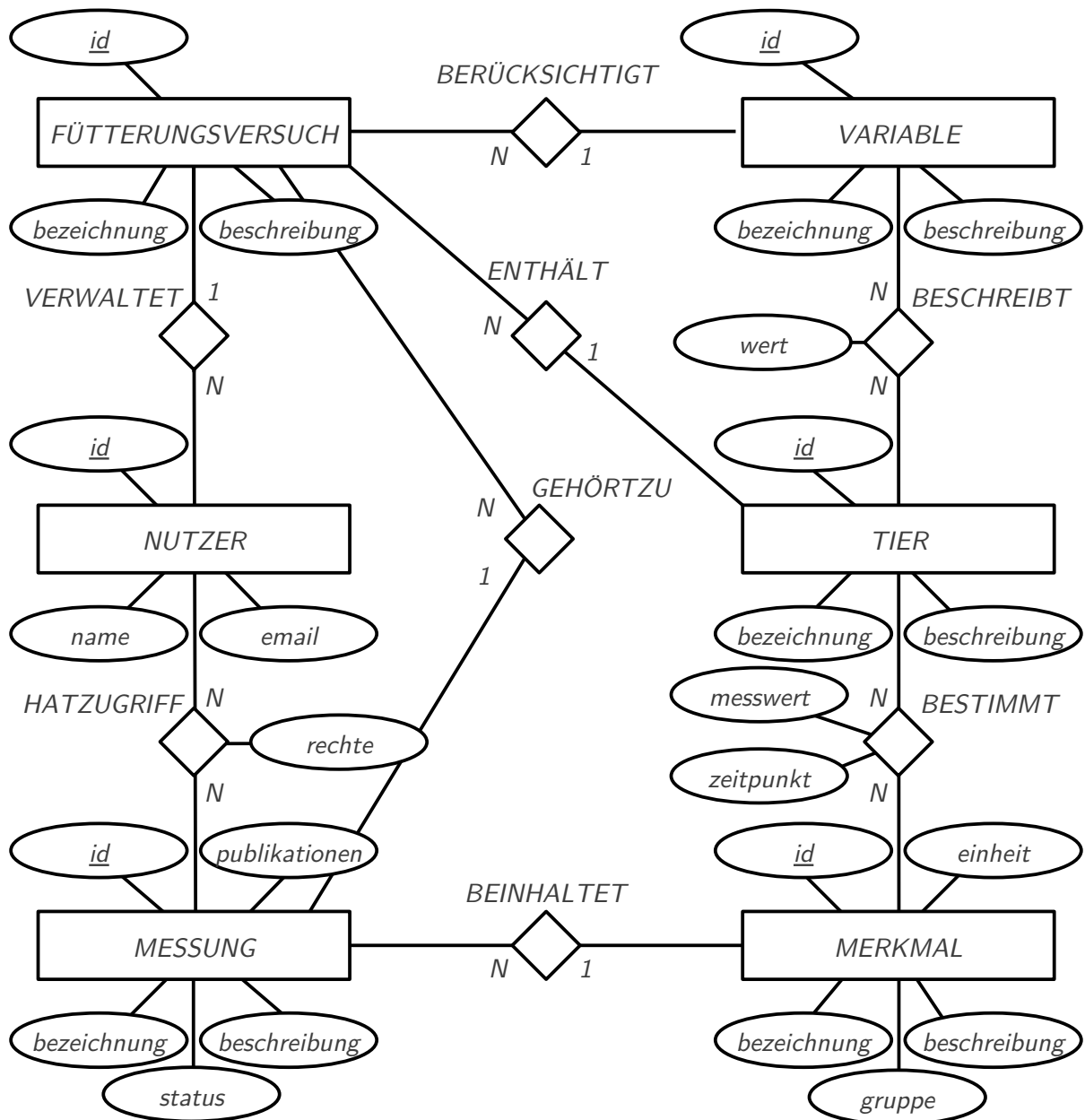
Die Datenbankanwendung speichert Daten zu den Nutzern der Datenbank. Die Datenbank repräsentiert einen Nutzer durch einen eindeutigen Namen zur Identifikation des Nutzers und durch eine E-Mail Adresse.

## 3.2. Entwurf von Catria

Aufbauend auf den Anforderungen aus Abschnitt 3.1 erfolgte im nächsten Schritt der Entwurf der Datenbankstruktur. Zunächst erfolgte hierzu die Formulierung eines konzeptuellen Datenschemas und anschließend eine Transformation des konzeptuellen Datenschemas in ein relationales Datenbankschema. Das konzeptuelle Datenschema basierte auf der schriftlichen Beschreibung des Datenbankumfangs und diente zur Spezifikation der Datenbankstruktur und zur Kommunikation der Struktur mit den Nutzern der Datenbankanwendung. Das Datenschema wurde nach den Konzepten des *Entity-Relationship-Modells* (ER-Modell) spezifiziert (Chen 1976<sup>[19]</sup>, Elmasri und Navathe 2009<sup>[31]</sup> [Seite 57 ff]). Das relationale Datenbankschema baute auf dem konzeptuellen Datenschema auf und diente zur technischen Umsetzung der relationalen Datenbank. Dieses Vorgehen wurde gewählt, da das CMS Drupal eine relationale Datenbank zur Speicherung von Daten verwendet.

### 3.2.1. Das Catria-Datenschema

Das ER-Modell beschreibt Daten mit Objekten (Entität, engl. *Entity*) und mit Beziehungen (engl. *Relationship*) zwischen den Entitäten. Dabei lassen sich die Entitäten zu Entitätstypen zusammenfassen und die Beziehungen lassen sich zu Beziehungstypen zusammenfassen. Aus den Datenbankanforderungen (Abschnitt 3.1.2) ergaben sich die sechs Entitätstypen *FÜTTERUNGSVERSUCH*, *VARIABLE*, *NUTZER*, *TIER*, *MESSUNG* und *MERKMAL*. Die acht Beziehungstypen *BERÜCKSICHTIGT*, *VERWALTET*, *ENTHÄLT*, *BESCHREIBT*, *GEHÖRTZU*, *HATZUGRIFF*, *BESTIMMT* und *BEINHALTET* repräsentieren die Assoziationen zwischen einzelnen Entitäten. Abbildung 3.1 stellt das Catria-Datenschema als ER-Schemadiagramm dar.



**Abbildung 3.1.:** ER-Schemadiagramm für die Catria Datenbank. Das Diagramm zeigt den Zusammenhang zwischen sechs Entitätstypen (Rechtecke) und neun Beziehungstypen (Rauten) des Catria-Datenschemas. Die Attribute der Entitätstypen und Beziehungstypen sind in Ovalen dargestellt. Linien kennzeichnen die Zugehörigkeit von Attributen zu Entitätstypen und Beziehungstypen und die Zugehörigkeit von Entitätstypen zu Beziehungstypen. An den Beziehungstypen sind die zugehörigen Kardinalitäten angegeben (1 zu N, N zu 1 oder N zu N).



### 3.2.1.1. Entitätstypen des ER-Schemas

Die fünf Entitätstypen *FÜTTERUNGSVERSUCH*, *VARIABLE*, *TIER*, *MESSUNG* und *MERKMAL* repräsentieren die Metadaten aus den Fütterungsversuchen. Jeder dieser Entitätstypen hat die drei Attribute *id*, *bezeichnung* und *beschreibung*. Das Attribut *id* dient als interner Identifikator für jede Entität eines Entitätstyps. Mit dem Attribut *beschreibung* kann jede Entität durch einen zusätzlichen Text beschrieben werden. Das Attribut *status* des Entitätstyps *MESSUNG* dient zur Zugriffskontrolle auf die Messdaten. Der Entitätstyp *NUTZER* repräsentiert die Nutzer der Datenbankanwendung. Die Datenbank beschreibt jeden Nutzer der Datenbankanwendung mit einer E-Mail Adresse (*email*) und einem Namen (*name*).

### 3.2.1.2. Beziehungstypen des Catria ER-Schemas

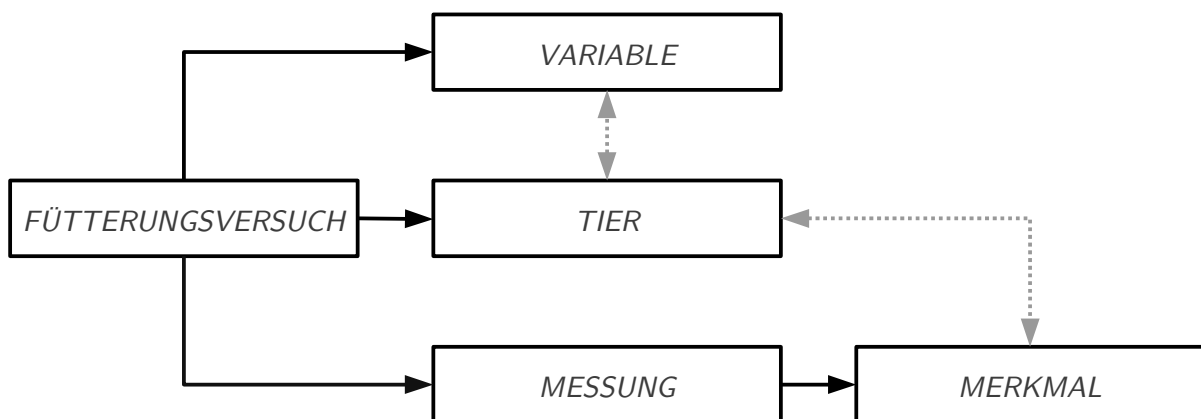
Die acht Beziehungstypen stellen die Beziehungen zwischen den einzelnen Entitäten dar. Jede Beziehung enthält zwei Entitäten der beteiligten Entitätstypen. Dabei hat jeder Beziehungstyp ein bestimmtes Kardinalitätsverhältnis. Dieses Kardinalitätsverhältnis definiert die Anzahl der möglichen Beziehungen, welche eine bestimmte Entität enthalten (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 73]).

- Der Beziehungstyp *BERÜCKSICHTIGT* ist ein  $N : 1$  Beziehungstyp zwischen den Entitätstypen *FÜTTERUNGSVERSUCH* und *VARIABLE*. Jeder Fütterungsversuch berücksichtigt eine beliebige Anzahl an Variablen und jede Variable wird von genau einem Fütterungsversuch berücksichtigt.
- Der Beziehungstyp *VERWALTET* ist ein  $N : 1$  Beziehungstyp zwischen den Entitätstypen *NUTZER* und *FÜTTERUNGSVERSUCH*. Jeder Nutzer kann beliebig viele Fütterungsversuche verwalten und jeder Fütterungsversuch wird von genau einem Nutzer verwaltet.
- Der Beziehungstyp *ENTHÄLT* ist ein  $N : 1$  Beziehungstyp zwischen den Entitätstypen *FÜTTERUNGSVERSUCH* und *TIER*. Jeder Fütterungsversuch enthält eine beliebige Anzahl an Tieren und jedes Tier ist in genau einem Fütterungsversuch enthalten.
- Der Beziehungstyp *BESCHREIBT* ist ein  $N : N$  Beziehungstyp zwischen den Entitätstypen *VARIABLE* und *TIER*. Jede dieser Beziehungen hat das Attribut *wert* zur Speicherung der Ausprägung einer Variable bei einem Tier. Jede Variable beschreibt eine beliebige Anzahl von Tieren und jedes Tier wird von einer beliebigen Anzahl von Variablen beschrieben.
- Der Beziehungstyp *GEHÖRTZU* ist ein  $N : 1$  Beziehungstyp zwischen den Entitätstypen *FÜTTERUNGSVERSUCH* und *MESSUNG*. Jede Messung gehört zu genau einem Fütterungsversuch und zu jedem Fütterungsversuch können mehrere Messungen gehören.
- Der Beziehungstyp *HATZUGRIFF* ist ein  $N : N$  Beziehungstyp zwischen den Entitätstypen *NUTZER* und *MESSUNG*. Jeder Nutzer kann Zugriffsrechte bei beliebig vielen Messungen haben und jede Messung kann Zugriffsrechte für beliebig vielen Nutzer haben. Der Beziehungstyp hat das Attribut *rechte* zur Unterscheidung von verschiedenen Zugriffsrechten.
- Der Beziehungstyp *BESTIMMT* ist eine  $N : N$  Beziehungstyp zwischen den Entitätstypen *TIER* und *MERKMAL*. Der Beziehungstyp hat die Attribute *messwert* und *zeitpunkt*. An einem Tier können beliebig viele Merkmale bestimmt werden und jedes Merkmal kann an beliebig vielen Tieren bestimmt werden.

- Der Beziehungstyp *BEINHALTET* ist ein  $N : 1$  Beziehungstyp zwischen den Entitätstypen *MESSUNG* und *MERKMAL*. Jede Messung beinhaltet mehrere Merkmale und jedes Merkmal gehört zu genau einer Messung.

### 3.2.1.3. Vereinfachte Darstellung des Catria-Datenschemas

Aus dem Catria-Datenschema lässt sich eine hierarchische Darstellung der Datenbank ableiten (Abbildung 3.2). Die Gruppe der Fütterungsversuche stellt hierbei die oberste Hierarchieebene dar. Zu jedem Fütterungsversuch gehört eine Menge von Variablen, eine Menge von Tieren und eine Menge von Messungen. Zu jeder Messung wiederum gehört eine Menge von Merkmalen.



**Abbildung 3.2.:** Vereinfachte Darstellung des Catria-Datenschemas. Das Diagramm zeigt den hierarchischen Zusammenhang zwischen den fünf Entitätstypen *FÜTTERUNGSVERSUCH*, *VARIABLE*, *TIER*, *MESSUNG* und *MERKMAL*. Schwarze Pfeile verdeutlichen die Hierarchie ( $N:1$  Beziehungstypen). Graue Pfeile markieren  $N:N$  Beziehungstypen zwischen den Entitätstypen.

### 3.2.2. Das Catria-Datenbankschema

Ein relationales Datenbankschema dient zur Spezifikation einer relationalen Datenbank (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 133ff]). Eine relationale Datenbank besteht in der Regel aus mehreren Relationen. Jede Relation lässt sich als eine Tabelle mit Einträgen betrachten. Hierbei repräsentieren die Spalten der Tabelle verschiedene Attribute und die Zeilen der Tabelle repräsentieren unterschiedliche Einträge. Ein Relationsschema definiert die Struktur einer Relation. Ein Relationsschema besteht aus einem Relationsnamen und einer Liste von Attributen. Die Gültigkeit einer Relation wird über Einschränkungen (engl. *constraints*) bestimmt (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 140ff]). Hierbei sind Einschränkungen der Wertebereiche (engl. *domain constrains*) und die Schlüsseleinschränkungen (engl. *key constrains*) von besonderer Bedeutung. Die Einschränkungen der Wertebereiche legen für jedes Attribut eine mögliche Wertemenge (Domäne) fest. Typische Wertemengen sind zum Beispiel die Menge der natürlichen Zahlen oder die Menge aller Zeichenkombinationen einer bestimmten Länge. Die Schlüsseleinschränkungen bestimmen für ein Relationsschema ein Attribut (oder eine Kombination von Attributen), deren Werte für die einzelnen Einträge (Tupel) der Relation einzigartig sein müssen und dienen zur Identifikation von Einträgen. Schlüsselattribute können im Relationsschema vermerkt werden, indem die Attribute unterstrichen notiert werden. Fremdschlüssel sind Attribute eines Relationsschemas, deren Werte sich auf Schlüssel in anderen Relationsschemas beziehen.

Durch die Anwendung des *ER-to-Relational Mapping* Algorithmus (Elmasri und Navathe 2009<sup>[31]</sup> [Seite 235ff]) wurde das Catria-Datenschema aus Abschnitt 3.2.1 in ein relationales Datenbankschema übersetzt. Das relationale Datenbankschema der Catria-Datenbank enthält neun Relationsschemas (Abbildung 3.3). Jeweils ein Relationsschema wurde für die sechs Entitätstypen definiert, ein Relationsschema wurde für den Beziehungstyp *BESCHREIBT* definiert, ein Relationsschema wurde für den Beziehungstyp *HATZUGRIFF* definiert und ein Relationsschema wurde für den Beziehungstyp *BESTIMMT* definiert. Für die Beziehungstypen *BERÜCKSICHTIGT*, *VERWALTET*, *ENTHÄLT GEHÖRTZU* und *BEINHALTET* wurden keine Relationsschemas definiert, da sich Beziehungstypen mit einem  $N : 1$  Kardinalitätsverhältnis über Fremdschlüssel in den zugehörigen Relationsschemas repräsentieren lassen. Die sechs Relationsschemas *FÜTTERUNGSVERSUCH*, *VARIABLE*, *NUTZER*, *TIER*, *MESSUNG* und *MERKMAL* enthalten das Schlüsselattribut *id* zur Identifizierung von einzelnen Einträgen in den jeweiligen Relationen. Die drei Relationsschemas *BESCHREIBT*, *HATZUGRIFF* und *BESTIMMT* enthalten zusammengesetzte Schlüssel aus Fremdschlüsseln, die auf Einträge in den beteiligten Relationen verweisen. Das relationale Datenbankschema verwendet insgesamt sechs unterschiedliche Domänen (*Dom\_id*, *Dom\_bez*, *Dom\_bes*, *Dom\_wert*, *Dom\_datum* und *Dom\_email*). Die Domänen der Attribute sind in Tabelle 3.1 dargestellt.

## 3.3. Programmierung von Catria

Die Implementierung von Catria erfolgte auf Basis des CMS Drupal 7<sup>[15]</sup>. Drupal bietet verschiedene Werkzeuge zum Aufbau und zur Verwaltung von dynamischen Websites an. Hierzu gehört unter anderem ein Nutzer-Management-System und die Unterstützung von einigen Standard-Datentypen wie zum Beispiel Artikel oder Blog-Einträge. Über individuelle Module lässt sich eine Drupal Installation mit neuen Datentypen und mit weiteren Funktionen erweitern (Vandyk 2009

$S = \{FÜTTERUNGSVERSUCH, VARIABLE, NUTZER, TIER, MESSUNG, MERKMAL, BESCHREIBT, HATZUGRIFF, BESTIMMT\}$

$FÜTTERUNGSVERSUCH(\underline{id}, \text{bezeichnung}, \text{beschreibung}, NUTZER.id)$

$VARIABLE(\underline{id}, \text{bezeichnung}, \text{beschreibung}, FÜTTERUNGSVERSUCH.id)$

$NUTZER(\underline{id}, \text{name}, \text{email})$

$TIER(\underline{id}, \text{bezeichnung}, \text{beschreibung}, FÜTTERUNGSVERSUCH.id)$

$MESSUNG(\underline{id}, \text{bezeichnung}, \text{status}, \text{beschreibung}, \text{publikationen}, NUTZER.id)$

$MERKMAL(\underline{id}, \text{bezeichnung}, \text{gruppe}, \text{einheit}, \text{beschreibung}, MESSUNG.id)$

$BESCHREIBT(VARIABLE.id, TIER.id, \text{wert})$

$HATZUGRIFF(NUTZER.id, MESSUNG.id, \text{rechte})$

$BESTIMMT(TIER.id, MERKMAL.id, \text{zeitpunkt}, \text{messwert})$

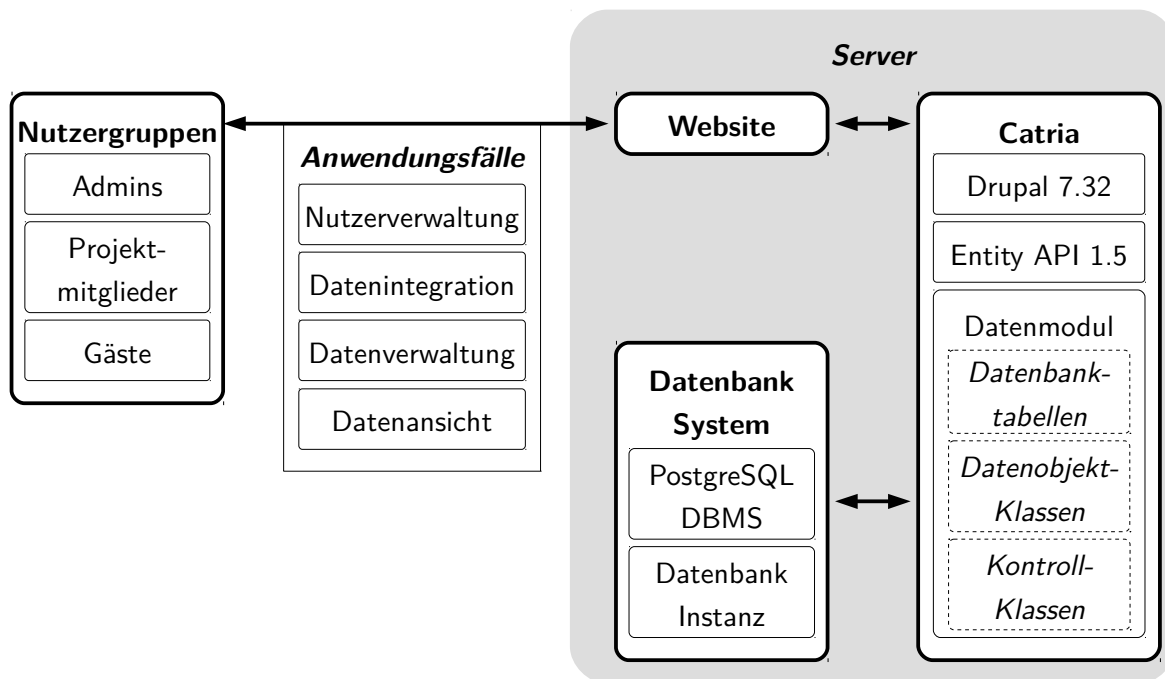
**Abbildung 3.3.:** Das relationale Datenbankschema  $S$  von Catria mit den neun enthaltenen Relationsschemas. Die Relationsschemas sind mit einem Namen und mit einer Menge von Attributen in Klammern dargestellt. Die Schlüsselattribute der Relationsschemas sind unterstrichen dargestellt.

[112][Seite 45ff]). Eine Distribution bezeichnet die Zusammenstellung einer Drupal-Installation mit individuellen Modulen und Anpassungen.

Catria besteht aus dem Drupal Kern (Version 7.32), dem Drupal Modul Entity API<sup>[128]</sup> (Version 7.x-1.5) und einem Datenmodul mit spezifischen Erweiterungen (Abbildung 3.4). Die Erweiterungen von Catria beinhalten die Definition der Datenbanktabellen, die Deklaration von Objektklassen zur Darstellung der Daten in einer Drupal-Umgebung und die Deklaration von Kontroll-Klassen zur Verarbeitung von Datenobjekten. Für das Catria-Datenmodul wurden insgesamt 34 Dateien in der Programmiersprache PHP (engl. *PHP: hypertext preprocessor*) erstellt (Tabelle A.1). Der Programmcode von Catria ist komplett offen und steht unter der Lizenz *GNU General Public License* (GNU GPL). Somit können Drupal-Entwickler den Code anpassen und neue Varianten von Catria entwerfen. Eine Validierung des Programmcodes erfolgte mit Hilfe des Programms *pareview.sh*<sup>[30]</sup>. Der Programmcode erfüllt die Standards der Drupal-Gemeinschaft. Diese Standards fordern unter anderem die Einhaltung eines festgelegten Formats, die ausführliche Kommentierung aller PHP-Klassen und aller PHP-Funktionen und die Einhaltung von Sicherheitsmaßnahmen. Das Catria-Datenmodul ist frei erhältlich und kann über die offizielle Drupal Website heruntergeladen werden<sup>[108]</sup>.

### 3.3.1. Definition der Datenbanktabellen

Das Catria-Datenmodul enthält die Definitionen der acht Relationsschemas *FÜTTERUNGSVERSUCH*, *VARIABLE*, *TIER*, *MESSUNG*, *MERKMAL*, *BESCHREIBT*, *HATZUGRIFF* und *BESTIMMT* aus dem relationalen Datenbankschema (Abbildung 3.3). Die Relation *NUTZER* gehört bereits zum Kern von Drupal. Die Definition der acht Relationsschemas erfolgt mit Hilfe



**Abbildung 3.4.:** Verwendung und Aufbau von Catria. Catria besteht aus dem Drupal-Kern, dem Modul Entity API und einem individuellen Datenmodul. Catria nutzt das *PostgreSQL* Datenbankmanagementsystem zur Speicherung der Daten. Die drei Nutzergruppen verwenden die Catria-Website zur Durchführung verschiedener Anwendungsfälle.

der PHP-Funktion *catria\_schema()* in der Datei *catria.install*. Mit der Installation von Catria erfolgt die Erstellung der Relationsschemas in einer zugewiesenen Datenbank-Instanz. Catria unterstützt dabei die Verwendung des DBMS PostgreSQL.

### 3.3.2. Deklaration der Datenobjekttypen

Catria stellt die Daten aus Fütterungsversuchen in einer Drupal-Umgebung mit acht Datenobjekttypen (engl. *Drupal entity type*) dar. Jeder Datenobjekttyp repräsentiert ein Relationsschema des relationalen Datenbankmodells (Abbildung 3.3). Die Deklaration der Datenobjekttypen erfolgt mit der PHP-Funktion *catria\_entity\_info()* in der Datei *catria.module*. Die Deklaration ordnet jeder Datenbankrelation eine Datenobjektklasse und eine Kontrollklasse zu.

Eine Datenobjektklasse stellt einzelne Einträge aus den Relationen als Datenobjekte dar. Die Datenobjektclassen definieren die Attribute der einzelnen Objekte, sie enthalten Funktionen zum Zugriff auf die Attribute der Objekte und sie bieten Funktionen zur grafischen Darstellung einzelner Objekte an. Die Attribute der Objekte einer Klasse entsprechen dabei den Attributen der zugehörigen Relation. Die Datenobjektclassen erben von der Klasse *Entity* aus dem Drupal Modul *Entity API*<sup>[128]</sup>. Jede Datenobjektklasse befindet sich als PHP-Klasse in einer eigenen PHP-Datei.

Die Kontrollklassen dienen zum Laden von Objekten aus der Datenbank und zum Speichern von Objekten in der Datenbank. Die Kontrollklassen enthalten Funktionen zur grafischen Darstellung von Mengen von Datenobjekten. Jede Kontrollklasse befindet sich als PHP-Klasse in einer eigenen PHP-Datei.

## 3.4. Verwendung von Catria

Die Installation von Catria auf einem Web-Server macht eine Catria-Website verfügbar. Je nach Konfiguration des Web-Servers kann die Catria-Website über eine spezifische Adresse (URL, engl. *Uniform resource locator*) mit einem Webbrowser von den Nutzern aufgerufen werden. Die Verwendung von Catria verläuft dann über diese Catria-Website (Abbildung 3.4).

### 3.4.1. Aufbau einer Catria-Website

Der Aufbau einer Catria-Website ist in die zwei Bereiche Menüleiste und Inhaltsbereich unterteilt (Abbildung 3.5). Die Menüleiste dient zur Navigation zwischen den Daten aus den Fütterungsversuchen (*Feedingtrials*), zur Verwaltung der eigenen Daten (engl. *My data*) und zur Verwaltung des eigenen Nutzerkontos (engl. *My account*). Im Inhaltsbereich stellt Catria die Messdaten und Metadaten dar. Jedes Datenobjekt ist über eine eigene Seite mit einer spezifischen URL aufrufbar. Die Darstellung eines Fütterungsversuchs zeigt beispielsweise die berücksichtigten Variablen, die enthaltenen Tiere und eine Übersicht über die durchgeführten Messungen an.

### 3.4.2. Datenintegration mit Catria

Zur Integration der Daten stellt Catria den Projektmitgliedern zwei Eingabemasken zur Verfügung. Catria unterscheidet zwischen der Integration von Metadaten zu einem Fütterungsversuch und der Integration von Messdaten.

#### 3.4.2.1. Speicherung eines Datenobjekts vom Typ *FÜTTERUNGSVERSUCH*

Die Eingabemaske „*Create new feedingtrial*“ dient zur Speicherung eines Datenobjekts vom Typ *FÜTTERUNGSVERSUCH*. Zuerst gibt der Nutzer eine Bezeichnung und eine Beschreibung des Fütterungsversuchs in die jeweiligen Felder ein. Catria erstellt anschließend ein Objekt des Typs *FÜTTERUNGSVERSUCH* und speichert die Bezeichnung und die Beschreibung des Fütterungsversuchs mit den Attributen *FÜTTERUNGSVERSUCH:bezeichnung* und *FÜTTERUNGSVERSUCH:beschreibung*. Bei der Verarbeitung der Eingabemaske vermerkt Catria auch den jeweiligen Nutzer als den Besitzer des Fütterungsversuchs mit dem Attribut *FÜTTERUNGSVERSUCH:NUTZER.id*.

Der Besitzer kann nach der Erstellung eines Fütterungsversuchs eine Datei mit Informationen über die betrachteten Variablen und eine Datei mit Informationen über die enthaltenen Tiere von einem lokalen Computer auf den Server hochladen. Catria validiert die beiden Dateien und speichert anschließend die Daten aus den Dateien in den zugehörigen Relationen der Datenbank-Instanz. Hierbei erstellt Catria für jede Variable ein Objekt vom Typ *VARIABLE* und speichert die Bezeichnung der Variable, die Beschreibung der Variable und einen Verweis auf den zugehörigen Fütterungsversuch mit den Attributen *VARIABLE:bezeichnung*, *VARIABLE:beschreibung* und *VARIABLE:FÜTTERUNGSVERSUCH.id*. Anschließend erstellt Catria für jedes Tier ein Objekt

The screenshot shows the Catria website interface. At the top, there are logos for 'PIGGELDAT' and 'CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN'. Below the logos, the interface is divided into two main sections: (a) Main menu and (b) P006 details.

**(a) Main menu:**

- Home
- Impressum

**(b) P006 details:**

**Feedingtrial: EFFT2010**

**Experimental variables:**

Variable	value
gender	male
agegroup	26 days
ageexact	26
startdate	2010-06-17
enddate	2010-07-13
treatmentgroup	probiotic
sow	5
boar	2

**Measurement: Zusammensetzung der intestinalen Mikrobiota**

Attribute	timepoint	value
Eubacteria_col	26 days	XXXX
Lactobacillus_col	26 days	XXXX
Cleptum_col	26 days	XXXX
Coccooides_col	26 days	XXXX
Enterococcus_col	26 days	XXXX
Streptococcus_col	26 days	XXXX
Enterobacter_col	26 days	XXXX
Bifidobacteria_col	26 days	XXXX
BacteroidesPrevotella_col	26 days	XXXX
Lacidonhilus_col	26 days	XXXX

**Abbildung 3.5.:** Aufbau einer Catria Website am Beispiel der Piggeldat-Datenbank. Der Aufbau ist in die zwei Bereiche (a) Menüleiste und (b) Inhaltsbereich unterteilt. Das Beispiel zeigt ein Tier mit der Bezeichnung *P006*.

vom Typ *TIER* und speichert eine Bezeichnung des Tiers, eine Beschreibung des Tiers und einen Verweis auf den zugehörigen Fütterungsversuch mit den Attributen *TIER:bezeichnung*, *TIER:beschreibung* und *TIER:FÜTTERUNGSVERSUCH.id*. Nach der Speicherung der Variablen und der Tiere erstellt Catria für jede Kombination von Variablen und Tieren ein Objekt vom Typ *BESCHREIBT* und speichert die Ausprägungen eines Tieres für eine Variable mit dem Attribut *BESCHREIBT:wert*.

### 3.4.2.2. Speicherung eines Datenobjekts vom Typ *MESSUNG*

Die Eingabemaske „*Create new measurement*“ dient zur Speicherung eines Datenobjekts vom Typ *MESSUNG*. Zuerst gibt der Nutzer eine Bezeichnung, eine Beschreibung und eine Liste mit relevanten Publikationen für die Messung in die jeweiligen Felder ein und ordnet die Messung einem Fütterungsversuch zu. Die Auflistung der relevanten Publikationen kann dabei in HTML (engl. *HyperText Markup Language*) erfolgen und Verweise auf die Websites der Journale enthalten. Catria erstellt anschließend ein Objekt vom Typ *MESSUNG* und speichert den Namen, die Beschreibung, die Liste mit relevanten Publikationen und einen Verweis auf den zugewiesenen Fütterungsversuch

mit den Attributen *MESSUNG:bezeichnung*, *MESSUNG:beschreibung*, *MESSUNG:publikationen* und *MESSUNG:FÜTTERUNGSVERSUCH.id*.

Anschließend kann der Nutzer eine Datei mit Informationen über die Merkmale und eine Datei mit den Messwerten auf den Server hochladen. Catria validiert beide Dateien und speichert anschließend die Daten in den zugehörigen Relationen der Datenbank-Instanz. Hierbei erstellt Catria für jedes Merkmal ein Objekt vom Typ *MERKMAL* und speichert die Bezeichnung des Merkmals, die Beschreibung des Merkmals, die Einheit des Merkmals und einen Verweis auf die zugehörige Messung mit den Attributen *MERKMAL:bezeichnung*, *MERKMAL:MESSUNG.id* und *MERKMAL:beschreibung*. Anschließend erstellt Catria für jeden Eintrag der Messwert-Datei ein Objekt vom Typ *BESTIMMT* und speichert den Messwert, den Messzeitpunkt, einen Verweis auf das zugehörige Tier und einen Verweis auf das zugehörige Merkmal mit den Attributen *BESTIMMT:messwert*, *BESTIMMT:zeitpunkt*, *BESTIMMT:TIER.id* und *BESTIMMT:MERKMAL.id*.

### 3.4.3. Datenverwaltung mit Catria

Nach der Integration von Messdaten kann ein Nutzer den Status eines Objekts vom Typ *MESSUNG* verändern. Catria speichert den Status einer Messung mit dem Attribut *MESSUNG:status*. Der Status einer Messung bestimmt die Zugriffsrechte der anderen Nutzer auf die Messwerte der Messung. Standardmäßig erfolgt die Speicherung des Status *closed* für eine neue Messung. Der Status *closed* verbietet den Zugriff für alle anderen Nutzer außer dem Besitzer der Daten. Der Status *restricted* erlaubt die Freigabe der Messdaten an ausgewählte Nutzer. Der Status *intern* gibt die Daten an alle registrierten Nutzer (Projektmitglieder) frei. Der Status *open* gibt freien Zugriff auf alle Messdaten der Messung an registrierte Nutzer und an nicht registrierte Nutzer (Gäste). Messwerte auf denen ein Nutzer kein Zugriff hat, werden mit der Zeichenfolge XXXX gekennzeichnet.



### 3.4.4. Datenexport mit Catria

Projektmitglieder und Gäste können mit dem Daten-Export-Werkzeug die Messdaten aus den Fütterungsversuchen herunterladen. Der Aufruf des Export-Werkzeugs erfolgt über die Maske *download*, welche bei der Übersicht eines einzelnen Fütterungsversuchs verfügbar ist. Bei der Verwendung des Werkzeuges erstellt Catria eine Datei im CSV-Format (engl. *comma-separated values*), welche die Daten eines Fütterungsversuchs enthält. Hierzu wählt der Nutzer mit der Eingabemaske zunächst Variablen und die Messungen aus, die in der Datei enthalten sein sollen. Die Export-Datei enthält dann die Ausprägungen zu den ausgewählten Variablen und die Messwerte zu den ausgewählten Messungen. Abbildung 3.6 zeigt den Aufbau einer Catria-Export-Datei.

Tier	Zeitpunkt	Variablen	Merkmale
1	$z_1$	$y_{1,1} \cdots y_{1,K}$	$x_{1,1} \cdots x_{1,K}$
$\vdots$	$\vdots$	$\vdots \quad \ddots \quad \vdots$	$\vdots \quad \ddots \quad \vdots$
$N$	$z_1$	$y_{N,1} \cdots y_{N,K}$	$x_{N,1} \cdots x_{N,K}$
$\vdots$	$\vdots$	$\vdots \quad \quad \quad \vdots$	$\vdots \quad \quad \quad \vdots$
1	$z_T$	$y_{1,1} \cdots y_{1,K}$	$x_{T,1,1} \cdots x_{T,1,K}$
$\vdots$	$\vdots$	$\vdots \quad \ddots \quad \vdots$	$\vdots \quad \ddots \quad \vdots$
$N$	$z_T$	$y_{N,1} \cdots y_{N,K}$	$x_{T,N,1} \cdots x_{T,N,K}$

**Abbildung 3.6.:** Aufbau der Catria-Export-Datei. Die Erstellung einer Catria-Export-Tabelle verläuft über die *download* Eingabemaske. Die Tabelle enthält zu jedem Tier (Nummer 1 bis  $N$ ) unterschiedliche Zeitpunkte ( $\mathbf{Z}$ ), die Ausprägungen der ausgewählten Variablen ( $\mathbf{Y}$ ) und die Messwerte der Messungen ( $\mathbf{X}$ ). Jede Messung umfasst dabei mehrere Merkmale.

**Tabelle 3.1.:** Domänen des relationalen Datenschemas von Catria

<b>Domäne</b>	<b>Wertebereich</b>	<b>Attribute</b>
<i>Dom_id</i>	Menge aller natürlichen Zahlen von 1 bis 2 <sup>31</sup>	<i>FÜTTERUNGSVERSUCH:id</i> <i>FÜTTERUNGSVERSUCH:NUTZER.id</i> <i>VARIABLE:id</i> <i>VARIABLE:FÜTTERUNGSVERSUCH.id</i> <i>NUTZER:id</i> <i>TIER:id</i> <i>TIER:FÜTTERUNGSVERSUCH.id</i> <i>MESSUNG:id</i> <i>MESSUNG:NUTZER.id</i> <i>MERKMAL:id</i> <i>MERKMAL:MESSUNG.id</i> <i>BESCHREIBT:VARIABLE.id</i> <i>BESCHREIBT:TIER.id</i> <i>HATZUGRIFF:NUTZER.id</i> <i>HATZUGRIFF:MESSUNG.id</i> <i>BESTIMMT:TIER.id</i> <i>BESTIMMT:MERKMAL.id</i>
<i>Dom_bez</i>	Menge aller Zeichenketten mit einer maximalen Länge von 128 Zeichen	<i>FÜTTERUNGSVERSUCH:bezeichnung</i> <i>VARIABLE:bezeichnung</i> <i>NUTZER:name</i> <i>TIER:bezeichnung</i> <i>MESSUNG:bezeichnung</i> <i>MESSUNG:status</i> <i>MERKMAL:bezeichnung</i> <i>MERKMAL:gruppe</i> <i>MERKMAL:einheit</i> <i>BESCHREIBT:wert</i> <i>HATZUGRIFF:rechte</i> <i>BESTIMMT:zeitpunkt</i>
<i>Dom_bes</i>	Menge aller Zeichenketten mit einer maximalen Länge von 3000 Zeichen	<i>FÜTTERUNGSVERSUCH:beschreibung</i> <i>VARIABLE:beschreibung</i> <i>TIER:beschreibung</i> <i>MESSUNG:beschreibung</i> <i>MESSUNG:publikationen</i> <i>MERKMAL:beschreibung</i>
<i>Dom_wert</i>	Menge aller Fließkommazahlen mit maximal 24 Stellen und 8 Nachkommastellen	<i>BESTIMMT:messwert</i>
<i>Dom_email</i>	Menge aller möglichen E-Mail Adressen	<i>NUTZER:email</i>

## 4. Die Piggeldat-Datenbank

Die Piggeldat-Datenbank dient dem Sonderforschungsbereich 852 zur Verwaltung der Messdaten und Metadaten aus sieben Fütterungsversuchen, die in den Jahren 2010 bis 2013 durchgeführt wurden (Tabelle 4.1). Das Ziel dieser Versuche war die Untersuchung der Auswirkungen der Fütterung des Probiotikums *Enterococcus faecium* NCIMB 10415 und die Auswirkungen der Fütterung von unterschiedlichen Zink-Konzentrationen auf die Gesundheit von Ferkeln. In *Challenge*-Versuchen wurde insbesondere das Zusammenspiel der beiden Nahrungsfaktoren mit der Infektion mit dem pathogenen Bakterium *Salmonella enterica* und mit Infektionen mit den Viren *Transmissible Gastroenteritis Coronavirus* (TGEV), *porcine reproductive and respiratory syndrome virus* (PRRSV) und Schweineinfluenza (SwIV, engl. *swine influenza virus*) untersucht. Hierfür führten insgesamt 15 Projektgruppen aus sechs Instituten (Bundesinstitut für Risikobewertung, Charité - Universitätsmedizin Berlin, Deutsches Institut für Ernährungsforschung, Freie Universität Berlin, Humboldt-Universität zu Berlin und Technische Universität Berlin) die sieben Fütterungsversuche durch und maßen diverse Messdaten mit Technologien aus unterschiedlichen biologischen Bereichen.

**Tabelle 4.1.:** Fütterungsversuche des SFB 852 in den Jahren 2010 bis 2013

Bezeichnung	Untersuchung
EFFT2010	Generelle Auswirkung der Fütterung des Probiotikums <i>Enterococcus faecium</i> NCIMB
EFST2010	Auswirkung der Fütterung des Probiotikums <i>Enterococcus faecium</i> NCIMB 10415 auf eine Infektion mit <i>Salmonella enterica</i> Serovar Typhimurium DT104
ZNFT2011	Generelle Auswirkung der Fütterung unterschiedlicher Zink-Konzentrationen
ZNST2011	Auswirkung der Fütterung unterschiedlicher Zink-Konzentrationen auf eine Infektion mit <i>Salmonella enterica</i> Serovar Typhimurium DT104
ZNTG2012	Auswirkung der Fütterung unterschiedlicher Zink-Konzentrationen auf eine Infektion mit TGEV
ZNPR2012	Auswirkung der Fütterung unterschiedlicher Zink-Konzentrationen auf eine Infektion mit PRRSV
EZIV2013	Auswirkung der Fütterung unterschiedlicher Zink-Konzentrationen und des Probiotikums <i>Enterococcus faecium</i> NCIMB 10415 auf eine Infektion mit SwIV

## 4.1. Installation der Datenbankanwendung

Die Piggeldat-Datenbank wurde auf einem virtuellen Server der Charité-Universitätsmedizin installiert. Auf dem Server lief das Betriebssystem Ubuntu in Version 14.04 LTS (engl. *long-time support*). Des Weiteren wurden die von Drupal benötigten Programme PHP5 (Version 5.5.9), Apache Webserver (Version 2.4.7) und PostgreSQL (Version 9.1.2) auf dem Server installiert und konfiguriert. Die Installation der Datenbankanwendung Catria 1.0 erfolgte durch den Transfer der Catria-PHP-Dateien in das *document root* Verzeichnis des Apache-Webservers (*/var/www/*). Mit der Ausführung des Installationsskripts über einen Webbrowser wurde eine PostgreSQL-Datenbank-Instanz zugeteilt und die Website mit Zugriff auf die Datenbank eingerichtet. Die Piggeldat-Website ist über das Internet aufrufbar<sup>[110]</sup>.

## 4.2. Integration der Daten

Die Integration der Forschungsdaten des SFB 852 in die Piggeldat-Datenbank erfolgte über die Piggeldat-Website. Vor der Integration erfolgte zunächst die Einsammlung aller Forschungsdaten in Form von 50 Excel Dateien von den Projektgruppen und anschließend die einheitliche Formatierung aller Forschungsdaten. Eine Schwierigkeit hierbei lag bei der uneinheitlichen Formatierung der Dateien durch die Projektgruppen. Hierdurch war die einzelne Bearbeitung aller Dateien erforderlich, wodurch auch die einheitliche Verwendung von Bezeichnungen der experimentellen Variablen, der Ferkel und der gemessenen Merkmale sichergestellt wurde. Die Integration der Dateien erfolgte dann anschließend mit den Eingabemasken der Catria-Datenbankanwendung (siehe Abschnitt 3.4.2).

### 4.2.1. Integration der Fütterungsversuche

Für jeden Fütterungsversuch wurde eine Tabelle mit Informationen über die enthaltenen Ferkel und eine Tabelle mit Informationen über die berücksichtigten Variablen im CSV-Format erstellt. Hierzu wurden einheitliche Bezeichnungen aller Tiere und aller Variablen für die sieben Fütterungsversuche festgelegt und sichergestellt. Insgesamt zehn unterschiedliche Variablen wurden in den 7 Fütterungsversuchen berücksichtigt (Tabelle 4.2). Die Bezeichnungen der Ferkel setzten sich aus dem Buchstaben *P* und einer fortlaufenden dreistelligen Nummer zusammen (z.B. *P004*). Für jede Variable eines Fütterungsversuchs enthielt die CSV-Tabelle der zugehörigen Ferkel eine Spalte, die die Ausprägung der Ferkel für die jeweilige Variable enthielt.

### 4.2.2. Integration der Messdaten

Zur Integration der Messdaten aus den Excel-Dateien erfolgte zuerst eine Einordnung der Messdaten in Messungen. Für jede Messung wurde dann eine Datei mit den Messwerten und eine Datei mit Informationen zu den gemessenen Merkmalen erstellt. Hierbei wurde eine einheitliche Formatierung der Bezeichnungen der Merkmale eingeführt und sichergestellt. Die Bezeichnung eines Merkmals setzte sich aus einer kurzen Bezeichnung für die Art des gemessenen Merkmals und aus einer kurzen Bezeichnung für den Probenentnahmeort zusammen. Die Merkmale der sieben Fütterungsversuche entstammten aus 29 unterschiedlichen Probenentnahmeorten (Tabelle 4.3). Die einheitlichen Bezeichnungen der Messungen und der Merkmale wurde dann in allen

**Tabelle 4.2.:** Experimentelle Variablen und Störvariablen der Piggeldat-Datenbank

Variable	Bedeutung
<i>ageexact</i>	Exaktes Alter der Ferkel in Lebenstagen
<i>agegroup</i>	Altersgruppe
<i>boar</i>	ID des Ebers
<i>enddate</i>	Datum der Probenentnahme
<i>gender</i>	Geschlecht der Ferkel ( <i>male</i> oder <i>female</i> )
<i>litter</i>	Nummer des Wurfs
<i>sow</i>	ID der Sau
<i>startdate</i>	Geburtsdatum der Ferkel
<i>treatmentgroup</i>	Behandlungsgruppe
<i>vacgroup</i>	Erhalt einer Impfung ( <i>yes</i> oder <i>no</i> )

Dateien umgesetzt. Die Beschreibung der Merkmale wurde den zugehörigen Publikationen entnommen und in persönlichen Gesprächen mit den Experimentatoren abgeglichen. Die Verwendung der Piggeldat-Website erlaubte dann die Integration der Messdaten und der Metadaten in der Datenbank.

## 4.3. Inhalt der Piggeldat-Datenbank

Die Piggeldat-Datenbank enthält Messdaten und Metadaten aus sieben Fütterungsversuchen (Tabelle 4.4). Dabei beinhaltet die Datenbank die Metadaten von insgesamt 490 Ferkeln, 57 Messungen und 1244 Merkmalen und sie beinhaltet insgesamt 81060 Messwerte<sup>[110]</sup>. Der Fütterungsversuch *ZNFT2011* repräsentiert mit 15 Messungen und 39833 Messwerten den umfangreichsten Datensatz der Piggeldat-Datenbank, während der Fütterungsversuch *ZNTG2012* mit 4 Messungen und 876 Messwerten den kleinsten Datensatz repräsentiert. Auf der Website sind 27 Nutzer registriert. Alle Messdaten sind für die registrierten Mitglieder des SFB 852 frei zugänglich.

### 4.3.1. Der Fütterungsversuch *EFFT2010*

Der SFB 852 führte den Fütterungsversuch *EFFT2010* am Institut für Tierernährung der Freien Universität Berlin durch (Bednorz et al. 2013<sup>[8]</sup>, Klingspor et al. 2013<sup>[56]</sup>, Kreuzer et al. 2012<sup>[60]</sup>, Martin et al. 2012<sup>[67]</sup>, Richter et al. 2011<sup>[83]</sup>, Scharek-Tedin et al. 2015<sup>[90]</sup>, Siepert et al. 2014<sup>[95]</sup> und Starke et al. 2013<sup>[101]</sup>). Abbildung 4.1 zeigt den Aufbau des Fütterungsversuchs. Insgesamt 24 trächtige Sauen wurden zufällig in eine Kontrollgruppe und eine Behandlungsgruppe aufgeteilt. Die Sauen in der Behandlungsgruppe erhielten vier Wochen vor der erwarteten Geburt eine Ernährung mit zusätzlichem *Enterococcus faecium* NCIMB 10415. Die Ferkel in der Behandlungsgruppe hatten ab dem 12. Lebenstag Zugang zu einer Ernährung mit zusätzlichem *E. faecium* NCIMB 10415. Die Absetzung der Ferkel von den Sauen erfolgte nach 26 Lebenstagen. Alle Ferkel wurden in vier Altersgruppen aufgeteilt. Nach 12, 26, 34 und 54 Lebenstagen wurden die Ferkel aus den entsprechenden Altersgruppen zur Probenentnahme geschlachtet.

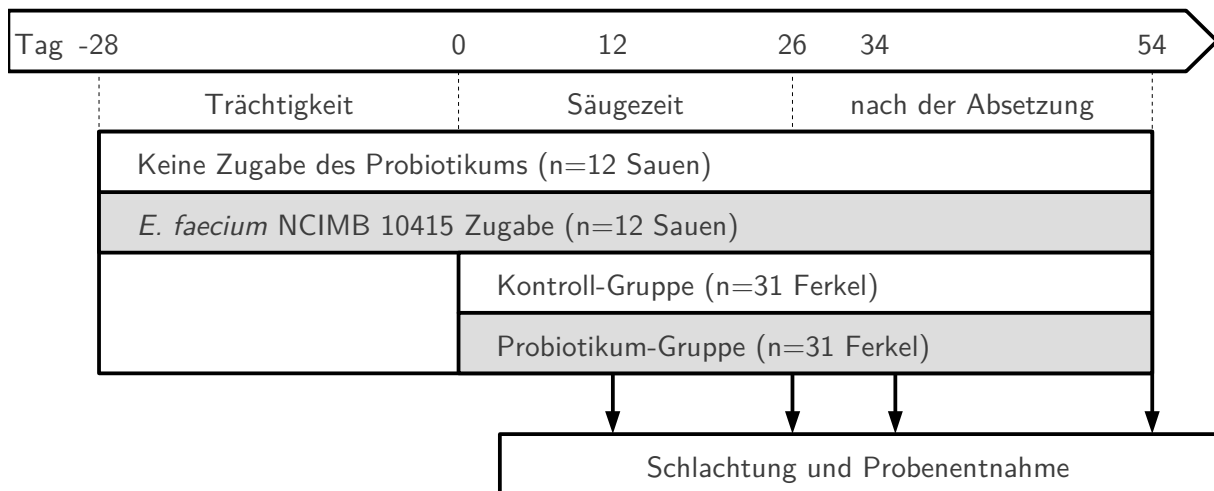
**Tabelle 4.3.:** Probenentnahmeorte der Messwerte in der Piggeldat-Datenbank

<b>Entnahmeort</b>	<b>Bedeutung</b>
<i>bdy</i>	Gesamter Körper eines Ferkels (engl. <i>body</i> )
<i>blo</i>	Blutprobe eines Ferkel(engl. <i>blood</i> )
<i>blo_pbmc</i>	Mononukleäre Zellen des peripheren Blutes
<i>bone</i>	Knochenprobe
<i>buc</i>	Abstrich des Mauls (engl. <i>buccal</i> )
<i>col</i>	Gewebeprobe des Colons (engl. <i>colon</i> )
<i>col_asc</i>	Aufsteigendes Colon (lat. <i>colon ascendens</i> )
<i>col_des</i>	Absteigendes Colon (lat. <i>colon descendens</i> )
<i>col_dig</i>	Inhalt des Colons (engl. <i>colon digesta</i> )
<i>col_muc</i>	Schleimhaut des Colons (engl. <i>colon mucosa</i> )
<i>fae</i>	Probe der Fäkalien (engl. <i>faeces</i> )
<i>ile</i>	Gewebeprobe des Ileums (engl. <i>ileum</i> )
<i>ile_LN</i>	Lymphknoten des Ileums (engl. <i>faeces</i> )
<i>ile_PP</i>	Peyer-Plaques des Ileums (engl. <i>faeces</i> )
<i>jej</i>	Gewebeprobe des Jejunums (engl. <i>faeces</i> )
<i>jej_prox</i>	Gewebeprobe des proximalen Jejunums
<i>jej_dist</i>	Gewebeprobe des distalen Jejunums
<i>jej_BBM</i>	Probe der Bürstensaummembran des Jejunums (engl. <i>jejunum brush border membrane</i> )
<i>jej_LN</i>	Lymphknoten des Jejunums (engl. <i>faeces</i> )
<i>jej_PP</i>	Peyer-Plaques des Jejunums (engl. <i>jejunum peyers patches</i> )
<i>kny</i>	Gewebeprobe der Niere (engl. <i>kidney</i> )
<i>liv</i>	Gewebeprobe der Leber (engl. <i>liver</i> )
<i>lun</i>	Gewebeprobe der Lunge
<i>mln</i>	<i>Mesenteric lymph nodes</i>
<i>nas</i>	Abstrich aus der Nase
<i>pacr</i>	Gewebeprobe der Bauchspeicheldrüse (engl. <i>pancreas</i> )
<i>spln</i>	Gewebeprobe der Milz (engl. <i>spleen</i> )
<i>sto</i>	Gewebeprobe des Magens(engl. <i>stomach</i> )
<i>tons</i>	Gewebeprobe der Mandeln (engl. <i>tonsils</i> )

**Tabelle 4.4.:** Anzahl der Datenobjekte in der Piggeldat Datenbank

<b>Fütterungsversuch</b>	<b>#Ferkel</b>	<b>#Messungen</b>	<b>#Merkmale</b>	<b>#Messwerte</b>
<i>EFFT2010</i>	62	14	492	18155
<i>EFST2010</i>	32	6	45	2200
<i>ZNFT2011</i>	120	15	596	39833
<i>ZNST2011</i>	72	3	25	1921
<i>ZNTG2012</i>	60	4	7	876
<i>ZNPR2012</i>	72	7	29	10643
<i>EZIV2013</i>	72	8	50	7432

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägung der



**Abbildung 4.1.:** Aufbau des Fütterungsversuchs EFFT2010 mit 24 Sauen und 62 Ferkeln. Tag 0 gibt den Zeitpunkt der Geburt der Ferkel an. Die nachfolgenden vier Tage geben die Zeitpunkte der Probenentnahmen an. Ab dem 12. Tag hatten alle Ferkel Zugang zu einer *starter*-Fütterung mit (Probiotikum-Gruppe) oder ohne (Kontroll-Gruppe) Zusatz des Probiotikums *Enterococcus faecium* NCIMB 10415. Die Absetzung der Ferkeln von den Sauen erfolgt nach 26 Tagen.

acht Variablen *agegroup*, *ageexact*, *startdate*, *enddate*, *treatmentgroup*, *gender*, *sow* und *boar*. Die Variable *treatmentgroup* hat die möglichen Ausprägungen *control* (Kontroll-Gruppe) und *probiotic* (Probiotikum-Gruppe). Die Variable *agegroup* hat die möglichen Ausprägungen *12 days*, *26 days*, *34 days* und *54 days* zur Notierung des Zeitpunktes der Probenentnahme.

Die integrierten Messungen beinhalten: Zusammensetzung der gastrointestinalen Mikrobiota, Konzentration bakterieller Metabolite, Vorkommen virulenter *Escherichia coli* Gene, Transportfunktion der Darmschleimhaut, Barrierenfunktion der Darmschleimhaut, Morphologie der Darmschleimhaut, Aktivität der Verdauungsenzyme, Genexpression des Darmepithels, Konzentration der Zytokine im Blut, Expression immunrelevanter Gene, Zusammensetzung der Immunzellpopulationen, Großes Blutbild, Gewicht der Ferkel und Größe einzelner Organe (Tabelle D.1). Die Piggeldat-Datenbank enthält zu den Messungen insgesamt 18155 Messwerte (Tabelle 4.4).

### 4.3.2. Der Fütterungsversuch EFST2010

Der SFB 852 führte den Fütterungsversuch *EFST2010* am Institut für Tierernährung der Freien Universität Berlin und am Bundesinstitut für Risikobewertung durch (Kreuzer et al. 2012<sup>[59]</sup> und Siepert et al. 2014<sup>[95]</sup>). Die Aufzucht der insgesamt 32 Ferkel erfolgte beim Institut für Tierernährung der Freien Universität Berlin. Zuerst wurden trächtige Sauen zufällig in eine Kontrollgruppe und eine Behandlungsgruppe aufgeteilt. Die Sauen in der Behandlungsgruppe erhielten drei Wochen vor der erwarteten Geburt eine Ernährung mit zusätzlichem *Enterococcus faecium* NCIMB 10415. Die Ferkel in der Behandlungsgruppe hatten ab dem 12. Lebenstag Zugang zu einer Ernährung mit zusätzlichem *E. faecium* NCIMB 10415. Nach der Absetzung mit

28 Lebenstagen erfolgte der Transport aller Ferkel zum Bundesinstitut für Risikobewertung. Die Tiere innerhalb der gleichen Behandlungsgruppen wurden paarweise in Gehege aufgeteilt. Mit 38 Lebenstagen erfolgte eine Infektion aller Tiere mit *Salmonella enterica* Serovar Typhimurium DT104. Sechs Tiere pro Behandlungsgruppe wurden zwei Tage nach der Infektion zur Probenentnahme geschlachtet. Die restlichen 10 Tiere pro Behandlungsgruppe wurden 28 Tage nach der Infektion zur Probenentnahme geschlachtet.

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägung der vier Variablen *agegroup*, *treatmentgroup*, *sow* und *gender*. Die Variable *treatmentgroup* hat die möglichen Ausprägungen *control* (Kontroll-Gruppe) und *probiotic* (Probiotikum-Gruppe). Die Variable *agegroup* hat die möglichen Ausprägungen *40 days* und *56 days* zur Notierung des Zeitpunktes der Probenentnahme.

Die integrierten Messungen beinhalten: Expression immunrelevanter Gene, Körpertemperatur, Vorkommen von Salmonellen, Gewicht der Ferkel, Konsistenz der Fäkalien und Vorkommen von Anti-Salmonellen Antikörpern (Tabelle D.1). Die Piggeldat-Datenbank enthält zu der Messung insgesamt 2200 Messwerte (Tabelle 4.4).

### 4.3.3. Der Fütterungsversuch ZNFT2011

Der SFB 852 führte den Fütterungsversuch ZNFT2011 am Institut für Tierernährung der Freien Universität Berlin mit insgesamt 120 Ferkeln durch (Bednorz et al. 2013<sup>[9]</sup>, Gefeller et al. 2014<sup>[35]</sup>, Liu et al. 2014<sup>[65]</sup>, Martin et al. 2013<sup>[68]</sup>, Starke et al. 2013<sup>[102]</sup>). Die Absetzung der Ferkel von den Sauen erfolgte 26 Tage nach der Geburt. Anschließend erfolgte die zufällige Aufteilung der Ferkel in drei gleich große Behandlungsgruppen, wobei nach Geschlecht, Wurf und Gewicht balanciert wurde. Die Tiere innerhalb der gleichen Behandlungsgruppen wurden paarweise in Gehege aufgeteilt. Ab dem 12. Lebenstag hatten alle Tiere Zugriff auf das gleiche Futter (engl. *starter-diet*). Nach der Absetzung erhielten die Ferkel je nach Behandlungsgruppe eine Diät mit entweder wenig (50 mg Zn/kg Futter), mittel (150 mg Zn/kg Futter) oder viel (2500 mg Zn/kg Futter) zusätzlichen Zink-Ionen. Alle Ferkel wurden in vier Altersgruppen aufgeteilt. Nach 32, 39, 46 und 53 Lebenstagen wurden die Ferkel aus den entsprechenden Altersgruppen zur Probenentnahme geschlachtet.

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägungen der sieben Variablen *age\_group*, *age\_exact*, *treatment\_group*, *gender*, *sow*, *boar* und *litter*. Die Variable *treatment\_group* hat die möglichen Ausprägungen *low* (wenig zusätzliches Zink), *medium* (Kontroll-Gruppe) und *high* (viel zusätzliches Zink). Die Variable *agegroup* hat die möglichen Ausprägungen *32 days*, *39 days*, *46 days* und *53 days* zur Notierung des Zeitpunktes der Probenentnahme.

Die integrierten Messungen beinhalten: Zusammensetzung der intestinalen Mikrobiota, Konzentration bakterieller Metabolite, Konzentration von Zink im Verdauungstrakt, Vorkommen virulenter *Escherichia coli* Gene, Transport-Funktion der Darmschleimhaut, Morphologie der Darmschleimhaut, Aktivität der Verdauungsenzyme, Genexpression des Darmepithels, Expression immunrelevanter Gene, Zusammensetzung der Immunzellpopulationen, Großes Blutbild, Vorkommen von Metallen im Körper, das Gewicht der Ferkel und die Größe einzelner Organe (Tabelle D.1). Die Piggeldat-Datenbank enthält zu den Messungen insgesamt 39833 Messwerte (Tabelle 4.4).



#### 4.3.4. Der Fütterungsversuch ZNST2011

Der SFB 852 führte den Fütterungsversuch ZNST2011 am Institut für Tierernährung der Freien Universität Berlin und am Bundesinstitut für Risikobewertung durch (Janczyk et al. 2013 [51]). Die Aufzucht der insgesamt 72 Ferkel erfolgte beim Institut für Tierernährung der Freien Universität Berlin. Nach der Absetzung mit 28 Tagen erfolgte der Transport der Ferkel zum Bundesinstitut für Risikobewertung und dort erfolgte eine zufällige Aufteilung der Tiere in drei Behandlungsgruppen. Die Tiere innerhalb der gleichen Behandlungsgruppen wurden in sechs Gehege aufgeteilt. Die Ferkel erhielten je nach Behandlungsgruppe ein Futter mit entweder wenig (50 mg Zn/kg Futter), mittel (150 mg Zn/kg Futter) oder viel (2500 mg Zn/kg Futter) zusätzlichen Zink-Ionen. Mit 32 Lebenstagen erfolgte eine Infektion aller Tiere mit *Salmonella enterica* Serovar Typhimurium DT104. Die Hälfte der Tiere wurden zwei Tage nach der Infektion zur Probenentnahme geschlachtet. Die andere Hälfte der Tiere wurde 42 Tage nach der Infektion zur Probenentnahme geschlachtet.

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägungen der zwei Variablen *agegroup* und *treatmentgroup*. Die Variable *treatmentgroup* hat die möglichen Ausprägungen *low* (wenig zusätzliches Zink), *medium* (Kontroll-Gruppe) und *high* (viel zusätzliches Zink). Die Variable *agegroup* hat die möglichen Ausprägungen *34 days* und *74 days* zur Notierung des Zeitpunktes der Probenentnahme.

Die integrierten Messungen beinhalten: Expression immunrelevanter Gene, Vorkommen von Salmonellen und Gewicht der Ferkel (Tabelle D.1). Die Piggeldat Datenbank enthält zu den Messung insgesamt 1921 Messwerte (Tabelle 4.4).

#### 4.3.5. Der Fütterungsversuch ZNTG2012

Der SFB 852 führte den Fütterungsversuch ZNTG2012 am Institut für Tierernährung der Freien Universität Berlin und am Bundesinstitut für Risikobewertung durch (Chai et al. 2014 [18]). Die Aufzucht der insgesamt 60 Ferkel erfolgte beim Institut für Tierernährung der Freien Universität Berlin. Nach der Absetzung mit 21 Tagen erfolgte die Aufteilung der Tiere in drei Behandlungsgruppen. Nach der Absetzung erhielten die Ferkel je nach Behandlungsgruppe ein Futter mit entweder wenig (50 mg Zn/kg Futter), mittel (150 mg Zn/kg Futter) oder viel (2500 mg Zn/kg Futter) zusätzlichen Zink-Ionen. Mit 26 Tagen erfolgte der Transport der Ferkel zum Bundesinstitut für Risikobewertung und dort erfolgte eine zufällige Aufteilung der Ferkel in 6 Buchten pro Behandlungsgruppe. Nach 28 Tagen wurden alle Ferkel mit TGEV infiziert. Einen Tag nach der Infektion wurden 10 Ferkel pro Behandlungsgruppe zur Probenentnahme geschlachtet und die restlichen 40 Ferkel wurden 18 Tagen nach der Infektion zur Probenentnahme geschlachtet.

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägungen der zwei Variablen *age\_group* und *treatment\_group*. Die Variable *treatment\_group* hat die möglichen Ausprägungen *low* (wenig zusätzliches Zink), *medium* (Kontroll-Gruppe) und *high* (viel zusätzliches Zink). Die Variable *agegroup* hat die möglichen Ausprägungen *29 days* und *44 days* zur Notierung des Zeitpunktes der Probenentnahme.

Die integrierten Messungen beinhalten: Genexpression von Zytokinen im Darmepithel, Vorkommen von Anti-TGEV Antikörpern, Konsistenz der Fäkalien und Gewicht der Ferkel (Tabelle D.1). Die Piggeldat Datenbank enthält zu den Messungen insgesamt 876 Messwerte (Tabelle 4.4).

#### 4.3.6. Der Fütterungsversuch ZNPR2012

Der SFB 852 führte den Fütterungsversuch ZNPR2012 am Institut für Tierernährung der Freien Universität Berlin und am Bundesinstitut für Risikobewertung durch (Chai et al. 2014<sup>[17]</sup>). Die Aufzucht der insgesamt 72 Ferkel erfolgte beim Institut für Tierernährung der Freien Universität Berlin. Nach der Absetzung mit 28 Tagen erfolgte der Transport der Ferkel zum Bundesinstitut für Risikobewertung und dort erfolgte eine zufällige Aufteilung der Tiere in drei Behandlungsgruppen. Die Ferkel erhielten je nach Behandlungsgruppe ein Futter mit entweder wenig (50 mg Zn/kg Futter), mittel (150 mg Zn/kg Futter) oder viel (2500 mg Zn/kg Futter) zusätzlichen Zink-Ionen. Ferkel innerhalb einer Behandlungsgruppe wurden in zwei Buchten pro Gruppe aufgeteilt. Nach 35 Tagen erhielten die Hälfte der Ferkel (eine Bucht pro Behandlungsgruppe) eine Impfung gegen PRRSV. Nach 63 Tagen wurden alle Tiere mit PRRSV infiziert und nach 98 Tagen wurden alle Ferkel zur Probenentnahme geschlachtet.

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägungen der zwei Variablen *treatment\_group* und *vac\_group*. Die Variable *treatment\_group* hat die möglichen Ausprägungen *low* (wenig zusätzliches Zink), *medium* (Kontroll-Gruppe) und *high* (viel zusätzliches Zink). Die Variable *vac\_group* hat die möglichen Ausprägungen *yes* (mit Impfung) und *no* (ohne Impfung).

Die integrierten Messungen beinhalten: Gewicht der Ferkel, Körpertemperatur der Ferkel, Zusammensetzung von Lymphozyten im Blut, das Vorkommen des Virus, Humorale Immunantwort, Vorkommen von IFN- $\gamma$  sekretierende Zellen und Konzentration von Anti-PRRSV Antikörpern (Tabelle D.1). Die Piggeldat Datenbank enthält zu den Messungen insgesamt 10643 Messwerte (Tabelle 4.4).

#### 4.3.7. Der Fütterungsversuch EZIV2013

Der SFB 852 führte den Fütterungsversuch EZIV2013 am Institut für Tierernährung der Freien Universität Berlin und am Friedrich Löffler Institut auf der Insel Riems durch (Wang et al. <sup>[115]</sup>). Die Aufzucht der insgesamt 72 Ferkel erfolgte beim Institut für Tierernährung der Freien Universität Berlin. Nach der Absetzung mit 28 Tagen erfolgte die Aufteilung der Tiere in drei Behandlungsgruppen. Nach der Absetzung erhielten die Ferkel je nach Behandlungsgruppe ein Futter mit entweder wenig (50 mg Zn/kg Futter) oder viel (2500 mg Zn/kg Futter) zusätzlichen Zink-Ionen oder ein Futter mit zusätzlichem *E. faecium* NCIMB 10415. Tiere innerhalb der Behandlungsgruppen wurden in zwei Buchten pro Gruppe aufgeteilt. Eine Bucht pro Behandlungsgruppe wurde zufällig ausgewählt und die zugehörigen Ferkel haben eine Impfung gegen Schweineinfluenza erhalten. Mit 58 Lebenstagen erfolgte der Transport der Ferkel zum Friedrich Löffler Institut auf der Insel Riems, wo anschließend nach 63 Tagen alle Ferkel mit Schweineinfluenza Viren H3N2 infiziert wurden. Die Ferkel wurden dann entweder einen Tag oder 6 Tage nach der Infektion zur Probenentnahme geschlachtet.

Die Piggeldat-Datenbank enthält für jedes Ferkel des Fütterungsversuchs die Ausprägungen der zwei Variablen *treatment\_group* und *vac\_group*. Die Variable *treatment\_group* hat die möglichen Ausprägungen *probiotic* (Probiotikum-Gruppe), *low* (wenig zusätzliches Zink) und *high* (viel zusätzliches Zink). Die Variable *vac\_group* hat die möglichen Ausprägungen *yes* (mit Impfung) und *no* (ohne Impfung).

### 4.3. Inhalt der Piggeldat-Datenbank

Die integrierten Messungen beinhalten: Körpertemperatur der Ferkel, Zusammensetzung von Lymphozyten, Vorkommen des Virus, Vorkommen von Anti-Virus Antikörpern und Morphologie der Bronchien (Tabelle D.1). Die Piggeldat Datenbank enthält zu den Messungen insgesamt 7432 Messwerte (Tabelle 4.4).



# 5. Analyse der Auswirkung des Probiotikums *E. faecium* NCIMB 10415

Die integrierten Messdaten des Fütterungsversuchs *EFFT2010* (Abschnitt 4.3.1) dienen zur Analyse der Auswirkungen des Probiotikums *E. faecium* NCIMB 10415 auf das Zusammenspiel zwischen gastrointestinaler Mikrobiota und Wirt beim Schwein. Die Ergebnisse dieser Analyse wurden im Jahr 2014 publiziert (Twardziok et al. 2014<sup>[109]</sup>). Abbildung 5.1 stellt den Ablauf der durchgeführten Berechnungen dar. Zuerst erfolgte die Datenverarbeitung mit der Aufteilung der Daten in fünf Datenmatrizen (Abbildung 5.1 **(a)**). Anschließend erfolgte eine Filterung der Datenmatrizen und die Standardisierung der Datenmatrizen (Abbildung 5.1 **(b)**). Die Durchführung einer Hauptkomponentenanalyse (PCA, engl. *Principle Component Analysis*) diente zur Untersuchung der Varianz in der vollständigen Datenmatrix (Abbildung 5.1 **(c)**). Das Verfahren *Orthogonal Partial Least Square* (OPLS) diente zur Analyse der Auswirkung der Fütterung des Probiotikums anhand der fünf Datenmatrizen (Abbildung 5.1 **(d)**). Die Ergebnisse der OPLS-Analyse dienten zum Vergleich der Auswirkungen der Probiotikum-Fütterung zwischen den verschiedenen Zeitpunkten (Abbildung 5.1 **(e)**) und zur Identifizierung von Merkmalsgruppen, welche besonders stark durch die Zugabe des Probiotikums beeinflusst wurden (Abbildung 5.1 **(f)**). Im letzten Schritt erfolgte die Analyse eines Korrelationsnetzwerks (Abbildung 5.1 **(g)**). Sowohl die Datenverarbeitung als auch die Datenanalyse erfolgten mit eigenen Skripten in der Programmiersprache R (Version 3.1.1).

## 5.1. Beschreibung der integrierten Datensätze

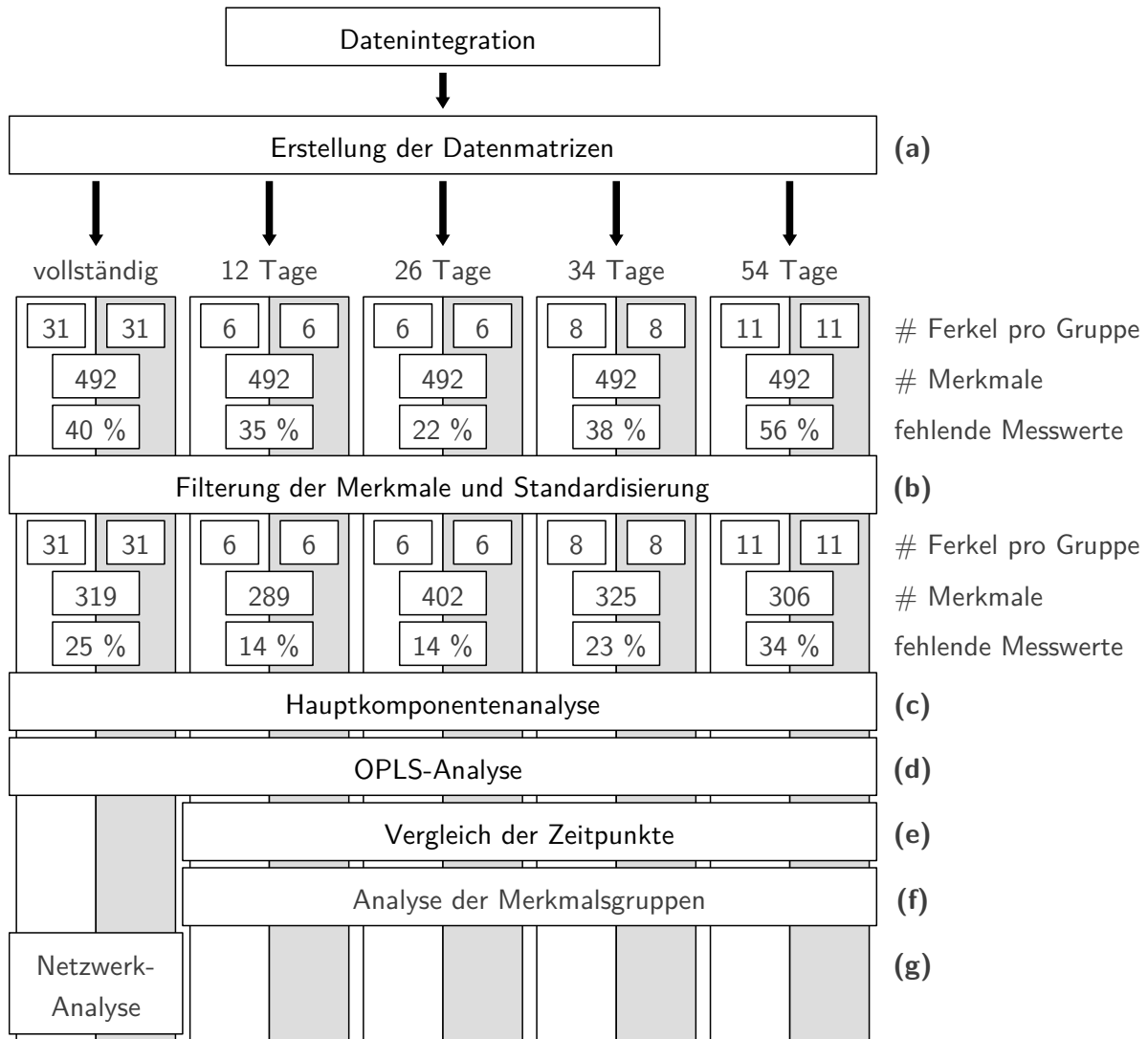
Die Sammlung der Messdaten erfolgte über die Piggeldat-Datenbank (siehe Kapitel 4). Der gesamte Datensatz setzte sich aus den folgenden 13 Datensätzen zusammen (Tabellen E.1 bis E.13):

- **Zusammensetzung der gastrointestinalen Mikrobiota**

Das Vorkommen spezifischer bakterieller OTUs wurde von Mitarbeitern des Instituts für Tierernährung der Freien Universität Berlin gemessen (Starke et al. 2013<sup>[101]</sup>). Mit der Anwendung einer *real-time-PCR* (engl. *polymerase chain reaction*) wurde das Vorkommen von 17 OTUs an 4 Entnahmeorten (Magen, Jejunum, Ileum und Colon) gemessen. Der Datensatz enthielt die Messwerte von 68 Merkmalen, welche bei 52 Ferkeln bestimmt wurden (Tabelle E.1).

- **Konzentration bakterieller Metabolite**

Die Konzentration bakterieller Metabolite wurde durch Mitarbeiter des Instituts für Tierernährung der Freien Universität Berlin gemessen. Mit verschiedenen Verfahren wurde die Konzentration von 16 Metaboliten (SCFAs, Laktate, Amine und Ammonium) an vier



**Abbildung 5.1.:** Ablauf der Datenverarbeitung und der Analyse der Messdaten aus dem *EFFT2010*-Fütterungsversuch. **(a)** Aufteilung der Messdaten in Datenmatrizen wurde nach dem Alter der Ferkel zum Messzeitpunkt durchgeführt. **(b)** Verarbeitung der Datenmatrizen entfernte Merkmale mit übermäßig vielen fehlenden Werten aus den Datenmatrizen und standardisierte die Merkmale. **(c)** Hauptkomponentenanalyse untersuchte die Varianz in den fünf Datenmatrizen. **(d)** Anwendung einer OPLS-Analyse untersuchte Effekte der Fütterung auf die Gesamtheit der Merkmale. **(d)** Vergleich der Zeitpunkte untersuchte Variationen bei der Auswirkung der Fütterung. **(e)** Analyse der Merkmalsgruppen identifizierte Mengen von Merkmalen mit besonders vielen Effekten. **(f)** Netzwerk-Analyse untersuchte die Effekte der Fütterung auf die Assoziation zwischen Merkmalsgruppen.

Entnahmeorten gemessen (Magen, Jejunum, Ileum und Colon). Der Datensatz enthielt die Messwerte von 64 Merkmalen, welche bei 54 Ferkeln bestimmt wurden (Tabelle E.2).

- **Vorkommen virulenter *Escherichia coli* Gene**

Die Bestimmung des Vorkommens unterschiedlicher *Escherichia coli* Klone im Verdauungstrakt der Ferkel und die genetische Charakterisierung der Klone wurde von Mitarbeitern des Instituts für Mikrobiologie und Tierseuchen durchgeführt (Bednorz et al. 2013<sup>[8]</sup>). Mit der Charakterisierung der Klone konnte das qualitative Vorkommen von 52 virulenten Genen an drei verschiedenen Entnahmeorten (Colon-Schleimhaut, Colon-Inhalt und Fäkalien) gemessen werden. Der erhaltende Datensatz enthielt die Messwerte von 141 Merkmalen, welche bei 24 Ferkeln bestimmt wurden (Tabelle E.3).

- **Transport-Funktion der Darmschleimhaut**

Die Transport- und Sekretions-Funktion der Darmschleimhaut wurden von Mitarbeitern des Instituts für Veterinär-Physiologie der Freien Universität Berlin bestimmt (Klingspor et al. 2013<sup>[56]</sup>). Die Ussing-Kammer-Versuche wurde mit 12 Proben pro Ferkel und mit 6 unterschiedlichen Untersuchungen durchgeführt. Zur Integration des Datensatzes wurde die Mittelwerte der Messwerte pro Untersuchung und pro Tier gebildet. Der Datensatz enthielt die Messwerte von 6 Merkmalen, welche bei 62 Tieren bestimmt wurden (Tabelle E.4).

- **Barrieren-Funktion der Darmschleimhaut**

Die Barrieren-Funktion der Darmschleimhaut wurde von Mitarbeitern des Instituts für Klinische Physiologie der Charité-Universitätsmedizin Berlin gemessen (Richter et al. 2011<sup>[83]</sup>). Mit Ussing-Kammer-Versuchen wurde der epitheliale Widerstand und die Permeabilität des Epithels für drei Marker-Moleküle bestimmt. Der Datensatz enthielt die Messwerte von 31 Merkmalen, welche bei 36 Tieren gemessen wurden (Tabelle E.5).

- **Morphologie der Darmschleimhaut**

Die Morphologie der Darmschleimhaut wurde durch Mitarbeiter des Instituts für Tierernährung der Freien Universität Berlin untersucht (Martin et al. 2012<sup>[67]</sup>). Hierbei erfolgten die Bestimmungen der durchschnittlichen Zotten-Länge und der durchschnittlichen Krypten-Tiefe. Der Datensatz enthielt die Messwerte von 3 Merkmalen, welche bei 50 Ferkeln bestimmt wurden (Tabelle E.6).

- **Aktivität der Verdauungsenzyme**

Die Aktivität der Verdauungsenzyme in der Bürstensaummembran des Jejunums wurde von Mitarbeitern des Instituts für Tierernährung der Freien Universität Berlin gemessen (Martin et al. 2012<sup>[67]</sup>). Hierbei wurde die Aktivität von fünf Enzymen (*Aminopeptidase N*, *Insteniale Alkalische Phosphatase*, *Sucrase-isomaltase*, *Lactase* und *Maltase*) bestimmt. Der Datensatz enthielt die Messwerte von 5 Merkmalen, welche bei 51 Tieren gemessen wurden (Tabelle E.7).

- **Genexpression des Darmepithels**

Die Expression der Marker-Gene im Darmepithel wurde von Mitglieder des Instituts für Tierernährung der Freien Universität Berlin durchgeführt (Martin et al. 2012<sup>[67]</sup>). Die Anwendung einer *real-time*-PCR bestimmte die relative Expression von insgesamt sieben Genen (ALPI, AMPN, CASP3, LCT, PCNA, SI, SLC5A1). Der Datensatz enthielt die Messwerte von 7 Merkmalen, welche bei 42 Tieren bestimmt wurden (Tabelle E.8).

- **Konzentration der Zytokine im Blut**

Die Konzentration von verschiedenen Zytokinen wurde von Mitarbeitern des Instituts für Mikrobiologie und Tierseuchen der Freien Universität Berlin bestimmt (Siepert et

al. 2014<sup>[95]</sup>). Die Anwendungen des Verfahrens ELISA bestimmte die Konzentration der Zytokine *Interferon gamma*, *Interleukin 6*, *Interleukin 8*, *Interleukin 13*, *Interleukin 17* und *Transforming growth factor beta* im Serum der Ferkel. Der Datensatz enthielt die Messwerte von 6 Merkmalen, welche bei 35 Ferkeln gemessen wurden (Tabelle E.9).

- **Expression immunrelevanter Gene**

Die Expression immunrelevanter Gene wurde durch Mitarbeiter des Instituts für Mikrobiologie und Tierseuchen der Freien Universität Berlin bestimmt (Siepert et al. 2014<sup>[95]</sup>). Die Anwendung einer *real-time*-PCR bestimmte die Expression von 10 Genen an 6 Entnahmeorten (Lymphknoten des Jejunums, Peyer-Plaques des Jejunums und des Ileums, Gewebe aus der Leber, Gewebe aus der Milz und Blutproben). Der Datensatz enthielt die Messwerte von 39 Merkmalen, welche bei 48 Ferkeln bestimmt wurden (Tabelle E.10).

- **Zusammensetzung der Immunzellpopulationen**

Der Anteil von bestimmten Immunzellpopulationen wurde von Mitarbeitern des Instituts für Züchtungsbiologie und molekulare Tierzucht der Humboldt-Universität zu Berlin bestimmt (Kreuzer et al. 2012<sup>[60]</sup>) und durch Mitarbeiter des Instituts für Tierernährung der Freien Universität Berlin bestimmt (Scharek-Tedin et al. 2015<sup>[90]</sup>). Mit der Anwendung des Verfahrens Durchflusszytometrie wurde das Vorkommen von insgesamt 16 Oberflächenmarkern in den Lymphknoten des Jejunums und des Ileums, in den Peyer-Plaques des Jejunums und des Ileums, im Blutüberstand des Ileums und in intraepithelialen Lymphozyten (IEL) des Jejunums und des Ileums gemessen. Der erste Datensatz enthielt die Messwerte von 61 Merkmalen, welche bei 46 Ferkeln bestimmt wurden. Der zweite Datensatz enthielt die Messwerte von 40 Merkmalen, welche bei 54 Ferkeln bestimmt wurden (Tabelle E.11).

- **Großes Blutbild**

Die Messung des Blutbildes wurde von Mitarbeitern des Instituts für Tierernährung der Freien Universität Berlin vorgenommen. Der Datensatz enthielt die Messwerte von 15 Merkmalen, welche bei 52 Ferkeln gemessen wurden (Tabelle E.12).

- **Gewicht der Tiere und Größe einzelner Organe**

Das Gewicht der Ferkel wurde durch Mitarbeiter des Instituts für Tierernährung der Freien Universität Berlin vor der Schlachtung der Ferkel bestimmt (Martin et al. 2012<sup>[67]</sup>). Das Gewicht einzelner Organe und die Länge des Dünndarms wurde nach der Entnahme der Organe gemessen. Der erhaltende Datensatz enthielt die Messwerte von 6 Merkmalen, welche bei 62 Ferkeln bestimmt wurden (Tabelle E.13).

## 5.2. Datenverarbeitung

Die integrierten Datensätze enthielten die Messwerte von insgesamt 492 unterschiedlichen Merkmalen, welche bei 62 Ferkeln gemessen wurden. Die Messungen erfolgten an vier unterschiedlichen Zeitpunkten (12 Tage, 26 Tage, 34 Tage und 54 Tage). Hierbei wurde jedes Merkmal an einem Tier nur an einem bestimmten Zeitpunkt gemessen. Durch die Integration ergab sich die Datenmatrix  $\tilde{\mathbf{X}}$ , welche die  $\tilde{K} = 492$  Merkmale als Spalten enthielt und die  $\tilde{N} = 62$  Ferkel als Zeilen enthielt. Der Anteil an unbekanntem Messwerten in der Matrix  $\tilde{\mathbf{X}}$  betrug 40,5 %. Mit der Aufteilung der Datenmatrix  $\tilde{\mathbf{X}}$  wurden vier weitere Datenmatrizen erstellt. Diese Datenmatrizen



**Tabelle 5.1.:** Merkmalsgruppen und Anzahl der Merkmale der *EFFT2010*-Datenmatrizen

Name	Beschreibung	X	X12	X26	X34	X54
BbmAct	Aktivität der Verdauungsenzyme	5	5	5	5	5
BloodCells	Großes Blutbild	11	11	11	11	11
CytoBlo	Konzentration der Zytokine im Blut	6	4	0	5	6
EcoliDig	Vorkommen der <i>E. coli</i> -Gene im Inhalt des Colons	0	29	36	0	0
EcoliFae	Vorkommen der <i>E. coli</i> -Gene in den Fäkalien	0	0	36	0	0
EcoliMuc	Vorkommen der <i>E. coli</i> -Gene an der Schleimhaut des Colons	0	0	29	0	0
EpiCol	Funktionalität des Colonepithels	13	0	13	13	6
Epille	Funktionalität des Ileumepithels	6	0	9	9	6
EpiJej	Funktionalität des Jejunumepithels	15	6	15	15	12
IsExprBlo	Expression immunrelevanter Genen in Blutzellen	3	3	3	3	3
IsExprLiv	Expression immunrelevanter Gene in der Leber	2	2	2	2	2
IsExprJejLN	Expression immunrelevanter Gene in den Lymphknoten des Jejunums	7	7	7	7	6
IsExprPPlle	Expression immunrelevanter Gene in den Peyer-Plaques des Ileums	10	10	10	10	6
IsExprPPJej	Expression immunrelevanter Gene in den Peyer-Plaques des Jejunums	10	10	9	10	8
IsExprSpln	Expression immunrelevanter Gene in der Milz	7	7	7	7	4
IsLymBlo	Zusammensetzung der Immunzellpopulation im Blut	30	30	30	30	33
IsLymIELlle	Zusammensetzung der Immunzellpopulation im Ileumepithel	9	9	9	9	9
IsLymIELJej	Zusammensetzung der Immunzellpopulation im Jejunumepithel	14	14	14	14	14
IsLymLNlle	Zusammensetzung der Immunzellpopulation in Lymphknoten des Ileums	13	13	13	13	16
IsLymJejLN	Zusammensetzung der Immunzellpopulation in Lymphknoten des Jejunums	10	10	10	10	13
IsLymPPlle	Zusammensetzung der Immunzellpopulation in Peyer-Plaques des Ileums	13	13	13	13	16
MarkerExpr	Genexpression des Darmepithels	7	7	7	7	0
MetabCol	Konzentration bakterieller Metabolite im Colon	14	9	14	15	14
Metablle	Konzentration bakterieller Metabolite im Ileum	11	1	1	14	14
MetabJej	Konzentration bakterieller Metabolite im Jejunum	13	0	10	13	12
MetabSto	Konzentration bakterieller Metabolite im Magen	13	13	12	13	13
MorphJej	Morphologie der Jejunumschleimhaut	3	3	3	3	3
OtuCol	Vorkommen bakterieller OTUs im Colon	17	17	17	17	17
Otulle	Vorkommen bakterieller OTUs im Ileum	17	16	17	17	17
OtuJej	Vorkommen bakterieller OTUs im Jejunum	17	17	17	17	17
OtuSto	Vorkommen bakterieller OTUs im Magen	17	17	17	17	17
Weight	Gewicht der Ferkel und Gewicht einzelner Organe	6	6	6	6	6

enthielten nur noch Messwerte von Ferkeln, deren Probenentnahme im gleichen Alter erfolgte (Abbildung 5.1 (a)).

Die Verarbeitung der Datenmatrizen erfolgte mit Skripten in der Programmiersprache R (Version 3.1.1) (b). Zuerst reduzierte die Anwendung eines Datenfilters den Anteil von fehlenden Messwerten in der Datenmatrix  $\mathbf{X}$  und in den vier weiteren Datenmatrizen. Der Filter entfernte alle Merkmale, welche einen Anteil von 50 % oder mehr unbekannte Werte enthielten. Dies war notwendig, da der Anteil unbekannter Werte in allen fünf Datenmatrizen relativ hoch war. Nach der Filterung erfolgte die Schätzung der übrigen unbekanntenen Werte mit dem *regularized iterative PCA* Algorithmus über die R-Funktion *estim\_ncpPCA()* aus dem R-Paket *missMDA* (Josse und Husson 2012<sup>[52]</sup>). Danach erfolgte die Standardisierung der Merkmale in den fünf Datenmatrizen, so dass alle Merkmale anschließend einen Mittelwert von 0 und eine Standardabweichung von 1 hatten. Die Bezeichnungen der fünf finalen Datenmatrizen lautete:  $\mathbf{X} = (x)_{N,K}$ ,  $\mathbf{X12} = (x)_{N12,K12}$ ,  $\mathbf{X26} = (x)_{N26,K26}$ ,  $\mathbf{X34} = (x)_{N34,K34}$  und  $\mathbf{X54} = (x)_{N54,K54}$ .

Neben den Datenmatrizen wurden fünf Objekt-Matrizen  $\mathbf{y} = (y)_{N,1}$ ,  $\mathbf{y12} = (y)_{N12,1}$ ,  $\mathbf{y26} = (y)_{N26,1}$ ,  $\mathbf{y34} = (y)_{N34,1}$  und  $\mathbf{y54} = (y)_{N54,1}$  gebildet. Diese Matrizen enthielten nur eine Spalte, die die Zugehörigkeit der Ferkel zu den zwei Behandlungsgruppen angab. Der Wert  $-1$  stand dabei für die Gruppe der Ferkel, die keine Zugabe des Probiotikums erhielten (Kontroll-Gruppe) und der Wert  $1$  stand für die Gruppe der Ferkel, die eine Zugabe des Probiotikums erhielten (Probiotikum-Gruppe). Die Anzahl der Zeilen der Objekt-Matrizen entsprach der Anzahl der Zeilen der zugehörigen Datenmatrizen.

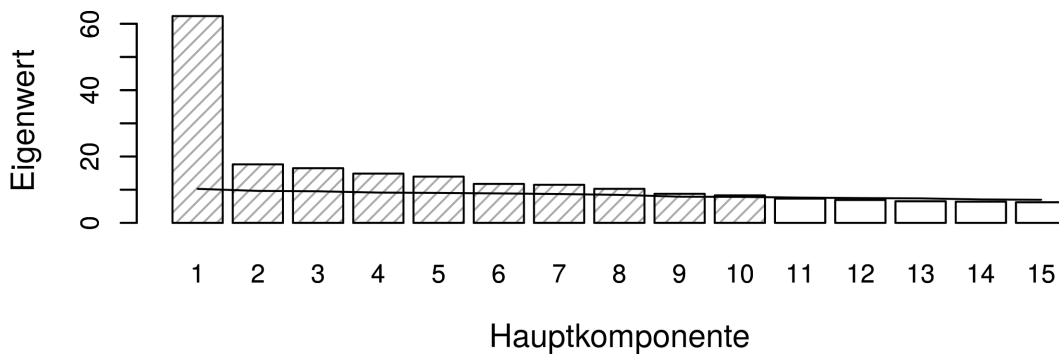
Alle gemessenen Variablen wurden auf Basis ihrer Bedeutung in  $L = 31$  Merkmalsgruppen  $\mathcal{L} = \{\mathcal{L}_1, \dots, \mathcal{L}_L\}$  kategorisiert. Die Zuordnung eines Merkmals ergab sich hierbei aus dem zugrunde liegenden Datensatz und aus dem zugehörigen Probenentnahmeort (siehe Abschnitt 5.1). Tabelle 5.1 zeigt eine Übersicht aller Merkmalsgruppen mit der Anzahl der enthaltenen Merkmale pro Gruppe. Eine Merkmalsgruppe  $\mathcal{L}_l = \{k : 1 \leq k \leq K\}$  war definiert als eine Menge von Indizes, wobei jeder Index  $k \in \{1, \dots, K\}$  ein Merkmal repräsentierte.

### 5.3. Hauptkomponentenanalyse

Die Durchführung der Hauptkomponentenanalyse (PCA, engl. *Principal Component Analysis*) diente zur Darstellung und zur Untersuchung der Variabilität der Datenmatrix  $\mathbf{X}$ . Eine PCA repräsentiert eine mehrdimensionale Datenmatrix mit einigen wenigen Hauptkomponenten (PC, *Principal Component*) (Handl 2010<sup>[39]</sup>[Seite 115ff]). Die Hauptkomponenten sind Linearkombinationen der Merkmale der Matrix, die die größten Variationen in der Matrix repräsentieren. Dabei sind die Hauptkomponenten untereinander unkorreliert und jede Hauptkomponente repräsentiert einen bestimmten Anteil der gesamten Varianz der Datenmatrix.

Die Berechnung der PCA wurden mit der R-Funktion *PCA()* aus dem R-Paket *FactoMineR* durchgeführt (Josse et al. 2008<sup>[63]</sup>). Die Anzahl der signifikanten Hauptkomponenten wurde mit der Methode *Parallel Analysis* berechnet (Horn 1965<sup>[44]</sup>, O'connor 2000<sup>[76]</sup>). Eine Hauptkomponente wurde als signifikant betrachtet, sobald sie einen höheren Eigenwert hatte als eine Hauptkomponente mit gleicher Ordnung von einer zufälligen, unkorrelierten, standard-normalverteilten Datenmatrix  $\mathbf{X}_{\text{random}} = (x)_{N,K}$ . Die PCA der Datenmatrix  $\mathbf{X}$  resultierte in zehn signifikante Hauptkomponenten (Abbildung 5.2). Die zehn signifikanten Hauptkomponenten repräsentierten insgesamt 55,1 % der Varianz der Datenmatrix  $\mathbf{X}$ .

Zur weiteren Untersuchung der Hauptkomponenten erfolgte die Durchführung einer Reihe von zweifaktoriellen Varianzanalysen (ANOVA, engl. *analysis of variance*) durch die Verwendung der



**Abbildung 5.2.:** Spektrum der Eigenwerte für die ersten 15 Hauptkomponenten von  $\mathbf{X}$ . Die schwarze Linie gibt die Eigenwerte der ersten 15 Hauptkomponenten für eine unkorrelierte standard-normalverteilte Datenmatrix an. Die Hauptkomponenten von  $\mathbf{X}$  mit einem höheren Eigenwert als die entsprechenden Hauptkomponenten der zufälligen Datenmatrix sind schraffiert dargestellt.

R-Funktion *aov()*. Jede ANOVA berücksichtigte die zwei Faktoren Behandlung und Zeitpunkt. Bei signifikanten Unterschieden zwischen den vier Zeitpunkten erfolgte die Anwendung eines *Tukey-Post-hoc*-Tests durch die R-Funktion *TukeyHSD()*. Tabelle 5.2 gibt die Ergebnisse der ANOVAs an.

Die ersten vier Hauptkomponenten zeigten deutliche Unterschiede zwischen den vier Zeitpunkten. Hierbei repräsentierten die vier Hauptkomponenten unterschiedliche Trennungen zwischen den Zeitpunkten. Die erste Hauptkomponente repräsentierte die stärkste Aufteilung der Ferkel bezüglich des Alters (Abbildung 5.3 **(a)**), wobei die Trennung zwischen den 26 und 34 Tage alten Ferkeln besonders ausgeprägt war. Die zweite Hauptkomponente hingegen zeigte eine klare Trennung zwischen den 34 und 54 Tage alten Ferkeln (Abbildung 5.3 **(b)**). Die dritte Hauptkomponente repräsentierte eine deutliche Trennung zwischen den 12 und 26 Tage alten Ferkeln (Abbildung 5.3 **(c)**). Die vierte Hauptkomponente repräsentierte eine Trennung zwischen den 26 Tage alten Ferkeln und den 34 beziehungsweise 54 Tage alten Ferkeln (Abbildung 5.3 **(d)**).

Die zehnte Hauptkomponente zeigte einen signifikanten Unterschied ( $P = 0,022$ ) zwischen den beiden Behandlungsgruppen (Tabelle 5.2), dabei repräsentierte die zehnte Hauptkomponente 2,6 % der gesamten Varianz der Datenmatrix (Abbildung 5.2).

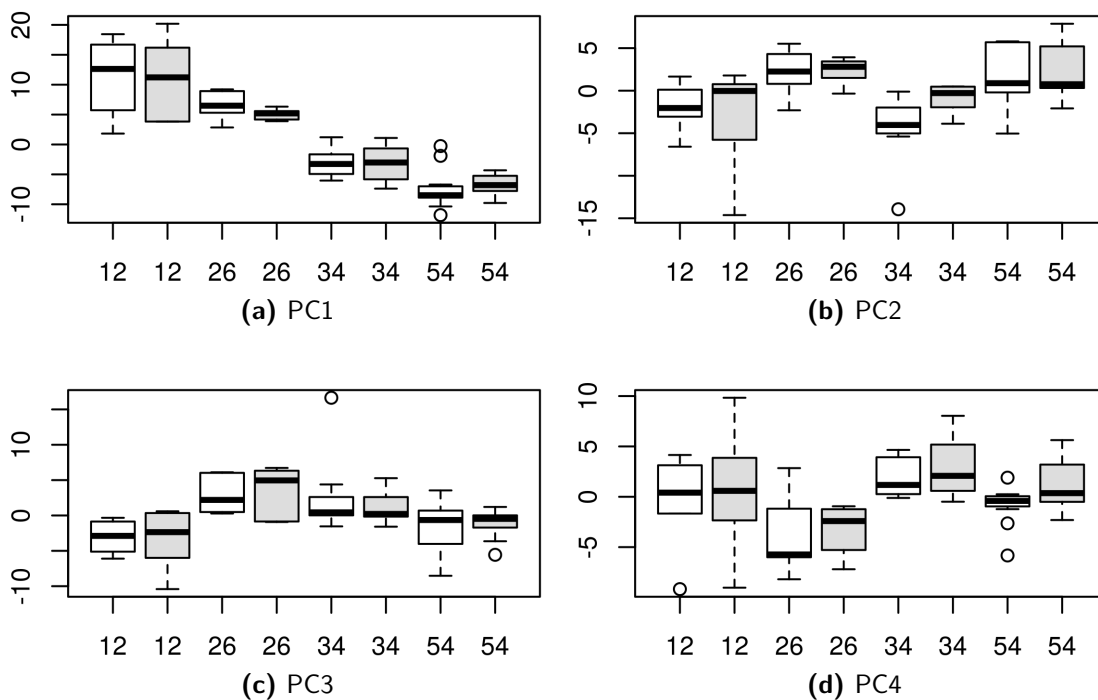
## 5.4. Orthogonal Projection to Latent Structures Analyse

Zur weiteren Untersuchung der Auswirkung der Fütterung des Probiotikums auf die gemessenen Merkmale wurde überprüft, ob sich die Behandlungsgruppen der Ferkel aus den Datenmatrizen vorhersagen ließen und welche Merkmale für eine mögliche Vorhersage von Bedeutung waren. Für diese Analysen wurden die Merkmale aus den Datenmatrizen entfernt, welche das Vorkommen von *E. faecium* NCIMB 10415 im Verdauungstrakt beschrieben. Die Analyse wurde mit den Matrizen  $\mathbf{X}$ ,  $\mathbf{X}_{12}$ ,  $\mathbf{X}_{26}$ ,  $\mathbf{X}_{34}$  und  $\mathbf{X}_{54}$  durchgeführt. Für jeden Datensatz, bestehend aus Datenmatrix und Gruppen-Matrix, wurde ein optimales Vorhersagemodell berechnet. Jedes Vorhersagemodell diente zur Vorhersage der Gruppenzugehörigkeiten der zugehörigen Ferkel. Die Vorhersagequalität eines Modells wurden über die Anwendungen einer Kreuzvalidierung bestimmt. Nach der Kreuzvalidierung wurde mit einem Permutationstest die Signifikanz der Vorhersagequalität bestimmt. Die

**Tabelle 5.2.:** Ergebnisse der zweifaktoriellen Hauptkomponenten-ANOVA mit den beiden Faktoren Zeitpunkt und Behandlung<sup>1</sup>

PC	P-Wert			Z Unterschiede
	Z	B	Z×B	
1	< 0.001***	0.879	0.882	(12) - (26) - (34) - (54)
2	< 0.001***	0.284	0.347	(12,34) - (26,54)
3	< 0.001***	0.856	0.810	(12,54) - (26,34)
4	< 0.001***	0.128	0.937	(12,26) - (12,34,54)
5	0.051	0.249	0.494	
6	0.173	0.587	0.725	
7	0.734	0.545	0.507	
8	0.068	0.929	0.625	
9	0.635	0.989	0.946	
10	0.717	0.022*	0.960	

<sup>1</sup>Z = Zeitpunkt; B = Behandlung; \* $P < 0.05$ ; \*\*\* $P < 0.001$



**Abbildung 5.3.:** Boxplots der ersten vier Hauptkomponenten der Datenmatrix **X**. Jede Teilgrafik zeigt eine andere Hauptkomponente: **(a)** PC1, **(b)** PC2, **(c)** PC3 und **(d)** PC4. Boxen verdeutlichen die Streuung der Hauptkomponenten (Y-Achse), getrennt nach Zeitpunkt (X-Achse) und Behandlungsgruppen (Farbe der Boxen). Weiß repräsentiert die Kontroll-Gruppe und eine graue Farbe repräsentiert die Probiotikum-Gruppe.

Kombination von Kreuzvalidierung mit einem Permutationstest stellt eine besonders zuverlässige Methode zur Validierung von Vorhersagemodellen dar (Westerhuis et al. 2008<sup>[118]</sup>).

### 5.4.1. Berechnung der optimalen Vorhersagemodelle

Da die Anzahl der Merkmale  $K$  die Anzahl der Ferkel  $N$  in allen fünf Datenmatrizen um ein vielfaches überstieg (Abbildung 5.1), erfolgte die Berechnung der Vorhersagemodelle mit dem Verfahren *Orthogonal Projection to Latent Structures* (OPLS) unter Verwendung des NIPALS (engl. *Nonlinear Iterative Partial Least Squares*) Algorithmus (Trygg und Wold 2002<sup>[107]</sup>). Ein OPLS-Modell berechnet eine reduzierte Darstellung einer Datenmatrix unter der Berücksichtigung einer Objekt-Matrix. Im Falle von einer mehrdimensionalen Datenmatrix  $\mathbf{X}$  und einer eindimensionalen Objekt-Matrix  $\mathbf{y}$  stellt ein OPLS-Modell die Datenmatrix  $\mathbf{X}$  mit genau einer vorhersagenden Komponente und mit zusätzlichen orthogonalen Komponenten dar. Die Anzahl der orthogonalen Komponenten definiert die Komplexität eines OPLS-Modells (siehe Appendix B).

Für jeden Datensatz wurden verschiedenen OPLS-Modelle mit steigender Anzahl an orthogonalen Komponenten berechnet. Zur Bestimmung eines optimalen OPLS-Modells  $\mathcal{M}$  für einen Datensatz  $(\mathbf{y}, \mathbf{X})$  wurde der folgende Algorithmus verwendet (Algorithmus 1):

**Zeile 2:** Zuerst wurde für jede mögliche Anzahl von orthogonalen Komponenten ein OPLS-Modell berechnet.

**Zeile 3:** Die Vorhersagequalität der OPLS-Modelle wurde über eine Kreuzvalidierung berechnet

**Zeile 5:** und anschließend wurde die optimale Anzahl an orthogonalen Komponenten  $A_o$  bestimmt.

**Zeile 7:** Das Modell  $\mathcal{M}$  mit dem höchsten Prognose-Bestimmtheitsmaß  $Q^2$  (Vorhersagekraft) wurde als finales Modell bestimmt.

#### Algorithmus 1: Berechnung des optimalen OPLS-Modells zur Vorhersage von $\mathbf{y}$ aus $\mathbf{X}$

```

Input :  $\mathbf{X}, \mathbf{y}, A_{o_{max}}$ 
1 for  $A_{o_{temp}} \in \{1 \dots A_{o_{max}}\}$  do
2   | Berechnung eines OPLS-Modells  $\mathcal{M}_{A_{o_{temp}}}$  mit  $(\mathbf{X}, \mathbf{y}, A_{o_{temp}})$ 
3   | Berechnung des durchschnittlichen Prognose-Bestimmtheitsmaßes
   |  $Q^2(\mathcal{M}_{A_{o_{temp}}}) = Q^2_{A_{o_{temp}}}$  durch Kreuzvalidierung
4 end
5  $A_o = \underset{A_{o_{temp}} \in \{1, \dots, A_{o_{max}}\}}{\operatorname{argmax}} Q^2(\mathcal{M}_{A_{o_{temp}}})$ 
6  $Q^2 = Q^2(\mathcal{M}_{A_o})$ 
7  $\mathcal{M} = \mathcal{M}_{A_o}$ 
Output :  $(\mathcal{M}, Q^2, A_o)$ 

```

### 5.4.1.1. Kreuzvalidierung

Die Validierung der Vorhersagequalität eines OPLS-Modells  $\mathcal{M}_{A_{otemp}}$  mit  $A_{otemp}$  orthogonalen Komponenten erfolgte durch die Anwendung einer Kreuzvalidierung (CV, engl. *cross-validation*) mit der Brechung der durchschnittlichen Vorhersagekraft  $Q^2(\mathcal{M}_{A_{otemp}}) = Q_{A_{otemp}}^2$  (Abbildung 5.4). Zuerst wurden zufällige Paare von Indizes gezogen, welche die Kreuzvalidierungsgruppen (CV-Gruppen) bildeten (Abbildung 5.4 (a)). Jedes Index-Paar stand für ein Ferkel aus der Kontroll-Gruppe und für ein Ferkel aus der Probiotikum-Gruppe. Für jede CV-Gruppe wurde ein neues OPLS-Modell berechnet (Abbildung 5.4 (b)). Die Komplexität der neuen Modelle entsprach dabei der Komplexität des ursprünglichen Modells  $\mathcal{M}_{A_{otemp}}$ . Zur Modellberechnung wurden die Daten der zugehörigen Ferkel-Paare aus der Datenmatrix und aus der Gruppen-Matrix entfernt. Das CV-Modell wurde dann benutzt, um die Gruppenzugehörigkeit der beiden ausgelassenen Ferkel vorherzusagen (Abbildung 5.4 (c)). Die Modellberechnung und die anschließende Vorhersage wurde für alle CV-Gruppen durchgeführt. Mit den vorhergesagten Gruppenzugehörigkeiten  $\widehat{y}^{cv} = (\widehat{y}^{cv})_{N,1}$  konnte  $Q_{cv}^2$  für einen CV-Durchgang als 1 minus der Variation der Residuen geteilt durch die Varianz von  $\mathbf{y}$  berechnet werden (Westerhuis et al. 2008<sup>[118]</sup>) (Abbildung 5.4 (d)):

$$Q_{cv}^2 = 1 - \frac{\sum_{n=1}^N (y_n - \widehat{y}^{cv}_n)^2}{\sum_{n=1}^N (y_n - \bar{y})^2} \quad (5.1)$$

Die durchschnittliche Vorhersagekraft  $Q_{A_{otemp}}^2$  für das Modell errechnete sich dann aus dem Mittelwert von 30 Wiederholungen der Kreuzvalidierung (Abbildung 5.4 (e)).

### 5.4.2. Permutationstest

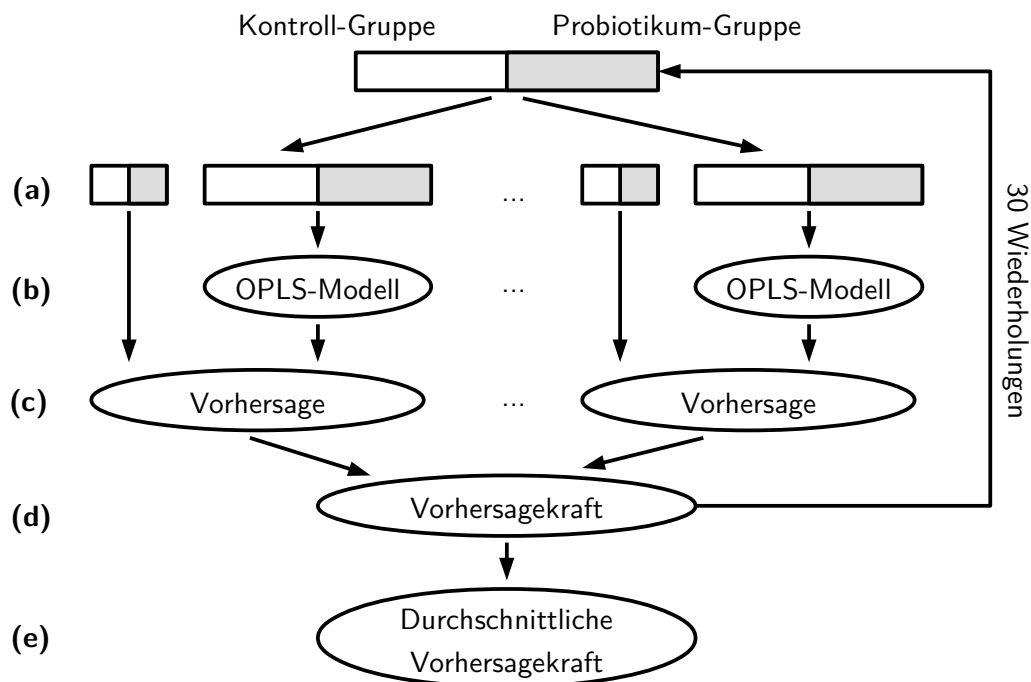
Nach der Berechnung der fünf finalen Modelle wurde die Bedeutung der  $Q^2$ -Werte mit der Berechnung eines Monte-Carlo-Permutationstests bestimmt (Radmacher et al. 2002<sup>[82]</sup> und Westerhuis et al. 2008<sup>[118]</sup>). Für jeden Datensatz, bestehend aus Datenmatrix und Objekt-Matrix, wurden 1000 zufällige Permutation der Gruppenzugehörigkeiten erzeugt. Für jeden permutierten Datensatz wurde ein optimales OPLS-Modell berechnet. Der Anteil der permutierten Modelle mit einem gleichen oder höheren  $Q^2$ -Wert als das ursprüngliche Modell wurde als P-Wert für die Vorhersagequalität des ursprünglichen Modells angenommen.

### 5.4.3. Bedeutung der Merkmale

Die Bedeutung der einzelnen Merkmale für die Vorhersage der Gruppenzugehörigkeiten wurde mit dem *Variable Importance of the Projection* (VIP)-Wert bestimmt. Die Berechnung der VIP-Werte ist eine verbreitete Methode und eignet sich insbesondere zur Selektion von Merkmalen (Chong und Jun 2005<sup>[21]</sup>). Der VIP-Wert eines Merkmals  $j$  berechnet sich aus der *predictive-weight*-Matrix  $\mathbf{W} = (w)_{1,K}$  (siehe Appendix B) des zugehörigen OPLS-Modells (Chong und Jun 2005<sup>[21]</sup>):

$$VIP_j = \sqrt{K} \left| \frac{W_{1,j}}{\|\mathbf{W}_{1,\bullet}\|_2} \right| \quad (5.2)$$

Der VIP-Wert für ein Merkmal  $j$  entspricht dem Gewicht des Merkmals geteilt durch die euklidische Norm des Vektors  $\mathbf{W}_{1,\bullet}$ , multipliziert mit der Wurzel der Anzahl der Merkmale  $K$ . Je höher ein VIP-Wert, desto größer ist die Bedeutung eines Merkmals für die Vorhersage der Gruppenzugehörigkeit.



**Abbildung 5.4.:** Durchführung der Kreuzvalidierung für ein OPLS-Modell. Die Berechnung basiert auf einem Datensatz mit den zwei Behandlungsgruppen Kontroll-Gruppe (weiß) und Probiotikum-Gruppe (grau). Gezeigt werden die fünf Schritte zur Berechnung der durchschnittlichen Vorhersagekraft: **(a)** Aufteilung des Datensatzes in CV-Gruppen, **(b)** Berechnung von OPLS-Modellen, **(c)** Vorhersage der Gruppenzugehörigkeit der ausgelassenen Ferkel und **(d)** Berechnung der Vorhersagekraft. Zur Berechnung der durchschnittlichen Vorhersagekraft **(e)** wurde die Schritte **(a)** bis **(d)** 30 mal wiederholt.

Die Matrizen  $\mathbf{VIP} = (VIP)_{1,K}$ ,  $\mathbf{VIP12} = (VIP)_{1,K12}$ ,  $\mathbf{VIP26} = (VIP)_{1,K26}$ ,  $\mathbf{VIP34} = (VIP)_{1,K34}$  und  $\mathbf{VIP54} = (VIP)_{1,K54}$  enthielten die VIP-Werte für die Merkmale der fünf optimalen OPLS-Modelle.

#### 5.4.4. Ergebnisse der OPLS-Analyse

Das optimale Modell für die Datenmatrix  $\mathbf{X}$  ergab eine unzureichende Vorhersagekraft mit  $Q^2 = -0,34$  (Tabelle 5.3). Der Wert deutlich unter Null zeigte, dass sich das Modell nicht zur Vorhersage der Behandlungsgruppen eignete. Die beiden Modelle für die Datenmatrix der 12 Tage alten Tiere ( $\mathbf{X12}$ ) und für die Datenmatrix der 54 Tage alten Tiere ( $\mathbf{X54}$ ) ergaben ein ähnliches Ergebnis ( $Q^2 = -0,25$ ,  $Q^2 = -0,13$ , Tabelle 5.3). Keines der beiden Modelle ließ eine stabile Vorhersage der Behandlungsgruppen zu. Die beiden Modelle für die Daten der 26 Tage alten Tiere ( $\mathbf{X26}$ ) und für die Daten der 34 Tage alten Tiere ( $\mathbf{X34}$ ) erreichten eine positive Vorhersagekraft ( $Q^2 = 0,11$ ,  $Q^2 = 0,08$ ). Ein Wert  $Q^2 > 0$  zeigt, dass ein Modell eine bessere Vorhersage liefert als eine rein zufällige Zuordnung. Die Permutationstests ergaben in keinem der fünf Fälle einen P-Werte kleiner als 0,05.

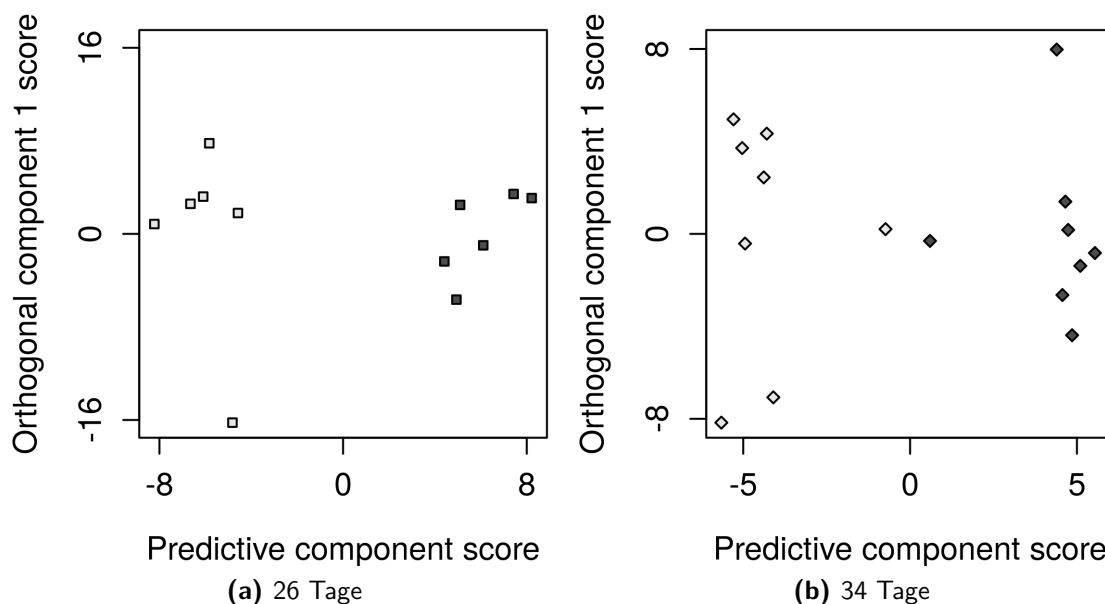
Die Berechnungen der OPLS-Modelle wurden mit reduzierten Datenmatrizen wiederholt. In

**Tabelle 5.3.:** Ergebnisse der OPLS-Diskriminanzanalyse<sup>1</sup>

Datenmatrix	100 %		85 %		75 %	
	Q <sup>2</sup>	P-Wert	Q <sup>2</sup>	P-Wert	Q <sup>2</sup>	P-Wert
<b>X</b>	-0,34	0,924	-0,10	0,470	0,07	
<b>X12</b>	-0,25	0,817	0,07	0,396	0,23	0,152
<b>X26</b>	0,11	0,264	0,27	0,091	0,41	0,022*
<b>X34</b>	0,08	0,354	0,24	0,121	0,35	0,026*
<b>X54</b>	-0,13	0,758	0,11	0,290	0,22	0,118

<sup>1</sup> \*P < 0.05

diesen Datenmatrizen waren 85 % beziehungsweise 75 % der Merkmale mit den höchsten VIP-Werten enthalten. Diese Berechnung konnte die Vorhersagequalität der Modelle der 26 Tage alten Tiere und der 34 Tage alten Tiere deutlich verbessern (**X26**: Q<sup>2</sup> = 0,41; **X34** Q<sup>2</sup> = 0,35) (Tabelle 5.3). Die Permutationstests zeigten, dass diese beiden Modelle signifikant besser waren als Datensätze mit permutierten Gruppenzugehörigkeiten (**X26**: P = 0,022; **X34**: P = 0,026). Abbildung 5.5 zeigt die klare Trennung der Behandlungsgruppen in dem reduzierten Datenraum. Mit beiden Modellen konnte eine Trennung der Datenpunkte durch eine gerade Linie erreicht werden.



**Abbildung 5.5.:** Projektion der Ferkel durch die Komponenten der OPLS-Modelle zeigt deutliche Trennungen der Ferkel nach den Behandlungsgruppen. Teilgrafik **(a)** zeigt 26 Tage alte Tiere und Teilgrafik **(b)** zeigt 34 Tage alte Tiere. Weiße Symbole repräsentieren Ferkel aus der Kontroll-Gruppe und graue Symbole repräsentieren Ferkel aus der Probiotikum-Gruppe.



## 5.5. Vergleich von unterschiedlichen Zeitpunkten

Im nächsten Schritt wurde eine zeitliche Veränderung der Auswirkung der Probiotikum-Fütterung untersucht. Um den Effekt der Fütterung zwischen den Altersgruppen zu vergleichen, wurde der Spearman-Korrelationskoeffizient mit den VIP-Werten der Merkmale (Formel 5.2) zwischen den vier Zeitpunkten berechnet. Mit dem Spearman-Korrelationskoeffizienten ließen sich die Reihenfolgen der Merkmale in Bezug auf deren VIP-Werte berücksichtigen. Bei den paarweisen Berechnungen wurden die Merkmale ausgelassen, die jeweils nur zu einem der beiden Zeitpunkte gemessen wurden. Die Bedeutung eines paarweisen Korrelationskoeffizienten wurde durch die Anwendung eines Monte-Carlo-Permutationstests mit  $10^3$  Iterationen berechnet. Für jede paarweise Berechnung wurden 1000 zufällige Permutation der beiden beteiligten Datensätze erzeugt und neue optimale OPLS-Modelle berechnet. Der Anteil der permutierten Paare mit einem höheren Korrelationskoeffizienten wurde als P-Wert für den Korrelationskoeffizienten des ursprünglichen Paares angenommen.

Die Berechnungen ergaben signifikante paarweise Korrelationen zwischen nachfolgenden Zeitpunkten (Tabelle 5.4). Besonders ausgeprägt war die Korrelation zwischen 26 Tage alten Tieren und 34 Tage alten Tieren ( $P = 0,004$ ) und zwischen 34 Tage alten Tieren und 54 Tage alten Tieren ( $P = 0,025$ ). In den beiden Fällen waren die Korrelationskoeffizienten zwar signifikant jedoch relativ gering. Der Vergleich der VIP-Werte der 26 und 54 Tage alten Tieren zeigte sogar eine geringe negative Korrelation ( $\rho = -0.16$ ,  $P = 0,010$ ). Die VIP-Werte der 12 Tage alten Tieren korrelierten am geringsten mit den VIP-Werten der anderen drei Zeitpunkte.

**Tabelle 5.4.:** Spearman-Korrelationskoeffizient der VIP-Werte zwischen den vier Zeitpunkten<sup>1</sup>

	26 Tage		34 Tage		54 Tage	
	$\rho$	P-Wert	$\rho$	P-Wert	$\rho$	P-Wert
<b>12 Tage</b>	0,01	0,467	0,11	0,084	0,02	0,397
<b>26 Tage</b>			0,18	0,004**	-0,16	0,010*
<b>34 Tage</b>					0,14	0,025*

<sup>1</sup> \*  $P < 0.05$ ; \*\*  $P < 0.01$

## 5.6. Analyse der Merkmalsgruppen

Im nächsten Schritt wurde die Auswirkung der Fütterung des Probiotikums auf die vorher definierten Merkmalsgruppen berechnet (Tabelle 5.1). Die Berechnung erfolgte mit einer Variation des *Gene Set Enrichment Algorithmus* (Subramanian et al. 2005<sup>[103]</sup>). Dieser Algorithmus dient üblicherweise zur Analyse von Daten aus Gen-Expressions-Experimenten und kann Gen-Gruppen identifizieren, die besonders stark durch eine Behandlung beeinflusst wurden. Die Anwendung des Algorithmus erfolgte für die Datenmatrizen der vier Zeitpunkte (**X12**, **X26**, **X34** und **X54**).

### 5.6.1. Berechnung der *Enrichment Analysis*

Die folgende Berechnung geht von einem Effekt-Vektor  $\mathbf{VIP} = (VIP_1, \dots, VIP_K)$  der Länge  $K$  und einer Menge von  $L$  Merkmalsgruppen  $\mathcal{L} = \{\mathcal{L}_1, \dots, \mathcal{L}_L\}$  aus. Jeder Eintrag von  $\mathbf{VIP}$  repräsentiert die Stärke der Auswirkung der Probiotikum-Fütterung auf ein zugehöriges Merkmal (Gleichung 5.2). Ein hoher VIP-Wert steht für ein hohen Effekt der Fütterung auf das Merkmal. Jedes Merkmal ist dabei ein Element von genau einer Merkmalsgruppe  $\mathcal{L}_l \in \mathcal{L}$ . Zur Bestimmung des Effekts der Fütterung des Probiotikums auf eine Merkmalsgruppe  $\mathcal{L}_l$  lassen sich für jede Position  $k$  in  $\mathbf{VIP}$  die zwei Werte  $A_{k,l}$  und  $B_{k,l}$  berechnen. Die beiden Matrizen  $\mathbf{A} = (A)_{K,L}$  und  $\mathbf{B} = (B)_{K,L}$  repräsentieren diese Werte für alle Merkmalsgruppen  $l \in \{1, \dots, L\}$  und für alle Merkmale  $k \in \{1, \dots, K\}$ .

- A** Ein Wert  $A_{kl}$  gibt die Anzahl der Merkmale  $k'$  an, die zu der Merkmalsgruppe  $\mathcal{L}_l$  gehören und einen höheren oder gleichen VIP-Wert wie  $\mathbf{VIP}_k$  haben, geteilt durch die Anzahl aller Merkmale, die zu der Merkmalsgruppe  $\mathcal{L}_l$  gehören.

$$A_{k,l} = \frac{|\{k' : 1 \leq k' \leq K \text{ und } k' \in \mathcal{L}_l \text{ und } VIP_{k'} \geq VIP_k\}|}{|\{k' : 1 \leq k' \leq K \text{ und } k' \in \mathcal{L}_l\}|} \quad (5.3)$$

- B** Ein Wert  $B_{kl}$  gibt die Anzahl der Merkmale  $k'$  an, die nicht zu der Merkmalsgruppe  $\mathcal{L}_l$  gehören und einen höheren oder gleichen VIP-Wert wie  $\mathbf{VIP}_k$  haben, geteilt durch die Anzahl aller Merkmale die nicht zu der Merkmalsgruppe  $\mathcal{L}_l$  gehören.

$$B_{k,l} = \frac{|\{k' : 1 \leq k' \leq K \text{ und } k' \notin \mathcal{L}_l \text{ und } VIP_{k'} \geq VIP_k\}|}{|\{k' : 1 \leq k' \leq K \text{ und } k' \notin \mathcal{L}_l\}|} \quad (5.4)$$

Der *Enrichment-Score* (ES) für eine Merkmalsgruppe  $\mathcal{L}_l$  ist die maximale Differenz zwischen den zugehörigen Einträgen aus den Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  über alle Positionen  $k \in \{1, \dots, K\}$ :

$$ES(l) = \max_{k \in \{1, \dots, K\}} (A_{k,l} - B_{k,l}) \quad (5.5)$$

Die Berechnung liefert für eine Merkmalsgruppe den maximal möglichen ES-Wert von 1, falls kein Merkmal aus anderen Merkmalsgruppen einen höheren VIP-Wert hat, als die Merkmale, die zu dieser Merkmalsgruppe gehören.

Die ES-Werte wurde für alle Merkmalsgruppen innerhalb der vier Messzeitpunkte bestimmt. Die Bedeutung der ES-Werte wurde mit einem Monte-Carlo-Permutationstests bestimmt. Für jeden der vier Datensätze wurden 1000 zufällige Permutationen der Gruppenzugehörigkeiten erzeugt und anschließend wurden die VIP-Werte neu berechnet. Für jede Kombination von Merkmalsgruppe und Datensatz wurde der Anteil der Permutationen mit einem höheren ES-Wert berechnet und als P-Wert des Tests angenommen.

### 5.6.2. Ergebnisse der *Enrichment Analysis*

Insgesamt sechs Merkmalsgruppen zeigten einen signifikanten ES-Wert (Tabelle 5.5). Bei den 12 Tage alten Ferkeln wurde kein Effekt auf die Merkmalsgruppen identifiziert. Bei den 26 Tage alten Ferkeln konnte eine Auswirkung auf das Vorkommen virulenter *E. coli*-Gene an der Schleimhaut des Colons (*EcoliMuc*), auf die Genexpression des Darmepithels (*MarkerExpr*) und auf die Konzentration bakterieller Metabolite im Jejunum (*MetabJej*) identifiziert werden. Bei

den 34 Tage alten Ferkeln konnte eine Auswirkung auf die Expression immunrelevanter Gene in den Peyers-Plaques des Ileums (*IsExprPPIle*) und auf das Vorkommen bakterieller OTUs im Ileum (*Otulle*) identifiziert werden. Bei den 54 Tage alten Tiere konnte eine Auswirkung auf die Expression immunrelevanter Gene in den Peyers-Plaques des Ileums (*IsExprPPIle*) und auf das Gewicht der Ferkel und die Größe einzelner Organe (*Weight*) nachgewiesen werden.

**Tabelle 5.5.:** Ergebnisse der *Enrichment Analysis*<sup>1</sup>

Merkmalsgruppe	12 Tage	26 Tage	34 Tage	54 Tage
<i>BbmAct</i>	0,50 (0,088)	0,04 (0,659)	0,00 (0,934)	0,09 (0,606)
<i>BloodCells</i>	0,21 (0,241)	-0,01 (0,998)	0,00 (0,737)	0,00 (0,985)
<i>CytoBlo</i>	0,29 (0,341)		0,23 (0,504)	0,44 (0,190)
<i>EcoliDig</i>	0,29 (0,333)	0,20 (0,429)		
<i>EcoliFae</i>		0,33 (0,677)		
<i>EcoliMuc</i>		<b>0,56 (0,033)*</b>		
<i>EpiCol</i>		0,00 (0,794)	0,15 (0,258)	0,00 (0,985)
<i>Epille</i>		0,03 (0,644)	0,14 (0,401)	0,21 (0,348)
<i>EpiJej</i>	0,21 (0,322)	0,14 (0,293)	0,05 (0,627)	0,13 (0,368)
<i>IsExprBlo</i>	0,28 (0,421)	0,52 (0,143)	0,32 (0,363)	0,00 (0,982)
<i>IsExprPPIle</i>	0,16 (0,544)	-0,02 (0,994)	<b>0,69 (0,009)**</b>	<b>0,59 (0,034)*</b>
<i>IsExprJejLN</i>	0,03 (0,664)	-0,03 (0,999)	0,36 (0,134)	0,27 (0,344)
<i>IsExprPPJej</i>	0,11 (0,484)	0,06 (0,643)	0,02 (0,787)	0,29 (0,174)
<i>IsExprLiv</i>	0,00 (0,788)	0,15 (0,671)	0,07 (0,732)	0,42 (0,277)
<i>IsExprSpln</i>	0,32 (0,168)	0,05 (0,886)	0,00 (0,888)	0,18 (0,683)
<i>IsLymBlo</i>	0,02 (0,658)	0,15 (0,176)	0,07 (0,388)	0,02 (0,617)
<i>IsLymLELle</i>	0,04 (0,663)	-0,03 (0,998)	0,00 (0,905)	0,26 (0,268)
<i>IsLymLELJej</i>	0,22 (0,227)	-0,03 (0,999)	0,04 (0,586)	0,00 (0,990)
<i>IsLymLNlle</i>	0,00 (0,921)	-0,03 (0,999)	0,02 (0,568)	0,23 (0,160)
<i>IsLymJejLN</i>	0,00 (0,865)	-0,03 (0,999)	0,00 (0,892)	0,05 (0,526)
<i>IsLymPPIle</i>	0,00 (0,852)	-0,03 (0,999)	0,00 (0,839)	0,02 (0,734)
<i>MarkerExpr</i>	0,20 (0,369)	<b>0,46 (0,032)*</b>	0,38 (0,076)	
<i>MetabCol</i>	0,19 (0,350)	-0,03 (0,999)	0,15 (0,311)	0,00 (0,994)
<i>Metablle</i>	0,78 (0,179)	-0,03 (0,999)	0,10 (0,485)	0,17 (0,288)
<i>MetabJej</i>		<b>0,34 (0,035)*</b>	0,14 (0,389)	0,14 (0,312)
<i>MetabSto</i>	0,00 (0,967)	0,07 (0,591)	0,00 (0,965)	0,33 (0,141)
<i>MorphJej</i>	0,00 (0,949)	0,26 (0,457)	0,23 (0,478)	0,16 (0,547)
<i>OtuCol</i>	0,06 (0,486)	-0,02 (0,998)	0,22 (0,135)	0,00 (0,968)
<i>Otulle</i>	0,25 (0,138)	0,06 (0,512)	<b>0,35 (0,020)*</b>	0,06 (0,439)
<i>OtuJej</i>	0,06 (0,416)	0,28 (0,216)	0,00 (0,938)	0,00 (0,955)
<i>OtuSto</i>	0,00 (0,804)	0,08 (0,441)	0,13 (0,291)	0,00 (0,984)
<i>Weight</i>	0,52 (0,073)	0,06 (0,721)	0,16 (0,358)	<b>0,62 (0,045)*</b>

<sup>1</sup> Die Bedeutung der Merkmalsgruppen stehen in Tabelle 5.1. Die Werte geben den ES-Wert der Merkmalsgruppen in den vier Datenmatrizen an. Die P-Werte stehen in Klammern; \* $P < 0,05$ ; \*\* $P < 0,01$

## 5.7. Netzwerkanalyse

Als nächstes wurde die Auswirkung der Fütterung des Probiotikums auf die Korrelation zwischen den Merkmalsgruppen untersucht. Hierfür wurden zwei Korrelationsnetzwerke berechnet. Ein Korrelationsnetzwerk ist ein kompletter, ungerichteter, gewichteter Graph  $G = (V, E, w)$ . Jeder Knoten  $v \in V$  steht hierbei für ein Merkmal. Die Menge  $E$  bezeichnet die Menge der Kanten in dem Graph. Es handelt sich hierbei um einen kompletten Graph, somit kommt jede mögliche Kante in dem Graph vor. Das Gewicht einer Kante wird mit der Gewichtsfunktion  $w$  bestimmt. Für das Gewicht der Kanten sind beliebige Funktionen denkbar. Im Folgenden berechnet sich das Kantengewicht als absoluter Pearson-Korrelationskoeffizient zwischen den zugehörigen Merkmalen. Somit gibt der Kantenwert des Graphen den linearen Zusammenhang der Merkmale an.

$$w(v_j, v_k) = \left| \frac{1}{N} \times \sum_{i=1}^N x_{i,j} \times x_{i,k} \right| \quad (5.6)$$

### 5.7.1. Berechnung der Graphen und Vergleich der Merkmalsgruppen

Die Berechnung mit der Datenmatrix  $\mathbf{X}$  resultierte in zwei Graphen. Der Graph  $GK = (V, E, w_K)$  (Kontroll-Graph) basierte auf den Messwerten der Ferkel aus der Kontroll-Gruppe und der Graph  $GP = (V, E, w_P)$  (Probiotikum-Graph) basierte auf den Messwerten der Ferkel aus der Probiotikum-Gruppe. Beide Graphen enthielten die identischen Mengen an Knoten und Kanten und unterschieden sich beim Gewicht der Kanten. Durch den Vergleich der beiden Graphen wurde die Auswirkung der Fütterung des Probiotikums auf die Assoziation zwischen den Merkmalsgruppen berechnet. Hierfür wurde für jedes Paar von zwei Merkmalsgruppen  $\mathcal{L}_l \in \mathcal{L}$  und  $\mathcal{L}_m \in \mathcal{L}$  der Wert *score* berechnet:

$$score(\mathcal{L}_l, \mathcal{L}_m) = \log \left( \frac{|\{e : e \in E_{lm} \text{ und } w_K(e) < w_P(e)\}|}{|\{e : e \in E_{lm} \text{ und } w_K(e) > w_P(e)\}|} \right) \quad (5.7)$$

Die Teilmenge  $E_{lm} \subset E$  enthielt hierbei alle Kanten, die zwischen den Knoten der zwei Merkmalsgruppen  $\mathcal{L}_l$  und  $\mathcal{L}_m$  lagen. Für zwei Merkmalsgruppen  $\mathcal{L}_l$  und  $\mathcal{L}_m$  entsprach der Wert  $score(\mathcal{L}_l, \mathcal{L}_m)$  dem Logarithmus des Verhältnisses von der Anzahl der Kanten zwischen den zwei Merkmalsgruppen, die ein höheres Kantengewicht im Probiotikum-Graphen hatten, zu der Anzahl der Kanten zwischen den zwei Merkmalsgruppen, die ein höheres Kantengewicht im Kontroll-Graphen hatten. Ein *score*-Wert von 0 bedeutete, dass die Anzahl der Kanten mit einem höheren Kantenwert im Probiotikum-Graphen gleich der Anzahl der Kanten mit einem niedrigeren Kantenwert im Probiotikum-Graph waren. Für den Fall, dass keine Effekte der Probiotikum-Fütterung auftraten, war ein *score*-Wert von 0 zu erwarten. Eine mögliche Division durch 0 stellte bei der Berechnung der Funktion nur ein theoretisches Problem dar, dass bei der Berechnung mit R keine Bedeutung hatte, da die Programmiersprache R die Division einer positiven Zahl durch 0 mit  $+\infty$  ersetzte.

Die Signifikanz der *score*-Werte wurde mit der Anwendung eines Monte-Carlo-Permutationstests berechnet. Für jedes Paar von Merkmalsgruppen wurden  $5 \times 10^4$  zufällige Permutation der Gruppenzugehörigkeiten erzeugt. Für jeden permutierten Datensatz wurde zwei neue Netzwerke berechnet. Der Anteil der Permutationen mit einem gleichen oder absolut höheren *score*-Wert als der ursprüngliche *score*-Wert wurde als P-Wert angenommen.

## 5.7.2. Ergebnisse der Netzwerkanalyse

Der Vergleich der beiden Korrelationsnetzwerke zeigte signifikante Unterschiede bei den Assoziationen zwischen 28 Paaren von Merkmalsgruppen (Tabelle 5.6). Die Fütterung des Probiotikums erhöhte dabei die Assoziation zwischen 14 Paaren von Merkmalsgruppen und verringerte die Assoziation zwischen 14 Paaren von Merkmalsgruppen. Alle signifikant veränderten Assoziationen mit einem P-Wert kleiner als 0,05 betrafen entweder die Reaktion des Immunsystems (*Is\**), das Verhalten der gastrointestinalen Mikrobiota (*Metab\**) oder die Zusammensetzung der gastrointestinalen Mikrobiota (*Otu\**). Im Folgenden wird auf drei bemerkenswerte Ergebnisse eingegangen.

### 5.7.2.1. Verringerte Assoziation zwischen den Merkmalsgruppen *Otulle* und *IsExprIlePP*

Die Fütterung des Probiotikums verringerte die Assoziation zwischen dem Vorkommen bakterieller OTUs im Ileum (*Otulle*) und der Expression immunrelevanter Gene in den Peyers-Plaques des Ileums (*IsExprIlePP*) ( $score = -0,88$  und  $P = 0,013$ ) (Tabelle 5.6). Von den insgesamt 170 Kanten zwischen den beiden Merkmalsgruppen hatten 120 Kanten einen höheren Kantenwert im Kontroll-Graphen als im Probiotikum-Graphen. In der Kontroll-Gruppe korrelierten die Vorkommen von *Enterococcus faecium* (*Efaecium\_ile*) und *Enterococcus faecalis* (*Efaecalis\_ile*) mit den Expressionen der immunrelevanten Gene deutlich negativ (Abbildung 5.6). In der Probiotikum-Gruppe war diese deutliche Korrelation sichtbar geringer.

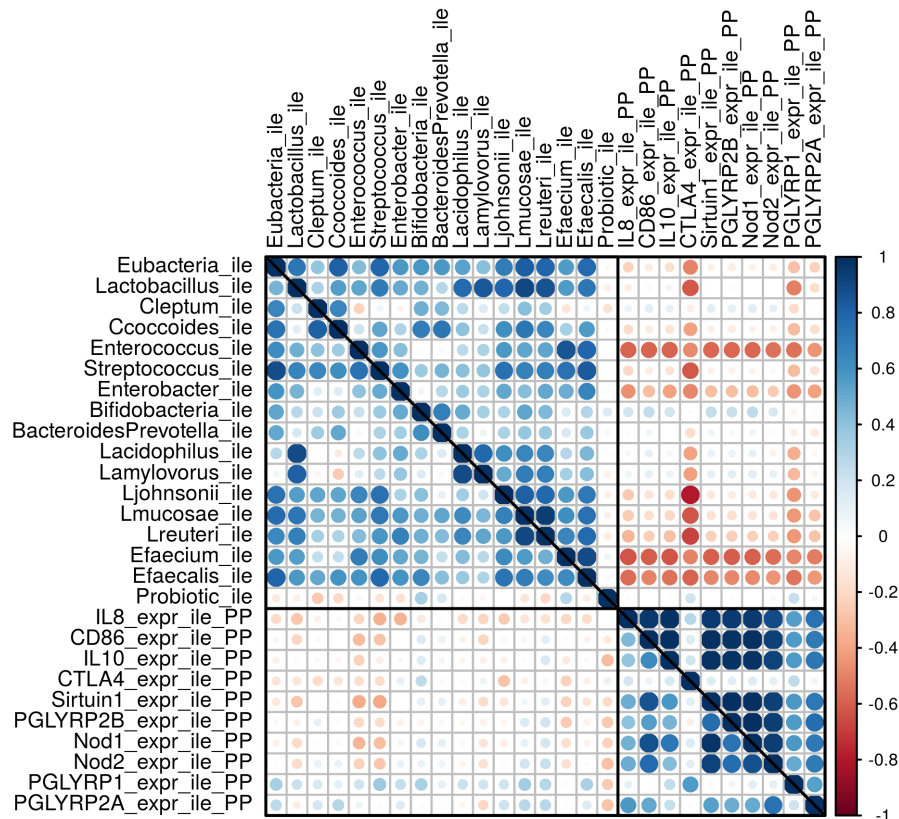
**Tabelle 5.6.:** Auswirkung der Probiotika-Fütterung auf den score-Wert zwischen den Merkmalsgruppen<sup>1</sup>.

Merkmalsgruppe 1	Merkmalsgruppe 2	#Kanten	#K > P	#K < P	score	P-Wert
<i>IsCytoBlo</i>	<i>EpiJej</i>	90	25	65	0,96	0,017*
<i>IsExprLiv</i>	<i>BloodCells</i>	22	3	19	1,85	0,005**
<i>IsExprLiv</i>	<i>EpiCol</i>	26	4	22	1,70	0,035*
<i>IsExprLiv</i>	<i>IsExprSpln</i>	14	1	13	2,56	0,022*
<i>IsExprJejLN</i>	<i>MarkerExpr</i>	49	35	14	-0,92	0,034*
<i>IsExprIlePP</i>	<i>EpiCol</i>	130	92	38	-0,88	0,005**
<i>IsExprJejPP</i>	<i>IsLymIleEL</i>	90	63	27	-0,85	0,026*
<i>IsExprJejPP</i>	<i>MorphJej</i>	30	24	6	-1,39	0,026*
<i>IsExprSpln</i>	<i>IsCytoBlo</i>	42	10	32	1,16	0,023*
<i>IsExprSpln</i>	<i>IsLymIleEL</i>	63	47	16	-1,08	0,006**
<i>IsExprSpln</i>	<i>IsLymJejIleEL</i>	98	19	79	1,43	0,002**
<i>IsExprSpln</i>	<i>Weight</i>	42	3	39	2,56	0,034*
<i>IsLymIleEL</i>	<i>MorphJej</i>	27	21	6	-1,25	0,026*
<i>MetabCol</i>	<i>IsCytoBlo</i>	84	26	58	0,80	0,025*
<i>MetabCol</i>	<i>MetabIle</i>	154	98	56	-0,56	0,037*
<i>MetabJej</i>	<i>BloodCells</i>	143	93	50	-0,62	0,037*
<i>MetabJej</i>	<i>Epille</i>	78	57	21	-1,00	0,045*
<i>MetabJej</i>	<i>IsLymIleJej</i>	182	123	59	-0,73	0,018*
<i>OtuCol</i>	<i>Epille</i>	102	32	70	0,78	0,017*
<i>OtuCol</i>	<i>IsExprJejLN</i>	119	43	76	0,57	0,014*
<i>OtuCol</i>	<i>IsExprSpln</i>	119	39	80	0,72	0,029*
<i>OtuCol</i>	<i>Otulle</i>	289	92	197	0,76	0,038*
<i>OtuCol</i>	<i>Weight</i>	102	25	77	1,12	0,045*
<i>Otulle</i>	<i>EpiCol</i>	221	153	68	-0,81	0,040*
<i>Otulle</i>	<i>IsExprIlePP</i>	170	120	50	-0,88	0,013*
<i>Otulle</i>	<i>MetabSto</i>	221	75	146	0,67	0,022*
<i>OtuSto</i>	<i>IsLymIlePP</i>	221	165	56	-1,08	0,042*
<i>OtuSto</i>	<i>MarkerExpr</i>	119	89	30	-1,09	0,007**

<sup>1</sup>Die dritte Spalte gibt die Anzahl der Kanten zwischen den beiden Merkmalsgruppen an. Die vierte Spalte gibt die Anzahl der Kanten an, die einen höheren Wert im Kontroll-Graphen hatten. Die fünfte Spalte gibt die Anzahl der Kanten an, die einen höheren Wert im Probiotikum-Graphen hatten. Die siebte Spalte gibt die P-Werte an; \* $P < 0,05$ ; \*\* $P < 0,01$

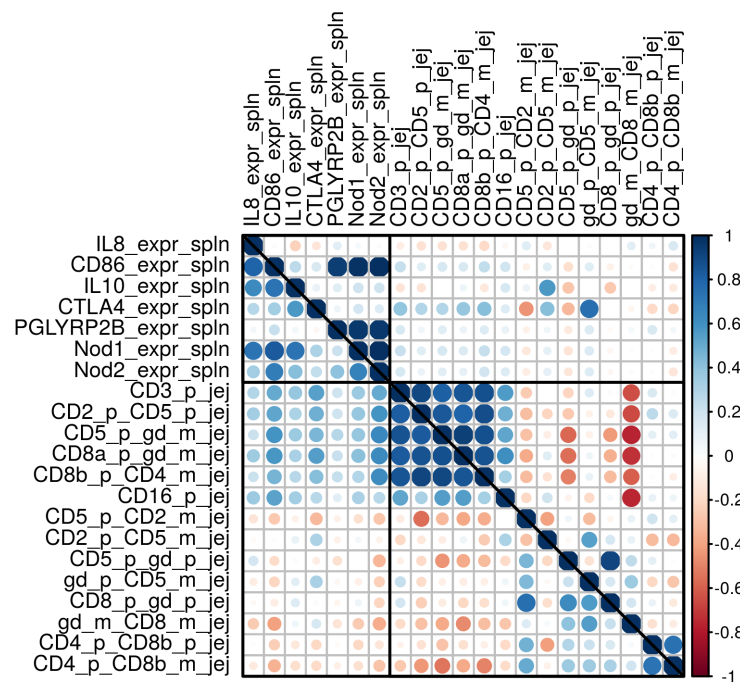
### 5.7.2.2. Erhöhte Assoziation zwischen den Merkmalsgruppen *IsExprSpln* und *IsLymJejIleEL*

Die Fütterung des Probiotikums erhöhte die Assoziation zwischen der Expression immunrelevanter Gene in der Milz (*IsExprSpln*) und der Zusammensetzung der IEL des Jejunums (*IsExprLymJejIleEL*) ( $score = 1,43$  und  $P = 0,002$ ) (Tabelle 5.6). Von den insgesamt 98 Kanten zwischen den beiden Merkmalsgruppen hatten 79 Kanten einen höheren Kantenwert im Probiotikum-Graphen als im Kontroll-Graphen (Abbildungen 5.7). Zum Beispiel korrelierte das Vorkommen



**Abbildung 5.6.:** Paarweise Korrelation zwischen den Merkmalen der Merkmalsgruppen *O-tulle* (\*\_ile) und *IsExprIlePP* (\*\_expr\_ile\_PP). Farbe der Kreise repräsentiert die Stärke der paarweisen Korrelationen. Die Diagonale (oben links nach unten rechts) teilt die Grafik in zwei Hälften. Obere rechte Hälfte der Matrix repräsentiert die Korrelationen mit Daten aus der Kontroll-Gruppe und untere linke Hälfte der Matrix repräsentiert die Korrelationen mit Daten aus der Probiotikum-Gruppe.

von  $CD3^+$  IEL (*CD3\_p\_jej*) stärker mit der Expression der sieben immunrelevanten Gene in der Probiotikum-Gruppe als in der Kontroll-Gruppe. Ebenfalls korrelierte das Vorkommen von  $CD8\alpha^+/\gamma\delta^-$  IEL (*CD8a\_p\_gd\_m\_jej*) stärker mit der Expression der sieben immunrelevanten Gene in der Probiotikum-Gruppe als in der Kontroll-Gruppe.



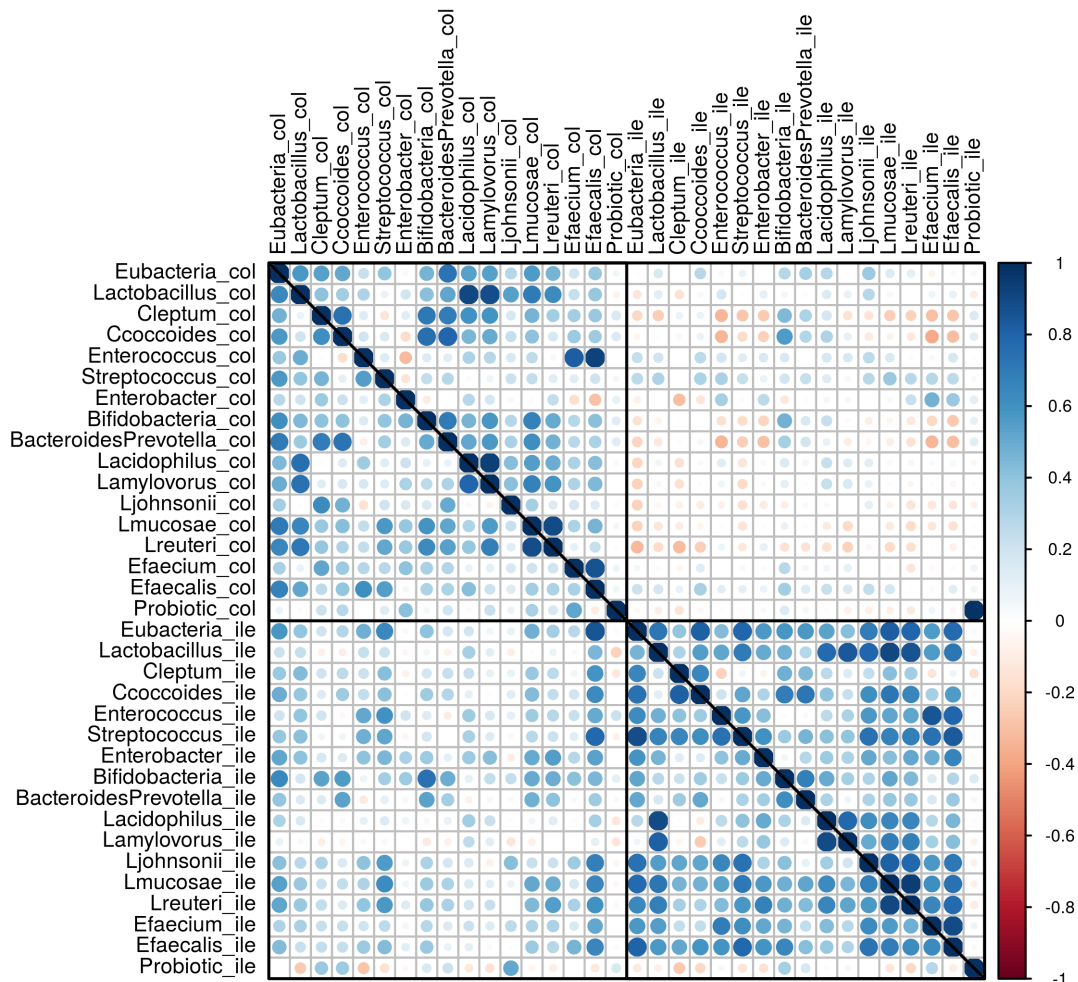
**Abbildung 5.7.:** Paarweise Korrelation zwischen den Merkmalen der Merkmalsgruppen *IsExprSpln* (\*\_expr\_spln) und *IsExprLymJejIEL* (\*\_jej). Farbe der Kreise repräsentiert die Stärke der paarweisen Korrelationen. Die Diagonale (oben links nach unten rechts) teilt die Grafik in zwei Hälften. Obere rechte Hälfte der Matrix repräsentiert die Korrelationen mit Daten aus der Kontroll-Gruppe und untere linke Hälfte der Matrix repräsentiert die Korrelationen mit Daten aus der Probiotikum-Gruppe.

### 5.7.2.3. Erhöhte Assoziation zwischen den Merkmalsgruppen *Otulle* und *OtuCol*

Die Fütterung des Probiotikums erhöhte die Assoziation zwischen dem Vorkommen bakterieller OTUs im Ileum (*Otulle*) und dem Vorkommen bakterieller OTUs im Colon (*OtuCol*) ( $score = 0,76$  und  $P = 0,038$ ) (Tabelle 5.6). Von den insgesamt 289 Kanten zwischen den beiden Merkmalsgruppen hatten 197 Kanten einen höheren Kantenwert im Probiotikum-Graphen als im Kontroll-Graphen. In der Kontroll-Gruppe traten hohe positive Korrelationen vor allem zwischen dem Vorkommen bakterieller OTUs innerhalb des Colons und innerhalb des Ileums auf, während die Vorkommen der Bakterien zwischen den beiden Darmabschnitten kaum und teilweise sogar negativ korrelierten (Abbildung 5.8). In der Probiotikum-Gruppe war die Korrelation der Vorkommen bakterieller OTUs zwischen den beiden Darmabschnitten im Vergleich zur Kontroll-Gruppe deutlich erhöht. Die Vorkommen der beiden Spezies *Lactobacillus reuteri* und *Lactobacillus mucosae* zeigten sowohl im Colon (*Lreutrei\_col* und *Lmucosae\_col*) als auch im Ileum (*Lreutrei\_ile* und *Lmucosae\_ile*) eine hohe positive Korrelation, während die Korrelationen zwischen den beiden Darmabschnitten in der Kontroll-Gruppe relativ gering waren. Die Fütterung des Probiotikums



erhöhte die Korrelationen der Vorkommen von *Lactobacillus reuteri* und *Lactobacillus mucosae* zwischen Ileum und Colon. Ein weiteres Beispiel für eine erhöhte Korrelation zwischen Ileum und Colon durch die Probiotikum-Fütterung ist die Korrelation des Vorkommens von Bifidobakterien im Colon (*Bifidobacteria\_col*) und im Ileum (*Bifidobacteria\_ile*).



**Abbildung 5.8.:** Paarweise Korrelation zwischen den Merkmalen der Merkmalsgruppen *Otu-Col* (\*\_col) und *Otu-Ile* (\*\_ile). Farbe der Kreise repräsentiert die Stärke der paarweisen Korrelationen. Die Diagonale (oben links nach unten rechts) teilt die Grafik in zwei Hälften. Obere rechte Hälfte der Matrix repräsentiert die Korrelationen mit Daten aus der Kontroll-Gruppe und untere linke Hälfte der Matrix repräsentiert die Korrelationen mit Daten aus der Probiotikum-Gruppe.



## 6. Diskussion

In der vorliegenden Arbeit wurde die Piggeldat-Datenbank zur Integration und zur Verwaltung von Messdaten aus Fütterungsversuchen des Sonderforschungsbereichs 852 erfolgreich implementiert. Die Piggeldat-Datenbank stellt ein umfangreiches Datenarchiv zur Auswirkung der beiden Nahrungsfaktoren *Enterococcus faecium* NCIMB 10415 und Zink bei Ferkeln dar und die Sammlung der Daten diente als Basis für die Durchführung von vielfältige Datenanalysen. Durch die Integration der Daten konnte insbesondere die Auswirkungen des Nahrungsfaktors *Enterococcus faecium* NCIMB 10415 bei Ferkeln genauer untersucht und beschrieben werden. Zur Implementierung der Datenbank erfolgte die Programmierung der Datenbankanwendung Catria, welche für den Aufbau von Datenbanken in zukünftigen Projekten zur freien Verfügung steht.

### 6.1. Die Datenbankanwendung Catria

Die Datenbankanwendung Catria wurde im Rahmen dieser Arbeit zur Verwaltung von Forschungsdaten aus Fütterungsversuchen entwickelt. Die Implementierung von Catria basierte dabei auf der Formulierung von vier Anwendungsfällen (Abschnitt 3.1.1.). Catria erfüllt die vier Anwendungsfälle Nutzerverwaltung, Datenintegration, Datenverwaltung und Datenansicht. Durch die Verwendung des CMS Drupal bietet Catria ein Nutzersystem an, welches die Registrierung und die Authentifizierung der Nutzer unterstützt. Die Datenintegration erlaubt die Speicherung beliebiger Messwerte in der Catria-Datenbank. Nach der Datenintegration können die Nutzer von Catria bestimmen, in welchem Umfang andere Nutzer auf die Messdaten zugreifen können. Die Catria-Website stellt die integrierten Forschungsdaten in übersichtlichen Tabellen dar. Des Weiteren bietet Catria den Export von Messdaten in eine CVS-Datei an. Hierdurch kann ein Nutzer die integrierten Messdaten auf einem lokalen Computer analysieren.

#### 6.1.1. Vorteile von *Open-Access* bei Fütterungsversuchen

Eine Catria-Installation kann über die zugehörige Datenbank-Website einen offenen Zugriff auf die eigenen Forschungsdaten erlauben. Generell wird der freie Zugriff auf Forschungsdaten mit dem Begriff *Open-Access* bezeichnet. Die "Berliner Erklärung über den offenen Zugang zu wissenschaftlichem Wissen" fordert den freien Zugriff auf Forschungsdaten aus öffentlichen Einrichtungen<sup>[11]</sup>. Diese Erklärung wird von vielen Wissenschaftlern und Instituten unterstützt (Klump et al. 2006<sup>[57]</sup>). *Open-Access* erhöht die Transparenz und die Reproduzierbarkeit von wissenschaftlichen Arbeiten (Molloy 2011<sup>[71]</sup>). Insbesondere bei der Untersuchung von Nahrungsfaktoren kann eine offene Publikation von Forschungsdaten zur Durchführung von Meta-Analysen beitragen, welche die Ergebnisse aus verschiedenen Fütterungsversuchen zusammenfassen (Sauvant et al. 2008<sup>[89]</sup>). Da Tierversuche generell kritisch zu betrachten sind, empfiehlt die *European Science Foundation* die effizientere Durchführung von Tierversuchen durch die Anerkennung der *three Rs* (*reduction, refinement* und *replacement*)<sup>[29]</sup>. Die offene Publikation von Forschungsdaten kann zur Reduzierung von Tierversuchen beitragen und eine effizientere Nutzung der Forschungsdaten

bewirken. Somit kann die Verwendung von Catria zur offenen Publikation von Forschungsdaten aus Fütterungsversuchen einen wichtigen Beitrag zur effizienteren Durchführung von Tierversuchen leisten.

### 6.1.2. Verwendung eines Open-Source-Content-Management-Systems

Catria basiert auf dem Open-Source-CMS Drupal. Im Allgemeinen dient ein CMS zum Aufbau und zur Verwaltung von Websites. Die Verwendung von CMS zur Verwaltung von biologischen Forschungsdaten ist in den letzten Jahren gestiegen und ein CMS kann bereits in vielen Fällen eine Alternative zur Verwendung eines kommerziellen *laboratory information management systems* (LIMS) darstellen (Mooney und Baenziger 2007<sup>[73]</sup>, Schmitt und Burchinal 2011<sup>[91]</sup>). Die Vorteile eines CMS bestehen darin, dass dieses bereits viele der benötigten Funktionalitäten implementiert und sich durch spezifische Erweiterungen an die eigenen Bedürfnisse anpassen lässt. Ein weiterer Vorteil eines CMS besteht darin, dass dieses sich in der Regel über lokale Server der Institute oder der universitären Rechenzentren betreiben lässt. Umfragen haben gezeigt, dass Wissenschaftler universitäre Datenbanken bei der Speicherung von Forschungsdaten bevorzugen (Simukovic et al. 2013<sup>[97]</sup>). Die Verwendung einer lokalen Datenbank kann somit die Bereitschaft zum geregelten Austausch von Forschungsdaten erhöhen. Der stetige Anstieg der Datenproduktion in den Biowissenschaften führt zu einer zunehmenden Anzahl an Forschungsprojekten, bei denen unterschiedliche Institute beteiligt sind. Diese Projekte benötigen die Implementierung von Infrastrukturen zur Verwaltung von Messdaten (Marx 2013<sup>[69]</sup>) und die Verwendung eines CMS kann ein Teil einer solchen Datenverwaltungsinfrastruktur darstellen. Durch die Verwendung einer Website zur Verwaltung von Forschungsdaten lässt sich die Kooperation zwischen unterschiedlichen Instituten ohne die Installation von Zusatzsoftware erreichen.

Ein Vorteil von Open-Source-Software besteht darin, dass sich durch die Veröffentlichung des Programmcodes Gemeinschaften von Entwicklern aufbauen lassen, welche neue Funktionen implementieren, den Code verbessern und auf Schwachstellen aufmerksam machen (Heron et al. 2013<sup>[42]</sup>). Dabei ist die Sicherheit von Software aus aktiven Open-Source-Projekten vergleichbar mit der Sicherheit von kommerzieller Software (Schryen 2011<sup>[93]</sup>). Bei der Auswahl eines Open-Source-CMS sind eine aktive Entwickler-Gemeinschaft und regelmäßige Aktualisierungen des Programmcodes für die Sicherheit der Anwendung entscheidend (Meike et al. 2009<sup>[70]</sup>).

### 6.1.3. Drupal zur Verwaltung von Forschungsdaten

Catria profitiert aus verschiedenen Gründen von der Verwendung von Drupal. Der Schwerpunkt von Drupal liegt beim Aufbau von Websites zur verteilten Verarbeitung von Informationen. Hierfür bietet Drupal standardmäßig ein Nutzer-Management-System an und es unterstützt einige Standard-Datentypen wie zum Beispiel Artikel oder Blog-Einträge. Drupal ist weltweit das am dritt-meisten verwendete CMS, wobei ca. 2% aller Websites Drupal nutzen<sup>[114]</sup>. Einige Stärken von Drupal sind die Programmierung von Drupal als Open-Source-Software, eine große und aktive Entwickler-Gemeinschaft und die Möglichkeit eine Drupal-Installation über eine Vielzahl möglicher Module an die eigenen Bedürfnisse anzupassen (Das et al. 2009<sup>[26]</sup>, Papanicolaou und Heckel 2010<sup>[78]</sup>, Hubble et al. 2011<sup>[45]</sup>, Sinha et al. 2011<sup>[98]</sup>, Iwanaga et al. 2013<sup>[50]</sup>). Die aktive Entwickler-Gemeinschaft validiert den Programmcode von neuen Drupal-Projekten und sorgt dafür, dass Schwachstellen gefunden und behoben werden. Beim offiziellen Drupal-Projekt Catria erfolgt somit auch eine fortlaufende Validierung des Programmcodes durch die Entwickler-Gemeinschaft. Dieses Vorgehen erhöht insbesondere die Sicherheit von Catria.

Verschiedene Projekte, zum Beispiel aus dem Fachgebiet der Ökologie, verwenden Drupal zur Publikation und zur Verwaltung von Forschungsdaten (Smith et al. 2009<sup>[99]</sup> und Drupal Ecological Information Management System 2014<sup>[88]</sup>). Das auf Drupal basierende Programm eXframe bietet die Verwaltung und die Publikation von Forschungsdaten aus dem Bereich der Genetik an (Sinha et al. 2011<sup>[98]</sup>). Das Projekt *Generic Model Organism Database* (GMOD) bietet drei Drupal-Module an, mit denen sich bioinformatische Analysen auf einer Drupal-Website durchführen lassen (Papanicolaou und Heckel 2010<sup>[78]</sup>). Catria profitiert hierdurch von der möglichen Erweiterung mit den vielen verfügbaren Drupal-Modulen.

### 6.1.4. Das Catria-Datenmodell

Catria implementiert ein maßgeschneidertes Datenmodell zur Verwaltung von Forschungsdaten aus Fütterungsversuchen innerhalb von Drupal. Das Catria-Datenmodell kann dabei Daten aus Tierversuchen erfassen, bei denen die einzelnen Tiere die Versuchsobjekte darstellen. Mit der Speicherung von experimentellen Variablen lassen sich komplexe Aufbauten von Fütterungsversuchen in der Datenbank modellieren. Dabei kann die Datenbank beliebige Informationen zu den Tieren, zu den Messungen und zu den gemessenen Merkmalen erfassen. Die Erfassung von Informationen über das Datenbank-Attribut *beschreibung* kann dabei in beliebigen Formaten erfolgen; zum Beispiel in dem gängigen XML-Format. Somit lassen sich mit dem Catria-Datenmodell die Metadaten aus Fütterungsversuchen auch nach den zwei Standards zur Datenspeicherung MIBBI (Taylor et al. 2008<sup>[105]</sup>) und DELSA (Kolker et al. 2014<sup>[58]</sup>) erfassen.

Wie jedes Datenmodell stellt das Catria-Datenmodell eine Abstrahierung der realen Gegebenheiten dar, das nicht alle möglichen Durchführungen von Fütterungsversuchen in der Forschung erfassen kann. Ein kritischer Punkt hierbei ist die Abgrenzung der Fütterungsversuche voneinander. Hierdurch ist es zum Beispiel nicht möglich, ein Teil der Tiere in zwei unterschiedlichen Fütterungsversuchen zu betrachten. Diese Einschränkung lässt sich umgehen, indem man die einzelnen Tiere in der Datenbank mehrfach speichert und dann den verschiedenen Fütterungsversuchen zuordnet. Des Weiteren können mit dem Catria-Datenmodell keine Studien erfasst werden, bei denen Populationen die Versuchseinheiten darstellen. Das relativ spezifische Anwendungsspektrum des Catria-Datenmodells beschränkt somit die Verwendung von Catria auf wenige spezielle Forschungsprojekte. Nichtsdestotrotz liefert das Catria-Datenmodell einen wichtigen Ansatz zur standardisierten Publikation von Forschungsdaten aus Fütterungsversuchen. Ein solcher Ansatz ist bisher nicht verbreitet. Bei Bedarf kann das Modell erweitert werden und es lassen sich Variationen des Datenmodells erstellen um weitere Anwendungsfälle zu erfassen. Durch die Programmierung als Open-Source Anwendung lassen sich gewünschte Änderungen des Datenmodells direkt im Programmcode vornehmen.

### 6.1.5. Vergleich von Catria mit *large-scale*-Systemen

Es existieren zahlreiche Programme zur Verwaltung von Forschungsdaten in großangelegten (engl. *large-scale*) Forschungsprojekten (Schmitt und Burchinal 2011<sup>[91]</sup>, Wruck et al. 2014<sup>[123]</sup>). Diese teilweise sehr komplexen Programme verfügen im Allgemeinen über eine abstrakte Datenmodellierung und können viele verschiedene Anwendungsfälle erfassen. Die Komplexität dieser Programme und derer Datenmodelle erschwert dabei die Anpassung der Programme an spezifische Anforderungen. Die Datenverwaltung in *large-scale*-Forschungsprojekten erfordert die Sammlung, die Integration und die Verteilung von Forschungsdaten (Wruck et al. 2014<sup>[123]</sup>). Catria kann diese Aufgaben ebenfalls erfüllen. Ein Nachteil von Catria besteht allerdings in der

fehlenden Automatisierung der Datenintegration. Das Programm openBIS kann zum Beispiel durch die Verwendung von Online-Verzeichnissen (engl. *dropbox*) eine automatische Integration von Messdaten in das System realisieren (Bauch et al. 2011<sup>[7]</sup>). Bei Catria müssen die Nutzer ihre Daten manuell integrieren. Ein Vorteil von Catria besteht darin, dass die Nutzer die volle Kontrolle auf den Zugriff der eigenen Daten haben, während die Messdaten bei Programmen wie openBIS innerhalb von Projekten automatisch freigegeben werden. Ein weiterer Vorteil von Catria besteht darin, dass Catria ein Teil eines großen und aktiven Open-Source Projekts ist. Unabhängige Anwendungen, die von einzelnen Instituten entwickelt und gepflegt werden, verfügen in der Regel nicht über eine breite Entwickler-Gemeinschaft.

## 6.2. Die Piggeldat-Datenbank

Der Sonderforschungsbereich 852 nutzte die Piggeldat-Datenbank zur Integration und zur Verwaltung der Forschungsdaten aus sieben verschiedenen Fütterungsversuchen mit Ferkeln. Für die Integration der Forschungsdaten erfolgte die Einführung und die Anwendung von einheitlichen Datenformaten und Bezeichnungen für die Ferkel, die Messungen und die gemessenen Merkmale. Durch die Menge und die Diversität der integrierten Forschungsdaten stellt die Piggeldat-Datenbank ein umfangreiches Datenarchiv für das Zusammenspiel zwischen Nahrungsfaktoren, der gastrointestinalen Mikrobiota und dem Wirt beim Schwein dar. Ein solches Datenarchiv ist in dieser Form bisher einmalig und kann einen wichtigen Beitrag zur Wissenschaft über die Grenzen des SFB 852 hinaus leisten.

### 6.2.1. Die Piggeldat-Datenbank als Fundament zur Datenanalyse

Die Piggeldat-Datenbank diente als Fundament zur Analyse der Auswirkungen der beiden Nahrungsfaktoren *E. faecium* NCIMB 10415 und Zink beim Ferkel. Die Sammlung der Daten in der zentralen Datenbank diente dabei als Grundlage zur Durchführung von verschiedenen Datenanalysen aus dem Bereich der Bioinformatik und aus dem Bereich der Statistik:

- Die Berechnung eines 2-faktoriellen linearen Mixed-Modells (Altersgruppe × Behandlungsgruppe) mit verschiedenen Messdaten zeigte die erhöhte Konsistenz der Fäkalien, eine erhöhte TGEV-spezifische Immunantwort und die verringerte Expression der Gene *IFN-α* und *OAS* im Gastrointestinaltrakt von Ferkeln durch die Fütterung einer hohen Zink-Konzentration bei einer Infektion der Ferkel mit TGEV (Chai et al. 2014<sup>[18]</sup>).
- Die Modellierung von Ussing-Kammer-Experimenten mit einem 4-faktoriellen linearen Mixed-Modell (Altersgruppe × Behandlungsgruppe × Experiment × Zink-Zugabe) zeigte verringerte Transport-Eigenschaften des Darmepithels von Ferkeln durch die direkte Zink-Zugabe (Gefeller et al. 2014<sup>[35]</sup>).
- Die Modellierung von Ussing-Kammer-Experimenten mit einem 3-faktoriellen linearen Mixed-Modell (Altersgruppe × Behandlungsgruppe × Experiment) zeigte erhöhte Transport-Eigenschaften und verbesserte Barrieren-Funktion des Darmepithels von Ferkeln durch die Fütterung mit dem Probiotikum *Enterococcus faecium* NCIMB 10415 (Klingspor et al. 2013<sup>[56]</sup>).
- Der Aufbau einer lokalen Sequenzdatenbank und der Einsatz von Programmen zur Sequenzanalyse unterstützte bei der Etablierung eines Protokolls zur simultanen Detektion von DNA-

und RNA-Viren aus den Fäkalien von Ferkeln (Sachsenröder et al. 2012<sup>[86]</sup>). Die Ergebnisse dieser Analyse lieferten eine Beschreibung des viralen Metagenoms im Gastrointestinaltrakt von Ferkeln und es konnte hierbei ein bis dahin unbekanntes Virus im Kot der Ferkel identifiziert werden.

- Die Analyse von Metagenom-Datensätzen resultierte in der Beschreibung der Variabilität des viralen Metagenoms im Gastrointestinaltrakt von Ferkeln und identifizierte deutliche Unterschiede zu dem viralen Metagenom im Gastrointestinaltrakt von Sauen (Sachsenröder et al. 2014<sup>[87]</sup>).
- Die Berechnung eines 2-faktoriellen linearen Modells (Altersgruppe × Behandlungsgruppe) mit Daten aus Genexpressionsanalysen zeigte die veränderte Expression von pro- und anti-inflammatorischen Zytokinen im Gastrointestinaltrakt von Ferkeln durch die Fütterung des Probiotikums *Enterococcus faecium* NCIMB 10415 (Siepert et al. 2014<sup>[95]</sup>).
- Die Integration diverser Datensätze diente zur explorativen Dataming-Analyse um die Auswirkung des Probiotikums *E. faecium* NCIMB 10415 zu beschreiben (Kapitel 5 und Twardziok et al. 2014<sup>[109]</sup>).
- Die Berechnung eines 3-faktoriellen linearen Mixed-Modells (Altersgruppe × Behandlungsgruppe × Impfung) mit verschiedenen Messdaten zeigte erhöhte Gewichtszunahme und eine erhöhte Anzahl an Virus-spezifischen Antikörpern im Blut von Ferkeln durch die Fütterung mit dem Probiotikum *Enterococcus faecium* NCIMB 10415 und durch die Fütterung von hohen Zink-Konzentrationen bei einer Infektion der Ferkel mit Schweineinfluenza (Wang et al. 2014<sup>[115]</sup>).

### 6.2.2. Die Piggeldat-Datenbank als Datenarchiv

Die ausführliche Dokumentation der integrierten Forschungsdaten unterstützt die weitere Verwendung der Piggeldat-Datenbank als Fundament zur Untersuchung der Auswirkungen von Nahrungsfaktoren bei Ferkeln. Der Aufbau eines Datenarchivs stellt den ersten entscheidenden Schritt für die Durchführung einer Data-Mining-Analyse dar (Runkler 2010<sup>[85]</sup>[Seite 1ff]). Zum Zeitpunkt der Anfertigung der vorliegenden Arbeit bietet die Piggeldat-Datenbank den Zugriff auf die Forschungsdaten aus insgesamt 17 wissenschaftlichen Veröffentlichungen an (Tabelle D.1, Bednorz et al. 2013<sup>[8]</sup>, Bednorz et al. 2013<sup>[9]</sup>, Chai et al. 2014<sup>[17]</sup>, Chai et al. 2014<sup>[18]</sup>, Gefeller et al. 2014<sup>[35]</sup>, Janczyk et al. 2013<sup>[51]</sup>, Klingspor et al. 2013<sup>[56]</sup>, Kreuzer et al. 2012<sup>[59]</sup>, Liu et al. 2014<sup>[65]</sup>, Martin et al. 2012<sup>[67]</sup>, Martin et al. 2013<sup>[68]</sup>, Richter et al. 2011<sup>[83]</sup>, Scharek-Tedin et al. 2015<sup>[90]</sup>, Siepert et al. 2014<sup>[95]</sup>, Starke et al. 2013<sup>[101]</sup>, Starke et al. 2013<sup>[102]</sup> und Wang et al. 2014<sup>[115]</sup>). Die Mitglieder des SFB 852 haben unbeschränkten Zugriff auf die integrierten Datensätze der Piggeldat-Datenbank. Durch die Dokumentation der Datensätze lassen sich die durchgeführten Messungen und die Verarbeitung der Daten nachvollziehen. Die Anwendung von Data-Mining-Werkzeugen mit den integrierten Datensätzen kann zur Analyse von Wechselwirkungen zwischen den Nahrungsfaktoren, der gastrointestinalen Mikrobiota und dem Wirt beitragen. Die Messdaten können auf Anfrage an externe Wissenschaftler freigegeben werden und somit kann die Datenbank zur weiteren Erforschung der Auswirkungen von Nahrungsfaktoren beim Ferkel beitragen. Durch diese dokumentierte Archivierung der Forschungsdaten wird die weitere Verwendung der Daten aus den Versuchstieren sichergestellt und eine optimale Nutzung ermöglicht.

### 6.2.3. Verwendung der Piggeldat-Datenbank im SFB 852

Bei der Verwendung der Datenbankanwendung Catria zum Aufbau der Piggeldat-Datenbank wurden einige Schwierigkeiten beim Aufbau eines zentralen Datenarchivs ersichtlich. Die Integration von Forschungsdaten aus einem Forschungsprojekt in eine zentrale Datenbank erfordert die strikte Einhaltung von einheitlichen Bezeichnungen und Formaten. Die Festlegung der Daten-Standards für die Piggeldat erfolgte teilweise während der Entwicklung der Catria-Datenbankanwendung in einer späten Phase des SFB 852. Dies lag vor allem daran, dass die Implementation der Datenbankanwendung parallel zum Forschungsprojekt verlief. Für künftige Projekte wird empfohlen, alle benötigten Daten-Standards und das verwendete Datenmodell bereits vor dem Beginn eines Forschungsprojekts festzulegen. Hierbei kann das Catria-Datenmodell als Richtlinie für zukünftige Fütterungsversuche dienen.

## 6.3. Auswirkung der Fütterung des Probiotikums *E. faecium* NCIMB 10415 beim Schwein

Die Ergebnisse der Data-Mining-Analyse mit den integrierten Datensätzen aus dem *EFFT2010*-Fütterungsversuch stimmen mit vorherigen Untersuchungen überein und zeigen mögliche Auswirkungen der Fütterung des Probiotikums auf das Zusammenspiel zwischen der gastrointestinalen Mikrobiota und dem Wirt. Insgesamt war die Auswirkung der Fütterung des Probiotikums auf die Gesamtheit der gemessenen Merkmale relativ gering. Die Hauptkomponentenanalyse zeigte, dass das Alter der Ferkel den stärksten Einfluss auf die gemessenen Merkmale hatte. Nichtsdestotrotz zeigte auch die Fütterung des Probiotikums einen signifikanten Effekt auf die Gesamtheit der gemessenen Merkmale (Tabelle 5.2). Bei den 26 Tage alten Ferkeln und 34 Tage alten Ferkeln konnte der deutlichste Unterschied zwischen den Behandlungsgruppen beobachtet werden (Abbildung 5.5). Dabei unterschieden sich die Auswirkungen des Probiotikums zwischen den vier Messzeitpunkten (Tabelle 5.4). Die Gruppierung der gemessenen Merkmale in Merkmalsgruppen diente zur Identifikation von Merkmalsgruppen, in denen besonders große Effekte der Fütterung auftraten. Die identifizierten Merkmalsgruppen standen hauptsächlich mit dem Verhalten und der Zusammensetzung der gastrointestinalen Mikrobiota und mit der Reaktion des Immunsystems im Zusammenhang (Tabelle 5.5). Die Untersuchung der Korrelationen zwischen den gemessenen Merkmalen zeigte veränderte Zusammenhänge zwischen der Aktivität und der Zusammensetzung der gastrointestinalen Mikrobiota und der Reaktion des Immunsystems durch die Probiotikum-Fütterung (Tabelle 5.6).

Das Probiotikum *Enterococcus faecium* wird unter anderem bei der Ernährung von Schweinen in der Landwirtschaft verwendet. Die Auswirkung der Fütterung von *Enterococcus faecium* auf die Leistungsfähigkeit von Ferkeln ist uneindeutig. Die tägliche Verabreichung von *E. faecium* NCIMB 10415 kann die Gewichtszunahme eines Ferkels erhöhen (Zeyner et al. 2006<sup>[127]</sup> und Wang et al. 2014<sup>[115]</sup>). In anderen Studien wurde keine Auswirkung der Fütterung von *Enterococcus faecium* auf das Körpergewicht beobachtet (Broom et al.<sup>[13]</sup> und Taras et al. 2006<sup>[104]</sup>) beziehungsweise sogar eine negative Auswirkung auf die Gewichtszunahme von Ferkeln beobachtet (Kreuzer et al. 2012<sup>[59]</sup>). Die Fütterung von *Enterococcus faecium* kann das Vorkommen von Durchfallerkrankungen verringern (Taras et al. 2006<sup>[104]</sup> und Zeyner et al. 2006<sup>[127]</sup>). In Zellkulturen-Experimenten konnte eine höhere Lebensfähigkeit von Zellen durch die Behandlung mit *E. faecium* NCIMB 10415 im Zusammenhang mit einer Infektion mit dem *transmissible gastroenteritis virus* (TGEV) und dem *swine influenza virus* (SwIV) nachgewiesen werden (Chai et al. 2013<sup>[16]</sup> und Wang et al.



2013<sup>[116]</sup>). Die Fütterung des Probiotikums an Ferkel kann die Produktion von Virus-spezifischen Antikörpern nach einer Infektion mit SwIV erhöhen (Wang et al. 2014<sup>[115]</sup>).

Verschiedene Studien mit den Messdaten aus dem *EFFT2010*-Fütterungsversuch konnten unterschiedliche Effekte der Fütterung des Probiotikums an Ferkel nachweisen: eine erhöhte Transport-Funktionen und verbesserte Barrieren-Funktion der Darmschleimhaut (Klingspor et al. 2013<sup>[56]</sup>), eine verringerte Futtermittelverwertung (Martin et al. 2012<sup>[67]</sup>), eine Aktivierung von zytotoxischen T-Zellen (Kreuzer et al. 2012<sup>[60]</sup>), ein verändertes Vorkommen von pathogenen Genen der gastrointestinalen *E. coli* Gemeinschaft (Bednorz et al. 2013<sup>[8]</sup>) und eine veränderte Genexpression von Zytokinen (Siepert et al. 2014<sup>[95]</sup>).

### 6.3.1. Durchführung der Datenverarbeitung

Die Integration der Messdaten aus dem *EFFT2010*-Fütterungsversuch wurde über die Piggeldat Datenbank durchgeführt. Eine Schwierigkeit bei der anschließenden Datenverarbeitung lag bei der hohen Rate an fehlenden Werten in den Datenmatrizen. Viele Messungen wurden nur an einem Teil der 62 Ferkel durchgeführt. Nach der Entfernung der Merkmale mit über 50 % fehlenden Werten erfolgte die Schätzung der restlichen fehlenden Werte mit dem *regularized iterative PCA* Algorithmus. Dieser Algorithmus berücksichtigt sowohl die Zusammenhänge zwischen den Objekten als auch zwischen den Merkmalen und stellt eine Verbesserung zur Schätzung von fehlenden Werte durch die Mittelwerte der Merkmale dar (Josse und Husson 2012<sup>[52]</sup>). Des Weiteren zeigt der Algorithmus eine höhere Stabilität als der NIPALS-Algorithmus, bei Datenmatrizen mit vielen fehlenden Werten (Josse und Husson 2012<sup>[52]</sup>). Die anschließende Normalisierung der Merkmale auf einen Mittelwert von 0 und einer Standardverteilung von 1 diente zur gleichen Gewichtung aller gemessenen Merkmale in den nachfolgenden Analysen.

### 6.3.2. Untersuchung der Varianz des Datensatzes

Die Durchführung der Hauptkomponenten Analyse erfolgte zur explorativen Untersuchung der Varianz der mehrdimensionalen Datenmatrix. Dabei konnte die Hauptkomponenten Analyse eine signifikante Auswirkung der Probiotikum-Fütterung auf die Gesamtheit der gemessenen Merkmale identifizieren. Die Probiotikum-Fütterung erklärte aber nur einen geringen Anteil der gesamten Varianz der Datenmatrix. Ein großer Anteil der Varianz in der Datenmatrix ließ sich durch die Aufteilung der Ferkel in die vier Altersgruppen erklären. Dabei entsprach die erste Hauptkomponente mit dem höchsten Eigenwert einer steigenden Aufteilung der Tiere nach dem Alter (Abbildung 5.3 **(a)**). Die drei weiteren Hauptkomponenten repräsentierten unterschiedliche Aufteilungen zwischen einzelnen Altersgruppen (Abbildung 5.3 **(b)** bis **(d)**). Die ersten vier Hauptkomponenten repräsentierten somit verschiedene Muster zur Entwicklung der Merkmale über die vier Zeitpunkte. Der hohe Anteil der Alterseffekte an der gesamten Varianz der Datenmatrix war zu erwarten, da Ferkel in den ersten Wochen eine hohe Entwicklung durchlaufen. Hierbei hat besonderes die Absetzung der Ferkel von den Sauen einen großen Effekt auf die Physiologie der Ferkel, wobei zwischen einer akuten Phase und einer Phase der Anpassung unterschieden werden kann (Montagne et al. 2007<sup>[72]</sup>). Die akute Phase ist von einer veränderten Morphologie und einer verringerten Aktivität von Verdauungsenzymen geprägt, während in der darauf folgenden Anpassungsphase eine Regeneration des Darmepithels eintritt. Die Ferkel des *EFFT2010*-Fütterungsversuchs zeigten ebenfalls nach der Absetzung unter anderem eine verkürzte Zotten-Länge und eine geringere Aktivität von Verdauungsenzymen (Martin et al. 2012<sup>[67]</sup>) und eine Veränderung der Transport-Eigenschaften des Darmepithels (Klingspor et al. 2013<sup>[56]</sup>). Die

weiteren Hauptkomponenten repräsentierten Variationen im Datensatz, die sich nicht über die Aufteilung der Ferkel in die Altersgruppen erklären ließen. Diese Variationen können zum Beispiel genetische Ursachen haben.

### 6.3.3. Vorhersage der Behandlungsgruppen und Berechnung der Auswirkung auf einzelne Merkmale

Die OPLS-Berechnung mit den kompletten Datenmatrizen erlaubte keine stabilen Vorhersagen der Behandlungsgruppen aus den gemessenen Merkmalen. Dieses Resultat lag vor allem an der relativ geringen Auswirkung der Probiotikum-Fütterung, an der geringen Zahl der verwendeten Tiere und an der hohen Rate an fehlenden Werten. Nichtsdestotrotz ließ die Reduzierung der Datenmatrizen nach den Variablen mit den höchsten VIP-Werten die stabile Vorhersage der Behandlungsgruppen der 26 Tage und 34 Tage alten Ferkel zu (Abbildung 5.5). Dieses Ergebnis deutet darauf hin, dass die Auswirkung der Probiotikum-Fütterung bei den 26 Tage und 34 Tage alten Ferkel am stärksten war.

Das Verfahren OPLS eignet sich insbesondere zur Analyse von Datenmatrizen, bei denen die Anzahl der Variablen die Anzahl der Objekte übersteigt ( $K > N$ ) und bei denen hohe Korrelationen zwischen den Variablen auftreten (Wold et al. 2001<sup>[121]</sup>, Hasegawa und Funatsu 2010<sup>[40]</sup>). Beides war hier der Fall. Mögliche nichtlineare Effekte der Probiotikum-Fütterung konnten durch diese Methode nicht identifiziert werden. Die stabile Anwendung von nichtlinearen Verfahren, wie zum Beispiel die Berechnung von neuronalen Netzwerken (Schneider und Wrede 1998<sup>[92]</sup>), erfordert in der Regel eine größere Anzahl an Versuchsobjekten. Mit den OPLS-Modellen konnten die VIP-Werte zur Beschreibung der Auswirkungen des Probiotikums auf einzelne Merkmale bestimmt werden. Die Tauglichkeit der VIP-Werte zur Charakterisierung der Auswirkung der Probiotikum-Fütterung zeigte sich unter anderem auch darin, dass die Selektion der Merkmale mit den höchsten VIP-Werten die Vorhersagekraft der OPLS-Modelle erhöhte (Tabelle 5.3).

### 6.3.4. Vergleich der Auswirkungen zwischen verschiedenen Zeitpunkten

Die Berechnung der paarweisen Korrelationen auf Basis der VIP-Werte der Merkmale zwischen den vier Zeitpunkten zeigten geringe aber signifikante Korrelationen zwischen nachfolgenden Zeitpunkten. Dies bedeutet, dass immerhin ein bemerkbarer Effekt der Fütterung des Probiotikums über nachfolgende Zeitpunkte bestehen blieb. Die geringe Höhe der Korrelationskoeffizienten zeigte aber auch, dass sich die Auswirkung der Fütterung über die vier Zeitpunkte deutlich änderte. Der Korrelationskoeffizient zwischen den 26 Tage alten Ferkeln und 54 Tage alten Ferkeln war sogar negativ. Diese Beobachtung stimmte damit überein, dass viele Effekte der Probiotikum-Fütterung nur zu einzelnen Zeitpunkten nachweisbar waren (siehe Kreuzer et al. 2012<sup>[60]</sup>, Bednorz et al. 2013<sup>[8]</sup>, Klingspor et al. 2013<sup>[56]</sup> und Siepert et al. 2014<sup>[95]</sup>). Generell lässt sich diese Beobachtung mit einer starken Ko-Entwicklung des Ferkels mit der gastrointestinalen Mikrobiota erklären. Die hohe Veränderung der gemessenen Merkmale über die vier Zeitpunkte wurde bereits durch die Berechnung der PCA deutlich (Abbildung 5.3). Die Fütterung des Probiotikums kann somit als eine Störung eines sehr komplexen Systems, bestehend aus gastrointestinaler Mikrobiota und Wirt, betrachtet werden, bei der sich der Zustand des Systems laufend verändert. Es macht Sinn, dass sich die Auswirkung einer solchen Störung ebenfalls ändert. Insgesamt deuten die Ergebnisse der Berechnungen darauf hin, dass die Auswirkungen der Fütterung des Probiotikums *E. faecium*

vom Entwicklungsstand des Ferkels abhängig sind.

### 6.3.5. Auswirkung auf Merkmalsgruppen

Durch die Analyse der Auswirkung des Probiotikums auf Merkmalsgruppen konnten mehrere Merkmalsgruppen identifiziert werden, die besonders viele beeinflusste Merkmale enthielten. Die Ergebnisse der Analyse stimmten mit vorherigen einzelnen Analysen der Messdaten überein. Das reduzierte Vorkommen von pathogenen Genen der gastrointestinalen *E. coli* Gemeinschaft wurde von Bednorz et al. 2013<sup>[8]</sup> publiziert. Die vorliegende Analyse identifizierte die reduzierte Genexpression in den Peyers-Plaques des Ileums von 34 Tage alten Ferkeln. Die reduzierte Expression der Gene IL-8, IL-10 and CTLA4 and CD86 wurde von Siepert et al. 2014<sup>[95]</sup> beschrieben und diskutiert. Die Ergebnisse der vorliegenden Analyse deuteten insgesamt auf ein verändertes Verhalten der Zellen in den Peyers-Plaques des Ileums hin. Die Auswirkung der Probiotikum-Fütterung auf das Vorkommen von einzelnen bakteriellen Strängen im Verdauungstrakt der Ferkel wurde von Starke et al. 2013<sup>[101]</sup> publiziert. Die Analyse der Merkmalsgruppen identifizierte nun die gastrointestinale Mikrobiota als eine der am stärksten beeinflussten Merkmalsgruppe innerhalb des gesamten Datensatzes.

Insgesamt konnte die Analyse nur bei wenigen Merkmalsgruppen eine Auswirkung der Probiotikum-Fütterung zeigen. Diese identifizierten Merkmalsgruppen standen hauptsächlich mit der gastrointestinalen Mikrobiota im Jejunum und im Ileum und der Reaktion des Immunsystems in Beziehung. Diese Beobachtung stimmt mit der früheren Beobachtungen überein, dass die Fütterung von *E. faecium* an Ferkel hauptsächlich die gastrointestinales Mikrobiota und die Reaktion des Immunsystems beeinflusst (Simon 2010<sup>[96]</sup>).

### 6.3.6. Auswirkung auf die Assoziation zwischen Merkmalsgruppen

Die Netzwerk-Analyse zeigte veränderte Assoziationen zwischen 28 verschiedenen Merkmalsgruppen (Tabelle 5.6). Von diesen Ergebnissen kann die weitere Untersuchung der Wirkungsweise des Probiotikums profitieren. Mehrere beeinflusste Assoziationen stehen mit der Struktur und mit dem Verhalten der gastrointestinalen Mikrobiota und der Reaktion des Immunsystems in Verbindung. Diese Beobachtung unterstützt die Vermutung, dass die Fütterung des Probiotikums nicht nur die gastrointestinale Mikrobiota sondern vielmehr die Wechselwirkung zwischen der gastrointestinalen Mikrobiota und dem Immunsystems beeinflusst (Simon 2010<sup>[96]</sup>).

Eine bemerkenswerte Beobachtung ist die reduzierte Assoziation zwischen dem Vorkommen bakterieller OTUs im Ileum und der Expression immunrelevanter Gene in den Peyer-Plaques des Ileums (Abbildung 5.6). Beide Merkmalsgruppen zeigten einen Einfluss der Probiotikum-Fütterung bei 34 Tage alten Ferkeln (Tabelle 5.5). Dieses Ergebnis lässt sich mit einer möglichen Stabilisierung der Mikrobiota durch die Fütterung des Probiotikums erklären, wodurch die Zusammensetzung der gastrointestinalen Mikrobiota weniger stark von der Reaktion des Immunsystems abhängt. Eine weitere mögliche Erklärung stellt die verringerte Reaktion des Immunsystems auf Veränderungen der Mikrobiota dar. Die erhöhte Korrelation zwischen der Zusammensetzung der IEL im Jejunum und der Expression immunrelevanter Gene in der Milz durch die Fütterung des Probiotikums deutet auf ein verändertes Verhalten des Immunsystems durch die Probiotikum-Fütterung hin (Abbildung 5.7).

Auf eine stabilisierende Wirkung des Probiotikums auf die Mikrobiota deutet der erhöhte Zusammenhang zwischen der Zusammensetzung bakterieller OTUs im Ileum und Colon hin (Abbildung 5.8). Eine Stabilisierung der gastrointestinalen Mikrobiota kann diese eventuell weniger

## *Diskussion*

anfällig für Pathogene machen. Nichtsdestotrotz betreffen die Messungen nur das Vorkommen von einigen bakteriellen OTUs und eine generelle Aussage über die Stabilisierung der gastrointestinalen Mikrobiota kann hier nicht getroffen werden. Es kann aber vermutet werden, dass die Stabilisierung der Mikrobiota eine der hauptsächlichen Auswirkungen des Probiotikums darstellt. Weitere ausführlichere *Metagenmics*-Studien sollten durchgeführt werden, um die stabilisierende Wirkung des Probiotikums auf die gastrointestinale Mikrobiota zu untersuchen.

## 7. Zusammenfassung

Die Auswirkung eines Nahrungsfaktors beim Hausschwein, *Sus scrofa domestica*, hängt von einem komplexen Zusammenspiel zwischen den Bestandteilen der Ernährung, der gastrointestinalen Mikrobiota und der Reaktion des Wirts ab. Die Untersuchung der Auswirkung eines Nahrungsfaktors profitiert hierdurch besonders von der Integration von Messdaten aus den Bereichen *Metagenomics*, *Metaproteomics*, *Metatranscriptomics* und *Metabolomics*. Der Sonderforschungsbereich 852 führte in den Jahren 2010 bis 2013 sieben Tierversuche durch, um die Auswirkung des Probiotikums *Enterococcus faecium* NCIMB 10415 und des Zusatzes von Zink auf die Entwicklung des Ferkels zu untersuchen. Die Projektgruppen des SFB 852 bestimmten dabei unter anderem die Zusammensetzung der gastrointestinalen Mikrobiota, die Expression immunrelevanter Gene, die Zusammensetzung von Immunzellpopulationen und physiologische Parameter.

Im Rahmen dieser Arbeit erfolgte die Installation und die Administration der Piggeldat-Datenbank zur Verwaltung und zur Archivierung der diversen Forschungsdaten des SFB 852. Die Piggeldat-Datenbank enthält die Metadaten von insgesamt 490 Ferkeln, 57 Messungen, 1244 Merkmalen und sie enthält insgesamt 81060 Messwerte. Damit bietet die Piggeldat-Datenbank den Zugriff auf die Forschungsdaten von 17 wissenschaftlichen Veröffentlichungen an. Die dokumentierten Daten sind innerhalb des Forschungsprojekts frei verfügbar und können bei Bedarf über die Piggeldat-Website veröffentlicht werden. Die zentrale Speicherung der Daten unterstütze die Durchführung von mehreren statistischen und bioinformatischen Datenanalysen, womit vielfältige Auswirkungen der Fütterung der beiden Nahrungsfaktoren gezeigt werden konnten. Die integrative Analyse der Messdaten aus einem Fütterungsversuch zeigte das veränderte Zusammenspiel zwischen gastrointestinaler Mikrobiota und dem Immunsystem des Wirts durch die Fütterung mit dem Probiotikum *Enterococcus faecium* NCIMB 10415.

Zur Installation der Piggeldat-Datenbank erfolgte die Programmierung der Datenbankanwendung Catria. Catria basiert auf dem Content-Management-System Drupal und dient zur Verwaltung von Forschungsdaten aus Fütterungsversuchen über eine Website. Hierbei bietet Catria die vier Funktionen der Nutzerverwaltung, Datenintegration, Datenverwaltung und Datenansicht an. Das Catria-Datenmodell unterstützt die Integration von allgemeinen Multi-Omics Messdaten aus Fütterungsversuchen und liefert damit ein allgemeines Muster zur strukturierten Publikation von Forschungsdaten aus Fütterungsversuchen. Nach der Integration der Forschungsdaten können die Nutzer die Daten mit anderen Nutzern teilen oder die eigenen Messdaten öffentlich publizieren. Catria ist frei verfügbar und lässt sich in zukünftigen Projekten zur Datenverwaltung einsetzen.

Die Integration von Messdaten aus unterschiedlichen biologischen Organisationsebenen diente zur explorativen Untersuchung der Auswirkung des Probiotikums *Enterococcus faecium* NCIMB 10415 auf das Zusammenspiel zwischen Ernährung, gastrointestinaler Mikrobiota und Wirt beim Ferkel. Insgesamt war die Auswirkung der Fütterung des Probiotikums auf die gemessenen Merkmale relativ gering. Die Anwendung einer Hauptkomponentenanalyse zeigte, dass starke Alterseffekte die Varianz im gesamten Datensatz dominierten. Hierbei konnten vier verschiedene Muster der Altersentwicklung identifiziert werden. Nichtsdestotrotz zeigte die Fütterung des Probiotikums einen signifikanten Effekt auf die Gesamtheit der gemessenen Merkmale. Bei den 26 Tage alten Ferkeln und 34 Tage alten Ferkeln konnte der deutlichste Unterschied zwischen

## Zusammenfassung

den Behandlungsgruppen beobachtet werden. Dabei unterschieden sich die Auswirkungen des Probiotikums zwischen den vier Messzeitpunkten deutlich. Besonders ausgeprägt war der Effekt der Probiotikum-Fütterung auf die Zusammensetzung und die Aktivität der gastrointestinalen Mikrobiota und auf die Reaktion des Immunsystems. Die Untersuchung der Korrelationen zwischen den gemessenen Merkmalen zeigte ein verändertes Zusammenspiel zwischen der gastrointestinalen Mikrobiota und der Reaktion des Immunsystems. Des Weiteren erhöhte *Enterococcus faecium* NCIMB 10415 den Zusammenhang der gastrointestinalen Mikrobiota zwischen Ileum und Colon.

## 8. Summary

### The Piggeldat database as a basis for the analysis of the effects of nutritional factors on piglets

The effect of a nutritional factor on piglets depends on a complex cross-talk between nutrition, intestinal microbiota and the host reaction. The analysis of nutritional factors benefits especially from the integration of measurement data from the fields of metagenomics, metaproteomics, metatranscriptomics and metabolomics. The collaborative research group SFB 852 performed seven animal trials between 2010 and 2013 to analyse the effects of the probiotic *Enterococcus faecium* NCIMB 10415 and of the addition of zinc to growing piglets. The project groups of the SFB 852 performed diverse measurements to determine the composition and activity of the intestinal microbiota, expression of immune-related genes, composition of immune cells and physiological parameters.

In the context of this work the Piggeldat database was installed and administered in order to manage and archive diverse measurement data from the SFB 852. The Piggeldat database contains metadata about 490 piglets, 57 measurements, 1244 attributes and it contains 81060 measurement values. With this, the Piggeldat database provides access to measurement data from 17 scientific publications. The documented data is freely available to all members of the research group and may be published online on the Piggeldat website on demand. The central storage of measurement data supported the application of several statistical and bioinformatical data analyses, which resulted in the identification of diverse effects of the two nutritional factors. For instance, an integrative analysis of measurement data from one feeding trial identified an effect of the probiotic *Enterococcus faecium* NCIMB 10415 on the interplay between microbial community and immune system.

The database software Catria was implemented in order to install the Piggeldat database. Catria is based on the content management system Drupal and allows for the management of research data from feeding trials via a website. Thereby, Catria provides the separate functions of user administration, data integration, data management and data view. The Catria data-model permits the integration of diverse types of multi-omics data from feeding trials. With this, Catria provides a general pattern for the structured publication of measurement data from feeding trials. After data integration, users may share their data with other users or give public access to their data sets. Catria is freely available and may be used for data management in other projects.

A comprehensive data set with measurements from different levels of biological areas was used for an exploratory data analysis to elucidate effects of the probiotic *Enterococcus faecium* NCIMB 10415 on the interplay between nutrition, intestinal microbiota and host in growing piglets. The overall effect of the probiotic treatment on the whole dataset was relatively low. The application of a principle component analysis indicated dominant age effects that were separated into four distinct patterns. Nevertheless, the analysis identified a significant probiotic effect. The difference between the treatment groups was strongest for piglets between 26 and 34 days old but the effect of the probiotic treatment varied between piglets from different age groups. Enrichment

## Summary

analysis indicated that the effect of the probiotic treatment was most recent on the activity and composition of the intestinal microbiota as well as on the reaction of the immune system. Correlation analyses identified altered associations between intestinal microbial communities and immune system reactions by probiotic supplementation. Furthermore, *Enterococcus faecium* NCIMB 10415 increased the association of the intestinal microbiota between ileum and colon.



# Literaturverzeichnis

- [1] A. Aarnink and M. Verstegen. Nutrition, key factor to reduce environmental load from pig production. *Livestock Science*, 109(1–3):194–203, 2007.
- [2] R. Aebersold and M. Mann. Mass spectrometry-based proteomics. *Nature*, 422(6928):198–207, 2003.
- [3] A. C. Ahn, M. Tewari, C.-S. Poon, and R. S. Phillips. The Limits of Reductionism in Medicine: Could Systems Biology Offer an Alternative? *PLoS Medicine*, 3(6):e208, 2006.
- [4] S. Albers, D. Klapper, U. Konradt, A. Walter, and J. H. Wolf. *Methodik der empirischen Forschung*. Gabler, Wiesbaden, 2007.
- [5] M. Arumugam, J. Raes, E. Pelletier, D. Le Paslier, T. Yamada, D. R. Mende, G. R. Fernandes, J. Tap, T. Bruls, M. Batto, Jean-Michel Bertalan, N. Borruel, F. Casellas, L. Fernandez, L. Gautier, T. Hansen, M. Hattori, T. Hayashi, M. Kleerebezem, K. Kurokawa, M. Leclerc, F. Levenez, C. Manichanh, H. B. Nielsen, T. Nielsen, N. Pons, J. Poulain, J. Qin, T. Sicheritz-Ponten, S. Tims, D. Torrents, E. Ugarte, E. G. Zoetendal, J. Wang, F. Guarner, O. Pedersen, W. M. de Vos, S. Brunak, J. Dorá, MetaHIT Consortium (additional members), J. Weissenbach, S. D. Ehrlich, and P. Bork. Enterotypes of the human gut microbiome. *Nature*, 473(7346):174–180, 2011.
- [6] J. Bassaganya-Riera and R. Hontecillas. CLA and n-3 PUFA differentially modulate clinical activity and colonic PPAR-responsive gene expression in a pig model of experimental IBD. *Clinical Nutrition*, 25(3):454–465, 2006.
- [7] A. Bauch, I. Adamczyk, P. Buczek, F.-J. Elmer, K. Enimanev, P. Glyzewski, M. Kohler, T. Pylak, A. Quandt, C. Ramakrishnan, C. Beisel, L. Malmström, R. Aebersold, and B. Rinn. openBIS: a flexible framework for managing and analyzing complex data in biology research. *BMC Bioinformatics*, 12(1):468, 2011.
- [8] C. Bednorz, S. Guenther, K. Oelgeschläger, B. Kinnemann, R. Pieper, S. Hartmann, K. Tedin, T. Semmler, K. Neumann, P. Schierack, A. Bethe, and L. H. Wieler. Feeding the Probiotic *Enterococcus faecium* Strain NCIMB 10415 to Piglets Specifically Reduces the Number of *Escherichia coli* Pathotypes That Adhere to the Gut Mucosa. *Applied and Environmental Microbiology*, 79(24):7896–7904, 2013.
- [9] C. Bednorz, K. Oelgeschläger, B. Kinnemann, S. Hartmann, K. Neumann, R. Pieper, A. Bethe, T. Semmler, K. Tedin, P. Schierack, L. H. Wieler, and S. Guenther. The broader context of antibiotic resistance: Zinc feed supplementation of piglets increases the proportion of multi-resistant *Escherichia coli* in vivo. *International Journal of Medical Microbiology*, 303(6):396–403, 2013.

- [10] E. Bendixen, M. Danielsen, K. Larsen, and C. Bendixen. Advances in porcine genomics and proteomics - a toolbox for developing the pig as a model organism for molecular biomedical research. *Briefings in Functional Genomics*, 9(3):208–219, 2010.
- [11] Berlin Declaration. Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. <http://openaccess.mpg.de/Berliner-Erklaerung>, Oktober 2003.
- [12] D. M. Bolser, P.-Y. Chibon, N. Palopoli, S. Gong, D. Jacob, V. D. Del Angel, D. Swan, S. Bassi, V. González, P. Suravajhala, S. Hwang, P. Romano, R. Edwards, B. Bishop, J. Eargle, T. Shtatland, N. J. Provart, D. Clements, D. P. Renfro, D. Bhak, and J. Bhak. MetaBase - the wiki-database of biological databases. *Nucleic Acids Research*, 40(D1):D1250–D1254, 2012.
- [13] L. Broom, H. Miller, K. Kerr, and J. Knapp. Effects of zinc oxide and *Enterococcus faecium* SF68 dietary supplementation on the performance, intestinal microbiota and immune status of weaned piglets. *Research in Veterinary Science*, 80(1):45–54, 2006.
- [14] F. J. Bruggeman and H. V. Westerhoff. The nature of systems biology. *Trends in Microbiology*, 15(1):45–50, 2007.
- [15] D. Buytaert. Drupal. <https://drupal.org/>, Oktober 2014.
- [16] W. Chai, M. Burwinkel, Z. Wang, C. Palissa, B. Esch, S. Twardziok, J. Rieger, P. Wrede, and M. F. G. Schmidt. Antiviral effects of a probiotic *Enterococcus faecium* strain against transmissible gastroenteritis coronavirus. *Archives of Virology*, 158(4):799–807, 2013.
- [17] W. Chai, Z. Wang, P. Janczyk, S. Twardziok, U. Blohm, N. Osterrieder, and M. Burwinkel. Elevated dietary zinc oxide levels do not have a substantial effect on porcine reproductive and respiratory syndrome virus (PPRSV) vaccination and infection. *Virology Journal*, 11(1):140, 2014.
- [18] W. Chai, S. Zakrzewski, D. Gunzel, R. Pieper, Z. Wang, S. Twardziok, P. Janczyk, N. Osterrieder, and M. Burwinkel. High-dose dietary zinc oxide mitigates infection with transmissible gastroenteritis virus in piglets. *BMC Veterinary Research*, 10(1):75, 2014.
- [19] P. Chen. The Entity-Relationship Model: Toward a Unified View of Data. *ACM Transactions on Database Systems*, 1:9–36, 1976.
- [20] J. Choi, J. Kim, S. Ingale, K. Kim, P. Shinde, I. Kwon, and B. Chae. Effect of potential multimicrobe probiotic product processed by high drying temperature and antibiotic on performance of weanling pigs. *Journal of Animal Science*, 89(6):1795–1804, 2011.
- [21] I. G. Chong and C. H. Jun. Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, 78(1):103–112, 2005.
- [22] J. C. Clemente, L. K. Ursell, L. W. Parfrey, and R. Knight. The impact of the gut microbiota on human health: an integrative view. *Cell*, 148(6):1258–1270, 2012.
- [23] G. L. Cromwell. Why and how antibiotics are used in swine production. *Animal Biotechnology*, 13(1):7–27, 2002.

- [24] C. Curdt, D. Hoffmeister, G. Waldhoff, C. Jekel, and G. Bareth. Development of a Metadata Management System for AN Interdisciplinary Research Project. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 1:7–12, 2012.
- [25] H. Daniel, A. M. Gholami, D. Berry, C. Desmarchelier, H. Hahne, G. Loh, S. Mondot, P. Lepage, M. Rothballer, A. Walker, C. Böhm, M. Wenning, M. Wagner, M. Blaut, P. Schmitt-Kopplin, B. Kuster, D. Haller, and T. Clavel. High-fat diet alters gut microbiota physiology in mice. *The ISME journal*, 8(2):295–308, 2014.
- [26] S. Das, L. Girard, T. Green, L. Weitzman, A. Lewis-Bowen, and T. Clark. Building biomedical web communities using a semantically aware content management system. *Briefings in Bioinformatics*, 10(2):129–138, 2009.
- [27] S. Datta and L. M. DePadilla. Feature selection and machine learning with mass spectrometry data for distinguishing cancer and non-cancer samples. *Statistical Methodology*, 3(1):79–92, 2006.
- [28] K. Dettmer, P. A. Aronov, and B. D. Hammock. Mass spectrometry-based metabolomics. *Mass Spectrometry Reviews*, 26(1):51–78, 2007.
- [29] Deutsche Forschungsgemeinschaft. *Tierversuche in der Forschung*. Lemmens Verlags- & Mediengesellschaft, Bonn, 2004.
- [30] P. Drotleff. *pareview,sh* - Resources for Drupal Project Application reviewers. <http://pareview.sh/>, Oktober 2014.
- [31] R. Elmasri and S. B. Navathe. *Grundlagen von Datenbanksystemen*. Pearson Studium, München, 2009.
- [32] H. J. Flint, K. P. Scott, P. Louis, and S. H. Duncan. The role of the gut microbiota in nutrition and health. *Nature Reviews Gastroenterology and Hepatology*, 9(10):577–589, 2012.
- [33] E. A. Franzosa, X. C. Morgan, N. Segata, L. Waldron, J. Reyes, A. M. Earl, G. Giannoukos, M. R. Boylan, D. Ciulla, D. Gevers, J. Izard, W. S. Garrett, A. T. Chan, and C. Huttenhower. Relating the metatranscriptome and metagenome of the human gut. *Proceedings of the National Academy of Sciences*, 111(22):E2329–E2338, 2014.
- [34] S. Friedhoff, C. Meier zu Verl, C. Pietsch, C. Meyer, J. Vompras, and S. Liebig. *Social Research Data: Documentation, Management, and Technical Implementation within the SFB 882*. SFB 882 Working Paper Series, vol. 16, Bielefeld: DFG Research Center (SFB) 882 From Heterogeneities to Inequalities, 2013.
- [35] E. Gefeller, H. Martens, J. Aschenbach, S. Klingspor, S. Twardziok, P. Wrede, R. Pieper, and U. Lodemann. Effects of age and zinc supplementation on transport properties in the jejunum of piglets. *Journal of Animal Physiology and Animal Nutrition*, 2014.
- [36] M. Goeken. *Entwicklung von Data-Warehouse-Systemen*. Deutscher Universitäts-Verlag, Wiesbaden, 2006.

- [37] M. J. Gosalbes, A. Durbán, M. Pignatelli, J. J. Abellan, N. Jiménez-Hernández, A. E. Pérez-Cobas, A. Latorre, and A. Moya. Metatranscriptomic Approach to Analyze the Functional Human Gut Microbiota. *PLoS ONE*, 6(3):e17447, 2011.
- [38] P. Groth, N. Pavlova, I. Kalev, S. Tonov, G. Georgiev, H.-D. Pohlenz, and B. Weiss. PhenomicDB: a new cross-species genotype/phenotype resource. *Nucleic Acids Research*, 35(suppl 1):D696–D699, 2007.
- [39] A. Handl. *Multivariate Analysemethoden, Statistik und ihre Anwendung, 2nd edition*. Springer-Verlag, Berlin Heidelberg, 2010.
- [40] K. Hasegawa and K. Funatsu. Evolution of PLS for Modeling SAR and omics Data. *Molecular Informatics*, 31(11-12):766–775, 2012.
- [41] S. N. Heinritz, R. Mosenthin, and E. Weiss. Use of pigs as a potential model for research into dietary modulation of the human gut microbiota. *Nutrition Research Reviews*, 26(02):191–209, 2013.
- [42] M. Heron, V. L. Hanson, and I. Ricketts. Open source and accessibility: advantages and limitations. *Journal of Interaction Science*, 1(1):1–10, 2013.
- [43] L. V. Hooper, D. R. Littman, and A. J. Macpherson. Interactions Between the Microbiota and the Immune System. *Science*, 336(6086):1268–1273, 2012.
- [44] J. L. Horn. A rationale and test for the number of factors in factor analysis. *Psychometrika*, 30(2):179–185, 1965.
- [45] A. Hubble, D. A. Murphy, and S. C. Perry. From Static and Stale to Dynamic and Collaborative: The Drupal Difference. *Information Technology and Libraries*, 30(4):190–197, 2011.
- [46] Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature*, 486(7402):207–214, 2012.
- [47] S. Huss, F. Fronhoffs, R. Büttner, and L. C. Heukamp. Web-based database for the management of tissue specimens in a transregional histological research facility. *Diagnostic Pathology*, 6(1):17, 2011.
- [48] T. Ideker, T. Galitski, and L. Hood. A new approach to decoding life: systems biology. *Annual Review of Genomics and Human Genetics*, 2(1):343–372, 2001.
- [49] E. Ivarsson, S. Roos, H. Y. Liu, and J. E. Lindberg. Fermentable non-starch polysaccharides increases the abundance of Bacteroides–Prevotella–Porphyromonas in ileal microbial community of growing pigs. *Animal: an international journal of animal bioscience*, 8(11):1777–1787, 2014.
- [50] T. Iwanaga, S. El Sawah, and A. Jakeman. Design and implementation of a web-based groundwater data management system. *Mathematics and Computers in Simulation*, 93:164–174, 2013.

- [51] P. Janczyk, S. Kreuzer, J. Assmus, K. Nöckler, and G. A. Brockmann. No Protective Effects of High-Dosage Dietary Zinc Oxide on Weaned Pigs Infected with *Salmonella enterica* Serovar Typhimurium DT104. *Applied and Environmental Microbiology*, 79(9):2914–2921, 2013.
- [52] J. Josse and H. François. Handling missing values in exploratory multivariate data analysis methods. *Journal de la Société Française de Statistique*, 153(2):79–99, 2012.
- [53] A. L. Kau, P. P. Ahern, N. W. Griffin, A. L. Goodman, and J. I. Gordon. Human nutrition, the gut microbiome and the immune system. *Nature*, 474(7351):327–336, 2011.
- [54] G. T. Keusch. What do *-omics* mean for the science and policy of the nutritional sciences? *The American Journal of Clinical Nutrition*, 83(2):520S–522S, 2006.
- [55] H. Kitano. Systems Biology: A Brief Overview. *Science*, 295(5560):1662–1664, 2002.
- [56] S. Klingspor, H. Martens, D. Caushi, S. Twardziok, J. R. Aschenbach, and U. Lodemann. Characterization of the effects of *Enterococcus faecium* on intestinal epithelial transport properties in piglets. *Journal of Animal Science*, 91(4):1707–1718, 2013.
- [57] J. Klump, R. Bertelmann, J. Brase, M. Diepenbroek, H. Grobe, H. Höck, M. Lautenschlager, U. Schindler, I. Sens, and J. Wächter. Data publication in the open access initiative. *Data Science Journal*, 5:79–83, 2006.
- [58] E. Kolker, V. Özdemir, L. Martens, W. Hancock, G. Anderson, N. Anderson, S. Aynacioglu, A. Baranova, S. R. Campagna, R. Chen, J. Choiniere, S. P. Dearth, W. C. Feng, L. Ferguson, G. Fox, D. Frishman, R. Grossman, A. Heath, R. Higdon, M. H. Hutz, I. Janko, L. Jiang, S. Joshi, A. Ke, J. W. Kemnitz, I. S. Kohane, N. Kolker, D. Lancet, E. Lee, W. Li, A. Lisitsa, A. Llerena, C. Macnealy-Koch, J. C. Marshall, P. Masuzzo, A. May, G. Mias, M. Monroe, E. Montague, S. Mooney, A. Nesvizhskii, S. Noronha, G. Omenn, H. Rajasimha, P. Ramamoorthy, J. Sheehan, L. Smarr, C. V. Smith, T. Smith, M. Snyder, S. Rapole, S. Srivastava, L. Stanberry, E. Stewart, S. Toppo, U. P. K. Verheggen, B. H. Voy, L. Warnich, S. W. Wilhelm, and G. Yandl. Toward More Transparent and Reproducible Omics Studies Through a Common Metadata Checklist and Data Publications. *Omics: A Journal of Integrative Biology*, 18(1):10–14, 2014.
- [59] S. Kreuzer, P. Janczyk, J. Assmus, M. F. G. Schmidt, G. A. Brockmann, and K. Nöckler. No Beneficial Effects Evident for *Enterococcus faecium* NCIMB 10415 in weaned pigs infected with *Salmonella enterica* serovar Typhimurium DT104. *Applied and Environmental Microbiology*, 78(14):4816–4825, 2012.
- [60] S. Kreuzer, P. Machnowska, J. Aßmus, M. Sieber, R. Pieper, M. F. G. Schmidt, G. A. Brockmann, L. Scharek-Tedin, and R. John. Feeding of the probiotic bacterium *Enterococcus faecium* NCIMB 10415 differentially affects shedding of enteric viruses in pigs. *Veterinary Research*, 43:58, 2012.
- [61] R. Lamendella, J. W. Santo Domingo, S. Ghosh, J. Martinson, and D. B. Oerther. Comparative fecal metagenomics unveils unique functional capacity of the swine gut. *BMC Microbiology*, 11(1):103, 2011.

- [62] R. Lamendella, N. VerBerkmoes, and J. K. Jansson. 'Omics' of the mammalian gut – new insights into function. *Current Opinion in Biotechnology*, 23(3):491–500, 2012.
- [63] S. Lê, J. Josse, and F. Husson. FactoMineR: An R Package for Multivariate Analysis. *Journal of Statistical Software*, 25(1):1–18, 2008.
- [64] T. D. Leser, J. Z. Amenuvor, T. K. Jensen, R. H. Lindecrona, M. Boye, and K. Møller. Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. *Applied and Environmental Microbiology*, 68(2):673–690, 2002.
- [65] P. Liu, R. Pieper, L. Tedin, L. Martin, W. Meyer, J. Rieger, J. Plendl, W. Vahjen, and J. Zentek. Effect of dietary zinc oxide on jejunal morphological and immunological characteristics in weaned piglets. *Journal of Animal Science*, pages jas–2013, 2014.
- [66] T. Looft, T. A. Johnson, H. K. Allen, D. O. Bayles, D. P. Alt, R. D. Stedtfeld, W. J. Sul, T. M. Stedtfeld, B. Chai, J. R. Cole, S. A. Hashshamb, J. M. Tiedjeb, and T. B. Stanton. In-feed antibiotic effects on the swine intestinal microbiome. *Proceedings of the National Academy of Sciences*, 109(5):1691–1696, 2012.
- [67] L. Martin, R. Pieper, S. Kröger, F. G. Boroojeni, W. Vahjen, K. Neumann, A. V. Kessel, and J. Zentek. Influence of age and *Enterococcus faecium* NCIMB 10415 on development of small intestinal digestive physiology in piglets. *Animal Feed Science and Technology*, 175(1–2):65–75, 2012.
- [68] L. Martin, R. Pieper, N. Schunter, W. Vahjen, and J. Zentek. Performance, organ zinc concentration, jejunal brush border membrane enzyme activities and mRNA expression in piglets fed with different levels of dietary zinc. *Archives of Animal Nutrition*, 67(3):248–261, 2013.
- [69] V. Marx. Biology: The big challenges of big data. *Nature*, 498(7453):255–260, 2013.
- [70] M. Meike, J. Sameting, and A. Wiesauer. Security in Open Source Web Content Management Systems. *IEEE Security & Privacy*, 7(4):44–51, 2009.
- [71] J. C. Molloy. The Open Knowledge Foundation: Open Data Means Better Science. *PLoS Biology*, 9(12):e1001195, 2011.
- [72] L. Montagne, G. Boudry, C. Favier, I. Le Huërou-Luron, J.-P. Lallès, and B. Sève. Main intestinal markers associated with the changes in gut architecture and function in piglets after weaning. *British Journal of Nutrition*, 97(1):45–57, 2007.
- [73] S. D. Mooney and P. H. Baenziger. Extensible open source content management systems and frameworks: a solution for many needs of a bioinformatics group. *Briefings in Bioinformatics*, 9(1):69–74, 2008.
- [74] National Center for Biotechnology Information, U.S. National Library of Medicine. Genbank and WGS statistics. [www.ncbi.nlm.nih.gov/genbank/statistics](http://www.ncbi.nlm.nih.gov/genbank/statistics), Juni 2014.
- [75] J. K. Nicholson, E. Holmes, J. Kinross, R. Burcelin, G. Gibson, W. Jia, and S. Pettersson. Host-gut microbiota metabolic interactions. *Science*, 336(6086):1262–1267, 2012.

- [76] B. P. O'connor. SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behavior Research Methods, Instruments, & Computers*, 32(3):396–402, 2000.
- [77] E. A. B. Pajarillo, J. P. Chae, M. P. Balolong, H. B. Kim, K.-S. Seo, and D.-K. Kang. Pyrosequencing-based analysis of fecal microbial communities in three purebred pig lines. *Journal of Microbiology*, 52(8):646–651, 2014.
- [78] A. Papanicolaou and D. G. Heckel. The GMOD Drupal Bioinformatic Server Framework. *Bioinformatics*, 26(24):3119–3124, 2010.
- [79] J. Pettigrew and M. Esnaola. Swine nutrition and pork quality: A review. *Journal of Animal Science*, 79(E.-Suppl.):E316–E342, 2001.
- [80] V. Poroyko, J. R. White, M. Wang, S. Donovan, J. Alverdy, D. C. Liu, and M. J. Morowitz. Gut Microbial Gene Expression in Mother-Fed and Formula-Fed Piglets. *PLoS ONE*, 5(8):e12459, 2010.
- [81] J. Qin, R. Li, J. Raes, M. Arumugam, K. S. Burgdorf, C. Manichanh, T. Nielsen, N. Pons, F. Levenez, T. Yamada, D. R. Mende, J. Li, J. Xu, S. Li, D. Li, J. Cao, B. Wang, H. Liang, H. Zheng, Y. Xie, J. Tap, P. Lepage, M. Bertalan, J.-M. Batto, T. Hansen, D. Le Paslier, A. Linneberg, H. B. Nielsen, E. Pelletier, P. Renault, T. Sicheritz-Ponten, K. Turner, H. Zhu, C. Yu, S. Li, M. Jian, Y. Zhou, Y. Li, X. Zhang, S. Li, N. Qin, H. Yang, J. Wang, S. Brunak, J. Doré, F. Guarner, K. Kristiansen, O. Pedersen, J. Parkhill, J. Weissenbach, MetaHIT Consortium, P. Bork, S. D. Ehrlich, and J. Wang. A human gut microbial gene catalogue established by metagenomic sequencing. *Nature*, 464(7285):59–65, 2010.
- [82] M. D. Radmacher, L. M. McShane, and R. Simon. A Paradigm for Class Prediction using Gene Expression Profiles. *Journal of Computational Biology*, 9(3):505–511, 2002.
- [83] J. F. Richter, J. D. Schulzke, and M. Fromm. Influence of *Enterococcus faecium* NCIMB 10415 on intestinal barrier function in weaned piglets. *Proceedings of the Society of Nutrition Physiology*, 2011.
- [84] H. Roider, N. Pavlova, I. Kirov, S. Slavov, T. Slavov, Z. Uzunov, and B. Weiss. Drug2Gene: an exhaustive resource to explore effectively the drug-target relation network. *BMC Bioinformatics*, 15(1):68, 2014.
- [85] T. A. Runkler. *Data Mining: Methoden und Algorithmen intelligenter Datenanalyse*. Vieweg + Teubner, Wiesbaden, 2010.
- [86] J. Sachsenröder, S. Twardziok, J. A. Hammerl, P. Janczyk, P. Wrede, S. Hertwig, and R. Johne. Simultaneous Identification of DNA and RNA Viruses Present in Pig Faeces Using Process-Controlled Deep Sequencing. *PLoS ONE*, 7(4):e34631, 2012.
- [87] J. Sachsenröder, S. O. Twardziok, M. Scheuch, and R. Johne. The General Composition of the Faecal Virome of Pigs Depends on Age, but Not on Feeding with a Probiotic Bacterium. *PLoS ONE*, 9(2):e88888, 2014.
- [88] I. San Gil. Drupal Ecological Information Management System. <https://www.drupal.org/project/deims>, Mai 2014.

- [89] D. Sauvant, P. Schmidely, J. J. Daudin, and N. R. St-Pierre. Meta-analyses of experimental data in animal nutrition. *Animal*, 2:1203–1214, 2008.
- [90] L. Scharek-Tedin, S. Kreuzer, S. O. Twardziok, B. Siepert, R. Klopffleisch, K. Tedin, R. Pieper, and J. Zentek. Probiotic treatment decreases the number of CD14-expressing cells in porcine milk which correlates with several intestinal immune parameters in the piglets. *Frontiers in immunology*, 6, 2015.
- [91] C. P. Schmitt and M. Burchinal. Data Management Practices for Collaborative Research. *Frontiers in Psychiatry*, 2, 2011.
- [92] G. Schneider and P. Wrede. Artificial neural networks for computer-based molecular design. *Progress in Biophysics and Molecular Biology*, 70(3):175–222, 1998.
- [93] G. Schryen. Is open source security a myth? *Communications of the ACM*, 54(5):130–140, 2011.
- [94] K. P. Scott, S. W. Gratz, P. O. Sheridan, H. J. Flint, and S. H. Duncan. The influence of diet on the gut microbiota. *Pharmacological Research*, 69(1):52–60, 2013.
- [95] B. Siepert, N. Reinhardt, S. Kreuzer, A. Bondzio, S. Twardziok, G. Brockmann, K. Nöckler, I. Szabó, P. Janczyk, R. Pieper, and K. Tedin. *Enterococcus faecium* NCIMB 10415 supplementation affects intestinal immune-associated gene expression in post-weaning piglets. *Veterinary Immunology and Immunopathology*, 157(1-2):65–77, 2014.
- [96] O. Simon. An interdisciplinary study on the mode of action of probiotics in pigs. *Journal of Animal and Feed Sciences*, 19(2):230–243, 2010.
- [97] E. Simukovic, M. Kindling, and P. Schirmbacher. Umfrage zum Umgang mit digitalen Forschungsdaten an der Humboldt-Universität zu Berlin. Umfragebericht, Version 1.0. Online verfügbar auf dem edoc-Server der Humboldt-Universität zu Berlin, 2013.
- [98] A. U. Sinha, E. Merrill, S. A. Armstrong, T. W. Clark, and S. Das. eXframe: reusable framework for storage, analysis and visualization of genomics experiments. *BMC Bioinformatics*, 12(1):452, 2011.
- [99] V. Smith, S. Rycroft, K. Harman, B. Scott, and D. Roberts. Scratchpads: a data-publishing framework to build, share and manage information on the diversity of life. *BMC Bioinformatics*, 10(Suppl 14):S6, 2009.
- [100] F. Sommer and F. Bäckhed. The gut microbiota - masters of host development and physiology. *Nature Reviews Microbiology*, 11(4):227–238, 2013.
- [101] I. C. Starke, R. Pieper, K. Neumann, J. Zentek, and W. Vahjen. Individual responses of mother sows to a probiotic *Enterococcus faecium* strain lead to different microbiota composition in their offspring. *Beneficial Microbes*, 4(4):345–356, 2013.
- [102] I. C. Starke, R. Pieper, K. Neumann, J. Zentek, and W. Vahjen. The impact of high dietary zinc oxide on the development of the intestinal microbiota in weaned piglets. *FEMS Microbiology Ecology*, 87(2):416–427, 2014.



- [103] A. Subramanian, P. Tamayo, V. K. Mootha, S. Mukherjee, B. L. Ebert, M. A. Gillette, A. Paulovich, S. L. Pomeroy, T. R. Golub, E. S. Lander, and J. P. Mesirov. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, 102(43):15545–15550, 2005.
- [104] D. Taras, W. Vahjen, M. Macha, and O. Simon. Performance, diarrhea incidence, and occurrence of *Escherichia coli* virulence genes during long-term administration of a probiotic *Enterococcus faecium* strain to sows and piglets. *Journal of Animal Science*, 84(3):608–617, 2006.
- [105] C. F. Taylor, D. Field, S.-A. Sansone, J. Aerts, R. Apweiler, M. Ashburner, C. A. Ball, P.-A. Binz, M. Bogue, T. Booth, A. Brazma, R. R. Brinkman, A. Michael Clark, E. W. Deutsch, O. Fiehn, J. Fostel, P. Ghazal, F. Gibson, T. Gray, G. Grimes, J. M. Hancock, N. W. Hardy, H. Hermjakob, R. K. Julian, Jr, M. Kane, C. Kettner, C. Kinsinger, E. Kolker, M. Kuiper, N. Le Novère, J. Leebens-Mack, S. E. Lewis, P. Lord, A.-M. Mallon, N. Marthandan, H. Masuya, R. McNally, A. Mehrle, N. Morrison, S. Orchard, J. Quackenbush, J. M. Reecy, D. G. Robertson, P. Rocca-Serra, H. Rodriguez, H. Rosenfelder, J. Santoyo-Lopez, R. H. Scheuermann, D. Schober, B. Smith, J. Snape, C. J. Stoeckert, Jr, K. Tipton, P. Sterk, A. Untergasser, J. Vandesompele, and S. Wiemann. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. *Nature Biotechnology*, 26(8):889–896, 2008.
- [106] V. Tremaroli and F. Bäckhed. Functional interactions between the gut microbiota and host metabolism. *Nature*, 489(7415):242–249, 2012.
- [107] J. Trygg and S. Wold. Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*, 16(3):119–128, 2002.
- [108] S. O. Twardziok. Catria - A Drupal Module for Management of Measurement Data from Feedingtrials. <https://www.drupal.org/sandbox/pi314/2360619>, Oktober 2014.
- [109] S. O. Twardziok, R. Pieper, J. R. Aschenbach, C. Bednorz, G. A. Brockmann, M. Fromm, S. Klingspor, S. Kreuzer, U. Lodemann, H. Martens, L. Martin, J. F. Richter, L. Scharek-Tedin, B. F. Siepert, I. C. Starke, K. Tedin, W. Vahjen, L. H. Wieler, S. S. Zakrzewski, J. Zentek, and P. Wrede. Cross-talk Between Host, Microbiome and Probiotics: A Systems Biology Approach for Analyzing the Effects of Probiotic *Enterococcus faecium* NCIMB 10415 in Piglets. *Molecular Informatics*, 33(3):171–182, 2014.
- [110] S. O. Twardziok and P. Wrede. Piggeldat - The SFB852 Database Website. <http://piggeldat.charite.de/>, Oktober 2014.
- [111] P. van Baarlen, M. Kleerebezem, and J. M. Wells. Omics approaches to study host–microbiota interactions. *Current Opinion in Microbiology*, 16(3):270–277, 2013.
- [112] J. K. VanDyk. *Das Drupal-Entwicklerhandbuch*. Addison-Wesley, München, 2009.
- [113] N. C. Verberkmoes, A. L. Russell, M. Shah, A. Godzik, M. Rosenquist, J. Halfvarson, M. G. Lefsrud, J. Apajalahti, C. Tysk, R. L. Hettich, and J. K. Jansson. Shotgun metaproteomics of the human distal gut microbiota. *The ISME Journal*, 3(2):179–189, 2009.

- [114] W3Techs. Usage of content management systems for websites. <http://w3techs.com/>, Mai 2014.
- [115] Z. Wang, M. Burwinkel, W. Chai, E. Lange, U. Blohm, A. Breithaupt, B. Hoffmann, S. Twardziok, J. Rieger, P. Janczyk, R. Pieper, and N. Osterrieder. Dietary *Enterococcus faecium* NCIMB 10415 and Zinc Oxide Stimulate Immune Reactions to Trivalent Influenza Vaccination in Pigs but Do Not Affect Virological Response upon Challenge Infection. *PLoS ONE*, 9(1):e87007, 2014.
- [116] Z. Wang, W. Chai, M. Burwinkel, S. Twardziok, P. Wrede, C. Palissa, B. Esch, and M. F. G. Schmidt. Inhibitory influence of *Enterococcus faecium* on the Propagation of Swine Influenza A Virus *In Vitro*. *PLoS ONE*, 8(1):e53043, 2013.
- [117] Z. Wang, M. Gerstein, and M. Snyder. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1):57–63, 2009.
- [118] J. A. Westerhuis, H. Hoefsloot, S. Smit, D. Vis, A. Smilde, E. Velzen, J. Duijnhoven, and F. Dorsten. Assessment of PLS-DA cross validation. *Metabolomics*, 4(1):81–89, 2008.
- [119] R. Winnenburg, M. Urban, A. Beacham, T. K. Baldwin, S. Holland, M. Lindeberg, H. Hansen, C. Rawlings, K. E. Hammond-Kosack, and J. Köhler. PHI-base update: additions to the pathogen-host interaction database. *Nucleic Acids Research*, 36(suppl 1):D572–D576, 2008.
- [120] D. S. Wishart, T. Jewison, A. C. Guo, M. Wilson, C. Knox, Y. Liu, Y. Djoumbou, R. Mandal, F. Aziat, E. Dong, S. Bouatra, I. Sinelnikov, D. Arndt, J. Xia, P. Liu, F. Yallou, T. Bjorn Dahl, R. Perez-Pineiro, R. Eisner, F. Allen, V. Neveu, R. Greiner, and A. Scalbert. HMDB 3.0 - The Human Metabolome Database in 2013. *Nucleic Acids Research*, page gks1065, 2012.
- [121] S. Wold, M. Sjöström, and L. Eriksson. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2):109 – 130, 2001.
- [122] J. C. Wooley, A. Godzik, and I. Friedberg. A Primer on Metagenomics. *PLoS Computational Biology*, 6(2):e1000667, 2010.
- [123] W. Wruck, M. Peuker, and C. R. Regenbrecht. Data management strategies for multinational large-scale systems biology projects. *Briefings in Bioinformatics*, 15(1):65–78, 2014.
- [124] G. D. Wu, J. Chen, C. Hoffmann, K. Bittinger, Y.-Y. Chen, S. A. Keilbaugh, M. Bewtra, D. Knights, W. A. Walters, R. Knight, R. Sinha, E. Gilroy, K. Gupta, R. Baldassano, L. Nessel, H. Li, F. D. Bushman, and J. D. Lewis. Linking Long-Term Dietary Patterns with Gut Microbial Enterotypes. *Science*, 334(6052):105–108, 2011.
- [125] J. Xu and J. I. Gordon. Honor thy symbionts. *Proceedings of the National Academy of Sciences*, 100(18):10452–10459, 2003.
- [126] D. Young, M. Ibuki, T. Nakamori, M. Fan, and Y. Mine. Soy-Derived Di- and Tripeptides Alleviate Colon and Ileum Inflammation in Pigs with Dextran Sodium Sulfate-Induced Colitis. *The Journal of Nutrition*, 142(2):363–368, 2012.
- [127] A. Zeyner and Boldt. Effects of a probiotic *Enterococcus faecium* strain supplemented from birth to weaning on diarrhoea patterns and performance of piglets. *Journal of Animal Physiology and Animal Nutrition*, 90(1–2):25–31, 2006.

[128] W. Ziegler. Entity API. <https://drupal.org/project/entity>, Oktober 2014.



# Anhang



# A. Das Catria-Datenmodul

Das Catria-Datenmodul ist über die offizielle Drupal-Website frei verfügbar und enthält insgesamt 34 Dateien<sup>[108]</sup>:

**Tabelle A.1.:** Dateien des Catria-Datenmoduls

Name	Beschreibung
<i>(/)</i>	
catia.info	Interne Info-Datei zum Laden des Moduls
catia.install	Definition der Datenbanktabellen
catia.module	Deklaration der Datenobjekttypen, Definition der Menü-Items und Zuordnung der <i>callback</i> -Funktionen
README.txt	Informationen zur Installation und zur Verwendung des Moduls
<i>(/class/entity/)</i>	
catia_animal.inc	Datenobjektklasse <i>CatriaAnimal</i>
catia_attribute.inc	Datenobjektklasse <i>CatriaAttribute</i>
catia_describes.inc	Datenobjektklasse <i>CatriaDescribes</i>
catia_determines.inc	Datenobjektklasse <i>CatriaDetermines</i>
catia_download.inc	Datenobjektklasse <i>CatriaDownload</i>
catia_entity.inc	Datenobjektklasse <i>CatriaEntity</i>
catia_feedingtrial.inc	Datenobjektklasse <i>CatriaFeedingtrial</i>
catia_has.inc	Datenobjektklasse <i>CatriaHas</i>
catia_measurement.inc	Datenobjektklasse <i>CatriaMeasurement</i>
catia_variable.inc	Datenobjektklasse <i>CatriaVariable</i>
<i>(/class/controller/)</i>	
catia_animal_controller.inc	Kontrollklasse <i>CatriaAnimalController</i>
catia_attribute_controller.inc	Kontrollklasse <i>CatriaAttributeController</i>
catia_describes_controller.inc	Kontrollklasse <i>CatriaDescribesController</i>
catia_determines_controller.inc	Kontrollklasse <i>CatriaDeterminesController</i>
catia_entity_controller.inc	Kontrollklasse <i>CatriaEntityController</i>
catia_feedingtrial_controller.inc	Kontrollklasse <i>CatriaFeedingtrialController</i>
catia_has_controller.inc	Kontrollklasse <i>CatriaHasController</i>
catia_measurement_controller.inc	Kontrollklasse <i>CatriaMeasurementController</i>
<i>(/include/)</i>	
catia_blocks.inc	Definition des Nutzermenüs
catia_callbacks_animal.inc	<i>Callback</i> -Funktionen für <i>CatriaAnimal</i>
catia_callbacks_attribute.inc	<i>Callback</i> -Funktionen für <i>CatriaAttribute</i>
catia_callbacks_download.inc	<i>Callback</i> -Funktionen für <i>CatriaDownload</i>
catia_callbacks_entity.inc	<i>Callback</i> -Funktionen für <i>CatriaEntity</i>

Fortsetzung auf der nächsten Seite

**Tabelle A.1 – Fortsetzung**

<b>Name</b>	<b>Beschreibung</b>
<i>catrta_callbacks_feedingtrial.inc</i>	<i>Callback-Funktionen für CatriaFeedingtrial</i>
<i>catrta_callbacks_measurement.inc</i>	<i>Callback-Funktionen für CatriaMeasurement</i>
<i>catrta_callbacks_variable.inc</i>	<i>Callback-Funktionen für CatriaVariable</i>
<i>catrta_forms_download.inc</i>	<i>Form-Funktionen für CatriaDownload</i>
<i>catrta_forms_entity.inc</i>	<i>Form-Funktionen für CatriaEntity</i>
<i>catrta_forms_feedingtrial.inc</i>	<i>Form-Funktionen für CatriaFeedingtrial</i>
<i>catrta_forms_measurement.inc</i>	<i>Form-Funktionen für CatriaMeasurement</i>



## B. Matrix-Notationen

In dieser Arbeit werden Matrizen und Vektoren **fett** dargestellt und einzelne Einträge einer Matrix werden in normaler Schriftart dargestellt. Die folgende Matrix  $\mathbf{X} = (X)_{N,K}$  enthält  $N = 4$  Zeilen und  $K = 3$  Spalten:

$$\mathbf{X} = \begin{pmatrix} X_{1,1} & X_{1,2} & X_{1,3} \\ X_{2,1} & X_{2,2} & X_{2,3} \\ X_{3,1} & X_{3,2} & X_{3,3} \\ X_{4,1} & X_{4,2} & X_{4,3} \end{pmatrix}$$

und die Matrix  $\mathbf{Y} = (Y)_{4,1}$  (ein Spaltenvektor) enthält vier Zeilen und eine Spalte:

$$\mathbf{Y} = \begin{pmatrix} Y_{1,1} \\ Y_{2,1} \\ Y_{3,1} \\ Y_{4,1} \end{pmatrix}.$$

Ein bestimmter Zeilenvektor oder ein bestimmte Spaltenvektor einer Matrix  $\mathbf{X}$  wird mit Hilfe des Zeichen  $\bullet$  notiert. Der Zeilenvektor  $\mathbf{X}_{1,\bullet}$  besteht aus der ersten Zeile der Matrix  $\mathbf{X}$ :

$$\mathbf{X}_{1,\bullet} = (X_{1,1} \quad X_{1,2} \quad X_{1,3})$$

und der Spaltenvektor  $\mathbf{X}_{\bullet,2}$  besteht aus der zweiten Spalte der Matrix  $\mathbf{X}$ :

$$\mathbf{X}_{\bullet,2} = \begin{pmatrix} X_{1,2} \\ X_{2,2} \\ X_{3,2} \\ X_{4,2} \end{pmatrix}.$$



## C. OPLS-Berechnung und Vorhersage der Gruppenzugehörigkeit

$\mathbf{T} = (T)_{N,1}$	<i>predictive-score</i> -Vektor
$\mathbf{To} = (To)_{N,Ao}$	<i>orthogonal-score</i> -Matrix
$\mathbf{P} = (P)_{1,K}$	<i>predictive-loading</i> -Vektor
$\mathbf{Po} = (Po)_{Ao,K}$	<i>orthogonal-loading</i> -Matrix
$\mathbf{W} = (W)_{1,K}$	<i>predictive-weight</i> -Vektor
$\mathbf{Wo} = (Wo)_{Ao,K}$	<i>orthogonal-weight</i> -Matrix
$\mathbf{E} = (E)_{N,K}$	Matrix der $\mathbf{X}$ -Residuen
$\mathbf{F} = (F)_{N,1}$	Vektor der $\mathbf{Y}$ -Residuen

Gegeben ist eine Matrix  $\mathbf{X} = (X)_{N,K}$  und ein Vektor  $\mathbf{Y} = (Y)_{N,1}$ . In der vorliegenden Arbeit enthält die Matrix  $\mathbf{X}$  die Messwerte von K Merkmalen und N Tieren und der Vektor  $\mathbf{Y}$  enthält die Zugehörigkeit der Tiere zu einer von zwei Behandlungsgruppen (codiert mit -1 und 1). Ein OPLS-Modell teilt die Matrix  $\mathbf{X}$  in einen beschreibenden und in einen orthogonalen Anteil auf [107]:

$$\mathbf{X} = \mathbf{TP} + \mathbf{ToPo} + \mathbf{E}$$

Der *predictive-score*-Vektor  $\mathbf{T}$  stellt dabei ein Schätzer für  $\mathbf{Y}$  dar ( $\mathbf{Y} = \mathbf{TC} + \mathbf{F}$ ) und die *orthogonal-scores*  $\mathbf{To}$  sind orthogonal zu  $\mathbf{Y}$  ( $\mathbf{To}^t\mathbf{Y} = 0$ ). Die OPLS-Berechnung mit dem NIPALS-Algorithmus erfolgt für eine feste Anzahl  $Ao$  an orthogonalen Komponenten. Dabei erfolgt die Berechnung der Matrizen  $\mathbf{Wo}$ ,  $\mathbf{Po}$  und  $\mathbf{To}$  iterativ über die Anzahl der orthogonalen Komponenten (Algorithmus 2):

**Zeile 2:** Zuerst erfolgt die Berechnung des *predictive-weight*-Vektors  $\mathbf{W}$ .

**Zeile 3:** Als nächstes berechnet sich der *predictive-score*-Vektor  $\mathbf{T}$  als Abbildung von  $\mathbf{X}$  auf  $\mathbf{W}$ .

**Zeile 4:** Der *predictive-loading*-Vektor  $\mathbf{P}$  stellt eine Drehung von  $\mathbf{W}$  durch die Matrix  $\mathbf{X}'\mathbf{X}$  dar.

**Zeilen 5 bis 8:** Mit den Vektoren  $\mathbf{W}$  und  $\mathbf{P}$  erfolgt die Berechnung des orthogonalen Anteils in  $\mathbf{X}$ . Die Abbildung von  $\mathbf{To}$  auf  $\mathbf{Po}$  gibt den orthogonalen Anteil von  $\mathbf{X}$  an.

**Zeile 9:** Der orthogonale Anteil wird von  $\mathbf{X}$  abgezogen und die nächste orthogonale Komponente kann berechnet werden.

**Algorithmus 2:** NIPALS-Algorithmus zur Berechnung des orthogonalen Anteils in der Datenmatrix  $\mathbf{X}$

```

Input :  $\mathbf{X}$ ,  $\mathbf{y}$ ,  $A_0$ 
1 for  $a \in \{1 \dots A_0\}$  do
2    $\mathbf{W} = \frac{\mathbf{Y}'\mathbf{X}}{\|\mathbf{Y}'\mathbf{X}\|}$ 
3    $\mathbf{T} = \mathbf{X}\mathbf{W}'$ 
4    $\mathbf{P} = \frac{\mathbf{T}'\mathbf{X}}{\mathbf{T}'\mathbf{T}} = \frac{\mathbf{W}\mathbf{X}'\mathbf{X}}{\mathbf{W}\mathbf{X}'\mathbf{X}\mathbf{W}'}$ 
5    $\mathbf{W}_{o_{a,\bullet}} = \mathbf{P} - \frac{\mathbf{W}\mathbf{P}'}{\mathbf{W}\mathbf{W}'} \times \mathbf{W}$ 
6    $\mathbf{W}_{o_{a,\bullet}} = \frac{\mathbf{W}_{o_{a,\bullet}}}{\|\mathbf{W}_{o_{a,\bullet}}\|}$ 
7    $\mathbf{T}_{o_{\bullet,a}} = \frac{\mathbf{X}\mathbf{W}_{o_{a,\bullet}}'}{\mathbf{W}_{o_{a,\bullet}}\mathbf{W}_{o_{a,\bullet}}'}$ 
8    $\mathbf{P}_{o_{a,\bullet}} = \frac{\mathbf{T}_{o_{\bullet,a}}'\mathbf{X}}{\mathbf{T}_{o_{\bullet,a}}'\mathbf{T}_{o_{\bullet,a}}}$ 
9    $\mathbf{X} = \mathbf{X} - \mathbf{T}_{o_{\bullet,a}}\mathbf{P}_{o_{a,\bullet}}$ 
10 end
Output :  $(\mathbf{P}_o, \mathbf{T}_o, \mathbf{W}_o)$ 

```

Nach der Berechnung des orthogonalen Anteils kann man den gesamten orthogonalen Anteil von  $\mathbf{X}$  abziehen und anschließend den *predictive-weight*-Vektor  $\mathbf{W}$  und den *predictive-score*-Vektor  $\mathbf{T}$  berechnen:

$$\mathbf{W} = \frac{\mathbf{Y}^t(\mathbf{X} - \mathbf{T}_o\mathbf{P}_o)}{\|\mathbf{Y}^t(\mathbf{X} - \mathbf{T}_o\mathbf{P}_o)\|}$$

$$\mathbf{T} = (\mathbf{X} - \mathbf{T}_o\mathbf{P}_o)\mathbf{W}^t$$

Zur Vorhersage der Gruppenzugehörigkeit  $\hat{y}$  aus einem Vektor  $\mathbf{x} = (x)_{1,K}$  wird zuerst der orthogonale Anteil von  $\mathbf{x}$  mit den Matrizen  $\mathbf{W}_o$  und  $\mathbf{P}_o$  abgezogen und anschließend  $\hat{y}$  berechnet:

$$\hat{y} = (\mathbf{x} - \mathbf{x}\mathbf{W}_o^t(\mathbf{P}_o\mathbf{W}_o^t)^{-1}\mathbf{P}_o)\mathbf{W}_o^t \mathbf{C}$$

# D. Piggeldat-Datenbank

**Tabelle D.1.:** Datensätze in der Piggeldat-Datenbank

Messung	Fütterungsversuch	Publikation
Zusammensetzung der intestinalen Mikrobiota	EFFT2010	Starke et al. 2013 <sup>[101]</sup>
Konzentration bakterieller Metabolite	EFFT2010	
Vorkommen virulenter <i>Escherichia coli</i> Gene	EFFT2010	Bednorz et al. 2013 <sup>[8]</sup>
Transport-Funktion der Darmschleimhaut	EFFT2010	Klingspor et al. 2013 <sup>[56]</sup>
Barrieren-Funktion der Darmschleimhaut	EFFT2010	Richter et al. 2011 <sup>[83]</sup>
Morphologie der Darmschleimhaut	EFFT2010	Martin et al. 2012 <sup>[67]</sup>
Aktivität der Verdauungsenzyme	EFFT2010	Martin et al. 2012 <sup>[67]</sup>
Genexpression des Darmepithels	EFFT2010	Martin et al. 2012 <sup>[67]</sup>
Konzentration der Zytokine im Blut	EFFT2010	Siepert et al. 2014 <sup>[95]</sup>
Expression immunrelevanter Gene	EFFT2010	Siepert et al. 2014 <sup>[95]</sup>
Zusammensetzung der Immunzellpopulationen (HU)	EFFT2010	Kreuzer et al. 2012 <sup>[60]</sup>
Zusammensetzung der Immunzellpopulationen (FU)	EFFT2010	Scharek-Tedin et al. 2015 <sup>[90]</sup>
Großes Blutbild	EFFT2010	
Gewicht der Tiere und Größe einzelner Organe	EFFT2010	Martin et al. 2012 <sup>[67]</sup>
Expression immunrelevanter Gene	EFST2010	Siepert et al. 2014 <sup>[95]</sup>
Körpertemperatur	EFST2010	Kreuzer et al. 2012 <sup>[59]</sup>
Vorkommen von Salmonellen	EFST2010	Kreuzer et al. 2012 <sup>[59]</sup>
Gewicht der Ferkel	EFST2010	Kreuzer et al. 2012 <sup>[59]</sup>
Konsistenz der Fäkalien	EFST2010	Kreuzer et al. 2012 <sup>[59]</sup>
Vorkommen von Anti-Salmonellen Antikörpern	EFST2010	Kreuzer et al. 2012 <sup>[59]</sup>
Zusammensetzung der intestinalen Mikrobiota	ZNFT2011	Starke et al. 2013 <sup>[102]</sup>
Konzentration bakterieller Metabolite	ZNFT2011	Starke et al. 2013 <sup>[102]</sup>
Konzentration von Zink im Verdauungstrakt	ZNFT2011	Starke et al. 2013 <sup>[102]</sup>
Vorkommen virulenter <i>Escherichia coli</i> Gene	ZNFT2011	Bednorz et al. 2013 <sup>[9]</sup>
Transport-Funktion der Darmschleimhaut	ZNFT2011	Gefeller et al. 2014 <sup>[35]</sup>
Zusammensetzung der Immunzellpopulationen (T-Zellen)	ZNFT2011	
Zusammensetzung der Immunzellpopulationen ( $\gamma\delta$ -T-Zellen und B-Zellen)	ZNFT2011	
Zusammensetzung der Immunzellpopulationen (IEL)	ZNFT2011	Liu et al. 2014 <sup>[65]</sup>
Aktivität der Verdauungsenzyme	ZNFT2011	Martin et al. 2013 <sup>[68]</sup>
Genexpression des Darmepithels	ZNFT2011	Martin et al. 2013 <sup>[68]</sup> , Liu et al. 2014 <sup>[65]</sup>
Morphologie der Darmschleimhaut	ZNFT2011	Liu et al. 2014 <sup>[65]</sup>

Fortsetzung auf der nächsten Seite

**Tabelle D.1 – Fortsetzung**

<b>Messung</b>	<b>Fütterungsversuch</b>	<b>Publikation</b>
Expression immunrelevanter Gene	ZNFT2011	
Großes Blutbild	ZNFT2011	
Gewicht der Tiere und Größe einzelner Organe	ZNFT2011	Martin et al. 2013 <sup>[68]</sup>
Konzentration von Metallen im Gewebe	ZNFT2011	Martin et al. 2013 <sup>[68]</sup>
Expression immunrelevanter Gene	ZNST2011	
Vorkommen von Salmonellen	ZNST2011	Janczyk et al. 2013 <sup>[51]</sup>
Gewicht der Ferkel	ZNST2011	Janczyk et al. 2013 <sup>[51]</sup>
Genexpression der Cytokine im intestinalen Gewebe	ZNTG2012	Chai et al. 2014 <sup>[18]</sup>
Quantifizierung der TGEV-Antikörper (ELISA)	ZNTG2012	Chai et al. 2014 <sup>[18]</sup>
Konsistenz der Fäkalien	ZNTG2012	Chai et al. 2014 <sup>[18]</sup>
Gewicht der Ferkel	ZNTG2012	Chai et al. 2014 <sup>[18]</sup>
Quantifizierung der PRRSV-Antikörper (ELISA)	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Quantifizierung IFN- $\gamma$ -sekretierender PBMC	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Charakterisierung mononukleärer Zellen des peripheren Blutes	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Körpertemperatur	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Quantifizierung der PRRSV-Viren	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Quantifizierung der PRRSV-Antikörper (Neutralisationstest)	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Gewichtszunahme	ZNPR2012	Chai et al. 2014 <sup>[17]</sup>
Quantifizierung der SIV-Antikörper (ELISA)	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Quantifizierung der SIV-Antikörper (HI-Assay)	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Vorkommen von Blutzellen	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Charakterisierung mononukleärer Zellen der Lunge	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Charakterisierung mononukleärer Zellen des peripheren Blutes	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Grad der Entzündung der Lunge	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Körpertemperatur	EZIV2013	Wang et al. 2014 <sup>[115]</sup>
Quantifizierung der Schweineinfluenza Viren	EZIV2013	Wang et al. 2014 <sup>[115]</sup>

# E. Merkmale des *EFFT2010* Fütterungsversuch

Tabelle E.1.: Zusammensetzung der intestinalen Mikrobiota

Nr.	Merkmalsgruppe	Beschreibung
1	OtuCol	Vorkommen von Bakterien der Gattungen <i>Bacteriodes</i> und <i>Prevotella</i> im Colon
2	OtuCol	Vorkommen von Bakterien der Gattung <i>Bifidobacterium</i> im Colon
3	OtuCol	Vorkommen von Bakterien der Spezies <i>Clostridium coccooides</i> im Colon
4	OtuCol	Vorkommen von Bakterien der Spezies <i>Clostridium Leptum</i> im Colon
5	OtuCol	Vorkommen von Bakterien der Spezies <i>Enterococcus faecalis</i> im Colon
6	OtuCol	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> im Colon
7	OtuCol	Vorkommen von Bakterien der Gattung <i>Enterobacter</i> im Colon
8	OtuCol	Vorkommen von Bakterien der Gattung <i>Enterococcus</i> im Colon
9	OtuCol	Vorkommen von Bakterien der Gattung <i>Eubacterium</i> im Colon
10	OtuCol	Vorkommen von Bakterien der Spezies <i>Lactobacillus acidophilus</i> im Colon
11	OtuCol	Vorkommen von Bakterien der Gattung <i>Lactobacillus</i> im Colon
12	OtuCol	Vorkommen von Bakterien der Spezies <i>Lactobacillus amylovorus</i> im Colon
13	OtuCol	Vorkommen von Bakterien der Spezies <i>Lactobacillus johnsonii</i> im Colon
14	OtuCol	Vorkommen von Bakterien der Spezies <i>Lactobacillus mucosae</i> im Colon
15	OtuCol	Vorkommen von Bakterien der Spezies <i>Lactobacillus reuteri</i> im Colon
16	OtuCol	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> NCIMB 10415 im Colon
17	OtuCol	Vorkommen von Bakterien der Gattung <i>Streptococcus</i> im Colon
18	Otulle	Vorkommen von Bakterien der Gattungen <i>Bacteriodes</i> und <i>Prevotella</i> im Ileum
19	Otulle	Vorkommen von Bakterien der Gattung <i>Bifidobacterium</i> im Colon
20	Otulle	Vorkommen von Bakterien der Spezies <i>Clostridium coccooides</i> im Ileum
21	Otulle	Vorkommen von Bakterien der Spezies <i>Clostridium Leptum</i> im Ileum
22	Otulle	Vorkommen von Bakterien der Spezies <i>Enterococcus faecalis</i> im Ileum
23	Otulle	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> im Ileum
24	Otulle	Vorkommen von Bakterien der Gattung <i>Enterobacter</i> im Colon
25	Otulle	Vorkommen von Bakterien der Gattung <i>Enterococcus</i> im Ileum
26	Otulle	Vorkommen von Bakterien der Gattung <i>Eubacterium</i> im Ileum
27	Otulle	Vorkommen von Bakterien der Spezies <i>Lactobacillus acidophilus</i> im Ileum
28	Otulle	Vorkommen von Bakterien der Gattung <i>Lactobacillus</i> im Ileum
29	Otulle	Vorkommen von Bakterien der Spezies <i>Lactobacillus amylovorus</i> im Ileum
30	Otulle	Vorkommen von Bakterien der Spezies <i>Lactobacillus johnsonii</i> im Ileum
31	Otulle	Vorkommen von Bakterien der Spezies <i>Lactobacillus mucosae</i> im Ileum
32	Otulle	Vorkommen von Bakterien der Spezies <i>Lactobacillus reuteri</i> im Ileum

Fortsetzung auf der nächsten Seite

**Tabelle E.1 – Fortsetzung**

Nr.	Merkmalsgruppe	Beschreibung
33	Otulle	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> NCIMB 10415 im Ileum
34	Otulle	Vorkommen von Bakterien der Gattung <i>Streptococcus</i> im Ileum
35	OtuJej	Vorkommen von Bakterien der Gattungen <i>Bacteriodes</i> und <i>Prevotella</i> im Jejunum
36	OtuJej	Vorkommen von Bakterien der Gattung <i>Bifidobacterium</i> im Jejunum
37	OtuJej	Vorkommen von Bakterien der Spezies <i>Clostridium coccooides</i> im Jejunum
38	OtuJej	Vorkommen von Bakterien der Spezies <i>Clostridium Leptum</i> im Jejunum
39	OtuJej	Vorkommen von Bakterien der Spezies <i>Enterococcus faecalis</i> im Jejunum
40	OtuJej	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> im Jejunum
41	OtuJej	Vorkommen von Bakterien der Gattung <i>Enterobacter</i> im Jejunum
42	OtuJej	Vorkommen von Bakterien der Gattung <i>Enterococcus</i> im Jejunum
43	OtuJej	Vorkommen von Bakterien der Gattung <i>Eubacterium</i> im Jejunum
44	OtuJej	Vorkommen von Bakterien der Spezies <i>Lactobacillus acidophilus</i> im Jejunum
45	OtuJej	Vorkommen von Bakterien der Gattung <i>Lactobacillus</i> im Jejunum
46	OtuJej	Vorkommen von Bakterien der Spezies <i>Lactobacillus amylovorus</i> im Jejunum
47	OtuJej	Vorkommen von Bakterien der Spezies <i>Lactobacillus johnsonii</i> im Jejunum
48	OtuJej	Vorkommen von Bakterien der Spezies <i>Lactobacillus mucosae</i> im Jejunum
49	OtuJej	Vorkommen von Bakterien der Spezies <i>Lactobacillus reuteri</i> im Jejunum
50	OtuJej	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> NCIMB 10415 im Jejunum
51	OtuJej	Vorkommen von Bakterien der Gattung <i>Streptococcus</i> im Jejunum
52	OtuSto	Vorkommen von Bakterien der Gattungen <i>Bacteriodes</i> und <i>Prevotella</i> im Magen
53	OtuSto	Vorkommen von Bakterien der Gattung <i>Bifidobacterium</i> im Magen
54	OtuSto	Vorkommen von Bakterien der Spezies <i>Clostridium coccooides</i> im Magen
55	OtuSto	Vorkommen von Bakterien der Spezies <i>Clostridium Leptum</i> im Magen
56	OtuSto	Vorkommen von Bakterien der Spezies <i>Enterococcus faecalis</i> im Magen
57	OtuSto	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> im Magen
58	OtuSto	Vorkommen von Bakterien der Gattung <i>Enterobacter</i> im Magen
59	OtuSto	Vorkommen von Bakterien der Gattung <i>Enterococcus</i> im Magen
60	OtuSto	Vorkommen von Bakterien der Gattung <i>Eubacterium</i> im Magen
61	OtuSto	Vorkommen von Bakterien der Spezies <i>Lactobacillus acidophilus</i> im Magen
62	OtuSto	Vorkommen von Bakterien der Gattung <i>Lactobacillus</i> im Magen
63	OtuSto	Vorkommen von Bakterien der Spezies <i>Lactobacillus amylovorus</i> im Magen
64	OtuSto	Vorkommen von Bakterien der Spezies <i>Lactobacillus johnsonii</i> im Magen
65	OtuSto	Vorkommen von Bakterien der Spezies <i>Lactobacillus mucosae</i> im Magen
66	OtuSto	Vorkommen von Bakterien der Spezies <i>Lactobacillus reuteri</i> im Magen
67	OtuSto	Vorkommen von Bakterien der Spezies <i>Enterococcus faecium</i> NCIMB 10415 im Magen
68	OtuSto	Vorkommen von Bakterien der Gattung <i>Streptococcus</i> im Magen



**Tabelle E.2.:** Konzentration bakterieller Metabolite

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	MetabCol	Konzentration von Aminen im Colon
2	MetabCol	Konzentration von Ammonium im Colon
3	MetabCol	Konzentration von Cadaverin im Colon
4	MetabCol	Konzentration von D-Milchsäure im Colon
5	MetabCol	Konzentration von Essigsäure im Colon
6	MetabCol	Konzentration von Histamin im Colon
7	MetabCol	Konzentration von i-Buttersäure im Colon
8	MetabCol	Konzentration von i-Valeriansäure im Colon
9	MetabCol	Konzentration von L-Milchsäure im Colon
10	MetabCol	Konzentration von n-Buttersäuren im Colon
11	MetabCol	Konzentration von n-Valeriansäure im Colon
12	MetabCol	Konzentration von Propionsäure im Colon
13	MetabCol	Konzentration von Putrescin im Colon
14	MetabCol	Konzentration von Spermidin im Colon
15	MetabCol	Konzentration von Spermin im Colon
16	MetabCol	Konzentration von Spermin im Colon
17	MetabCol	Konzentration von Tyramin im Colon
18	Metablle	Konzentration von Aminen im Ileum
19	Metablle	Konzentration von Ammonium im Ileum
20	Metablle	Konzentration von Cadaverin im Ileum
21	Metablle	Konzentration von D-Milchsäure im Ileum
22	Metablle	Konzentration von Essigsäure im Ileum
23	Metablle	Konzentration von Histamin im Ileum
24	Metablle	Konzentration von i-Buttersäure im Ileum
25	Metablle	Konzentration von i-Valeriansäure im Ileum
26	Metablle	Konzentration von L-Milchsäure im Ileum
27	Metablle	Konzentration von n-Buttersäuren im Ileum
28	Metablle	Konzentration von n-Valeriansäure im Ileum
29	Metablle	Konzentration von Propionsäure im Ileum
30	Metablle	Konzentration von Putrescin im Ileum
31	Metablle	Konzentration von Spermidin im Ileum
32	Metablle	Konzentration von Spermin im Ileum
33	Metablle	Konzentration von Tyramin im Ileum
34	MetabJej	Konzentration von Aminen im Jejunum
35	MetabJej	Konzentration von Ammonium im Jejunum
36	MetabJej	Konzentration von Cadaverin im Jejunum
37	MetabJej	Konzentration von D-Milchsäure im Jejunum
38	MetabJej	Konzentration von Essigsäure im Jejunum
39	MetabJej	Konzentration von Histamin im Jejunum
40	MetabJej	Konzentration von i-Buttersäure im Jejunum
41	MetabJej	Konzentration von i-Valeriansäure im Jejunum
42	MetabJej	Konzentration von L-Milchsäure im Jejunum
43	MetabJej	Konzentration von n-Buttersäuren im Jejunum
44	MetabJej	Konzentration von n-Valeriansäure im Jejunum
45	MetabJej	Konzentration von Propionsäure im Jejunum
46	MetabJej	Konzentration von Putrescin im Jejunum

*Fortsetzung auf der nächsten Seite*

**Tabelle E.2 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
47	MetabJej	Konzentration von Spermidin im Jejunum
48	MetabJej	Konzentration von Spermin im Jejunum
49	MetabJej	Konzentration von Tyramin im Jejunum
50	MetabSto	Konzentration von Aminen im Magen
51	MetabSto	Konzentration von Ammonium im Magen
52	MetabSto	Konzentration von Cadaverin im Magen
53	MetabSto	Konzentration von D-Milchsäure im Magen
54	MetabSto	Konzentration von Essigsäure im Magen
55	MetabSto	Konzentration von Histamin im Magen
56	MetabSto	Konzentration von i-Buttersäure im Magen
57	MetabSto	Konzentration von i-Valeriansäure im Magen
58	MetabSto	Konzentration von L-Milchsäure im Magen
59	MetabSto	Konzentration von n-Buttersäuren im Magen
60	MetabSto	Konzentration von n-Valeriansäure im Magen
61	MetabSto	Konzentration von Propionsäure im Magen
62	MetabSto	Konzentration von Putrescin im Magen
63	MetabSto	Konzentration von Spermidin im Magen
64	MetabSto	Konzentration von Tyramin im Magen

**Tabelle E.3.: Vorkommen virulenter *Escherichia coli* Gene**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	EcoliColDig	Vorkommen des Gens afa.draBC im Inhalt des Colons
2	EcoliColDig	Vorkommen des Gens astA(East1) im Inhalt des Colons
3	EcoliColDig	Vorkommen des Gens chuA im Inhalt des Colons
4	EcoliColDig	Vorkommen des Gens cnf1 im Inhalt des Colons
5	EcoliColDig	Vorkommen des Gens colV(cvi/cva) im Inhalt des Colons
6	EcoliColDig	Vorkommen des Gens csgA im Inhalt des Colons
7	EcoliColDig	Vorkommen des Gens cvaC im Inhalt des Colons
8	EcoliColDig	Vorkommen des Gens escv im Inhalt des Colons
9	EcoliColDig	Vorkommen des Gens est-Ia im Inhalt des Colons
10	EcoliColDig	Vorkommen des Gens est-II im Inhalt des Colons
11	EcoliColDig	Vorkommen des Gens etsB im Inhalt des Colons
12	EcoliColDig	Vorkommen des Gens etsC im Inhalt des Colons
13	EcoliColDig	Vorkommen des Gens feoB im Inhalt des Colons
14	EcoliColDig	Vorkommen des Gens fimC im Inhalt des Colons
15	EcoliColDig	Vorkommen des Gens focG im Inhalt des Colons
16	EcoliColDig	Vorkommen des Gens fyuA im Inhalt des Colons
17	EcoliColDig	Vorkommen des Gens gimB im Inhalt des Colons
18	EcoliColDig	Vorkommen des Gens hlyC im Inhalt des Colons
19	EcoliColDig	Vorkommen des Gens hlyF im Inhalt des Colons
20	EcoliColDig	Vorkommen des Gens hrlA/hek im Inhalt des Colons
21	EcoliColDig	Vorkommen des Gens iha.cds im Inhalt des Colons
22	EcoliColDig	Vorkommen des Gens ireA im Inhalt des Colons
23	EcoliColDig	Vorkommen des Gens iroN im Inhalt des Colons

*Fortsetzung auf der nächsten Seite*

**Tabelle E.3 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
24	EcoliColDig	Vorkommen des Gens irp2 im Inhalt des Colons
25	EcoliColDig	Vorkommen des Gens iss im Inhalt des Colons
26	EcoliColDig	Vorkommen des Gens iucD im Inhalt des Colons
27	EcoliColDig	Vorkommen des Gens iutA im Inhalt des Colons
28	EcoliColDig	Vorkommen des Gens kpsMTII im Inhalt des Colons
29	EcoliColDig	Vorkommen des Gens mat im Inhalt des Colons
30	EcoliColDig	Vorkommen des Gens neuC im Inhalt des Colons
31	EcoliColDig	Vorkommen des Gens ompA im Inhalt des Colons
32	EcoliColDig	Vorkommen des Gens ompT im Inhalt des Colons
33	EcoliColDig	Vorkommen des Gens papC im Inhalt des Colons
34	EcoliColDig	Vorkommen des Gens papEF im Inhalt des Colons
35	EcoliColDig	Vorkommen des Gens papGII,III im Inhalt des Colons
36	EcoliColDig	Vorkommen des Gens pic im Inhalt des Colons
37	EcoliColDig	Vorkommen des Gens pks im Inhalt des Colons
38	EcoliColDig	Vorkommen des Gens Rpai(malX) im Inhalt des Colons
39	EcoliColDig	Vorkommen des Gens sat im Inhalt des Colons
40	EcoliColDig	Vorkommen des Gens sfa.foc im Inhalt des Colons
41	EcoliColDig	Vorkommen des Gens sfaS im Inhalt des Colons
42	EcoliColDig	Vorkommen des Gens sitA im Inhalt des Colons
43	EcoliColDig	Vorkommen des Gens sitD(chromosomal) im Inhalt des Colons
44	EcoliColDig	Vorkommen des Gens sitD(episomal) im Inhalt des Colons
45	EcoliColDig	Vorkommen des Gens stx2 im Inhalt des Colons
46	EcoliColDig	Vorkommen des Gens tia im Inhalt des Colons
47	EcoliColDig	Vorkommen des Gens traT im Inhalt des Colons
48	EcoliColDig	Vorkommen des Gens tsh im Inhalt des Colons
49	EcoliColDig	Vorkommen des Gens vat im Inhalt des Colons
50	EcoliColDig	Vorkommen des Gens yqi im Inhalt des Colons
51	EcoliColMuc	Vorkommen des Gens afa.draBC an der Schleimhaut des Colons
52	EcoliColMuc	Vorkommen des Gens astA(East1) an der Schleimhaut des Colons
53	EcoliColMuc	Vorkommen des Gens chuA an der Schleimhaut des Colons
54	EcoliColMuc	Vorkommen des Gens cnf1 an der Schleimhaut des Colons
55	EcoliColMuc	Vorkommen des Gens colV(cvi/cva) an der Schleimhaut des Colons
56	EcoliColMuc	Vorkommen des Gens csgA an der Schleimhaut des Colons
57	EcoliColMuc	Vorkommen des Gens cvaC an der Schleimhaut des Colons
58	EcoliColMuc	Vorkommen des Gens escv an der Schleimhaut des Colons
59	EcoliColMuc	Vorkommen des Gens est-II an der Schleimhaut des Colons
60	EcoliColMuc	Vorkommen des Gens etsB an der Schleimhaut des Colons
61	EcoliColMuc	Vorkommen des Gens etsC an der Schleimhaut des Colons
62	EcoliColMuc	Vorkommen des Gens feoB an der Schleimhaut des Colons
63	EcoliColMuc	Vorkommen des Gens fimC an der Schleimhaut des Colons
64	EcoliColMuc	Vorkommen des Gens focG an der Schleimhaut des Colons
65	EcoliColMuc	Vorkommen des Gens fyuA an der Schleimhaut des Colons
66	EcoliColMuc	Vorkommen des Gens hlyA an der Schleimhaut des Colons
67	EcoliColMuc	Vorkommen des Gens hlyC an der Schleimhaut des Colons
68	EcoliColMuc	Vorkommen des Gens hlyF an der Schleimhaut des Colons
69	EcoliColMuc	Vorkommen des Gens hrlA/hek an der Schleimhaut des Colons
70	EcoliColMuc	Vorkommen des Gens iha.cds an der Schleimhaut des Colons

*Fortsetzung auf der nächsten Seite*

**Tabelle E.3 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
71	EcoliColMuc	Vorkommen des Gens ireA an der Schleimhaut des Colons
72	EcoliColMuc	Vorkommen des Gens iroN an der Schleimhaut des Colons
73	EcoliColMuc	Vorkommen des Gens irp2 an der Schleimhaut des Colons
74	EcoliColMuc	Vorkommen des Gens iss an der Schleimhaut des Colons
75	EcoliColMuc	Vorkommen des Gens iucD an der Schleimhaut des Colons
76	EcoliColMuc	Vorkommen des Gens iutA an der Schleimhaut des Colons
77	EcoliColMuc	Vorkommen des Gens kpsMTII an der Schleimhaut des Colons
78	EcoliColMuc	Vorkommen des Gens mat an der Schleimhaut des Colons
79	EcoliColMuc	Vorkommen des Gens neuC an der Schleimhaut des Colons
80	EcoliColMuc	Vorkommen des Gens ompA an der Schleimhaut des Colons
81	EcoliColMuc	Vorkommen des Gens ompT an der Schleimhaut des Colons
82	EcoliColMuc	Vorkommen des Gens papAH an der Schleimhaut des Colons
83	EcoliColMuc	Vorkommen des Gens papC an der Schleimhaut des Colons
84	EcoliColMuc	Vorkommen des Gens papEF an der Schleimhaut des Colons
85	EcoliColMuc	Vorkommen des Gens papGII,III an der Schleimhaut des Colons
86	EcoliColMuc	Vorkommen des Gens pic an der Schleimhaut des Colons
87	EcoliColMuc	Vorkommen des Gens pks an der Schleimhaut des Colons
88	EcoliColMuc	Vorkommen des Gens Rpai(malX) an der Schleimhaut des Colons
89	EcoliColMuc	Vorkommen des Gens sfa.foc an der Schleimhaut des Colons
90	EcoliColMuc	Vorkommen des Gens sitA an der Schleimhaut des Colons
91	EcoliColMuc	Vorkommen des Gens sitD(chromosomal) an der Schleimhaut des Colons
92	EcoliColMuc	Vorkommen des Gens sitD(episomal) an der Schleimhaut des Colons
93	EcoliColMuc	Vorkommen des Gens traT an der Schleimhaut des Colons
94	EcoliColMuc	Vorkommen des Gens tsh an der Schleimhaut des Colons
95	EcoliColMuc	Vorkommen des Gens vat an der Schleimhaut des Colons
96	EcoliColMuc	Vorkommen des Gens yqi an der Schleimhaut des Colons
97	EcoliFae	Vorkommen des Gens afa.draBC in den Fäkalien
98	EcoliFae	Vorkommen des Gens astA(East1) in den Fäkalien
99	EcoliFae	Vorkommen des Gens chuA in den Fäkalien
100	EcoliFae	Vorkommen des Gens cnf1 in den Fäkalien
101	EcoliFae	Vorkommen des Gens colV(cvi/cva) in den Fäkalien
102	EcoliFae	Vorkommen des Gens csgA in den Fäkalien
103	EcoliFae	Vorkommen des Gens cvaC in den Fäkalien
104	EcoliFae	Vorkommen des Gens escv in den Fäkalien
105	EcoliFae	Vorkommen des Gens est-II in den Fäkalien
106	EcoliFae	Vorkommen des Gens etsB in den Fäkalien
107	EcoliFae	Vorkommen des Gens etsC in den Fäkalien
108	EcoliFae	Vorkommen des Gens feoB in den Fäkalien
109	EcoliFae	Vorkommen des Gens fimC in den Fäkalien
110	EcoliFae	Vorkommen des Gens focG in den Fäkalien
111	EcoliFae	Vorkommen des Gens fyuA in den Fäkalien
112	EcoliFae	Vorkommen des Gens gimB in den Fäkalien
113	EcoliFae	Vorkommen des Gens hlyA in den Fäkalien
114	EcoliFae	Vorkommen des Gens hlyC in den Fäkalien
115	EcoliFae	Vorkommen des Gens hlyF in den Fäkalien
116	EcoliFae	Vorkommen des Gens hrlA/hek in den Fäkalien
117	EcoliFae	Vorkommen des Gens ireA in den Fäkalien

*Fortsetzung auf der nächsten Seite*

**Tabelle E.3 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
118	EcoliFae	Vorkommen des Gens <i>iroN</i> in den Fäkalien
119	EcoliFae	Vorkommen des Gens <i>irp2</i> in den Fäkalien
120	EcoliFae	Vorkommen des Gens <i>iss</i> in den Fäkalien
121	EcoliFae	Vorkommen des Gens <i>iucD</i> in den Fäkalien
122	EcoliFae	Vorkommen des Gens <i>iutA</i> in den Fäkalien
123	EcoliFae	Vorkommen des Gens <i>kpsMTII</i> in den Fäkalien
124	EcoliFae	Vorkommen des Gens <i>mat</i> in den Fäkalien
125	EcoliFae	Vorkommen des Gens <i>neuC</i> in den Fäkalien
126	EcoliFae	Vorkommen des Gens <i>ompA</i> in den Fäkalien
127	EcoliFae	Vorkommen des Gens <i>ompT</i> in den Fäkalien
128	EcoliFae	Vorkommen des Gens <i>papEF</i> in den Fäkalien
129	EcoliFae	Vorkommen des Gens <i>pic</i> in den Fäkalien
130	EcoliFae	Vorkommen des Gens <i>pks</i> in den Fäkalien
131	EcoliFae	Vorkommen des Gens <i>Rpai(malX)</i> in den Fäkalien
132	EcoliFae	Vorkommen des Gens <i>sat</i> in den Fäkalien
133	EcoliFae	Vorkommen des Gens <i>sfa.foc</i> in den Fäkalien
134	EcoliFae	Vorkommen des Gens <i>sitA</i> in den Fäkalien
135	EcoliFae	Vorkommen des Gens <i>sitD(chromosomal)</i> in den Fäkalien
136	EcoliFae	Vorkommen des Gens <i>sitD(episomal)</i> in den Fäkalien
137	EcoliFae	Vorkommen des Gens <i>stx2</i> in den Fäkalien
138	EcoliFae	Vorkommen des Gens <i>traT</i> in den Fäkalien
139	EcoliFae	Vorkommen des Gens <i>tsh</i> in den Fäkalien
140	EcoliFae	Vorkommen des Gens <i>vat</i> in den Fäkalien
141	EcoliFae	Vorkommen des Gens <i>yqi</i> in den Fäkalien

**Tabelle E.4.: Transport-Eigenschaften der Darmschleimhaut**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	EpiJej	Kurzschlussstrom ( <i>I<sub>sc</sub></i> ) nach der Zugabe von 1 mmol/L Glucose
2	EpiJej	Kurzschlussstrom ( <i>I<sub>sc</sub></i> ) nach der Zugabe von 4 mmol/L Glucose
3	EpiJej	Kurzschlussstrom ( <i>I<sub>sc</sub></i> ) nach der Zugabe von 12 mmol/L Glucose
4	EpiJej	Kurzschlussstrom ( <i>I<sub>sc</sub></i> ) nach der Zugabe von 0.01 $\mu$ mol/L PGE
5	EpiJej	Kurzschlussstrom ( <i>I<sub>sc</sub></i> ) nach der Zugabe von 0.5 $\mu$ mol/L PGE
6	EpiJej	Kurzschlussstrom ( <i>I<sub>sc</sub></i> ) nach der Zugabe von 5 $\mu$ mol/L PGE

**Tabelle E.5.: Barrieren-Eigenschaften der Darmschleimhaut**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	EpiCol	Kurzschlußstrom gemessen über Mukosa (Colon)
2	EpiCol	Ohmscher epithelialer Widerstand gemessen über Mukosa (Colon)
3	EpiCol	Ohmscher epithelialer Widerstand gemessen über Mukosa (Colon)
4	EpiCol	Ohmscher subepithelialer Widerstand gemessen über Mukosa (Colon)

*Fortsetzung auf der nächsten Seite*

**Tabelle E.5 – Fortsetzung**

Nr.	Merkmalsgruppe	Beschreibung
5	EpiCol	Ohmscher subepithelialer Widerstand gemessen über Mukosa (Colon)
6	EpiCol	Ohmscher Widerstand gemessen über Mukosa (Colon)
7	EpiCol	Ohmscher Widerstand gemessen über Mukosa (Colon)
8	EpiCol	Ohmscher Widerstand gemessen über Mukosa (Colon)
9	EpiCol	Permeabilität der Mukosa für 3H-Mannitol (182 Da) (Colon)
10	EpiCol	Permeabilität der Mukosa für FITC-Dextran4000 (4000 Da) (Colon)
11	EpiCol	Permeabilität der Mukosa für HRP (44000 Da) (Colon)
12	EpiCol	Stimulierbare Aktivität des epithelialen Natriumkanals im Colon (Colon)
13	EpiCol	Verhältnis der mukosalen Permeabilitäten von Natrium und Chlorid (Colon)
14	Epille	Kurzschlußstrom gemessen über Mukosa (Ileum)
15	Epille	Ohmscher epithelialer Widerstand gemessen über Mukosa (Ileum)
16	Epille	Ohmscher subepithelialer Widerstand gemessen über Mukosa (Ileum)
17	Epille	Ohmscher Widerstand gemessen über Mukosa (Ileum)
18	Epille	Ohmscher Widerstand gemessen über Mukosa (Ileum)
19	Epille	Permeabilität der Mukosa für 3H-Mannitol (182 Da) (Ileum)
20	Epille	Permeabilität der Mukosa für FITC-Dextran4000 (4000 Da) (Ileum)
21	Epille	Permeabilität der Mukosa für HRP (44000 Da) (Ileum)
22	Epille	Verhältnis der mukosalen Permeabilitäten von Natrium und Chlorid (Ileum)
23	EpiJej	Kurzschlußstrom gemessen über Mukosa (Jejunum)
24	EpiJej	Ohmscher epithelialer Widerstand gemessen über Mukosa (Jejunum)
25	EpiJej	Ohmscher subepithelialer Widerstand gemessen über Mukosa (Jejunum)
26	EpiJej	Ohmscher Widerstand gemessen über Jejunum Schleimhaut (Jejunum)
27	EpiJej	Ohmscher Widerstand gemessen über Mukosa (Jejunum)
28	EpiJej	Permeabilität der Mukosa für 3H-Mannitol (182 Da) (Jejunum)
29	EpiJej	Permeabilität der Mukosa für FITC-Dextran4000 (4000 Da) (Jejunum)
30	EpiJej	Permeabilität der Mukosa für HRP (44000 Da) (Jejunum)
31	EpiJej	Verhältnis der mukosalen Permeabilitäten von Natrium und Chlorid (Jejunum)

**Tabelle E.6.: Morphologie der Darmschleimhaut**

Nr.	Merkmalsgruppe	Beschreibung
1	MorphJej	Krypten-Tiefe
2	MorphJej	Verhältnis von Zotten zu Krypten
3	MorphJej	Zotten-Länge

**Tabelle E.7.: Aktivität der Verdauungsenzyme**

Nr.	Merkmalsgruppe	Beschreibung
1	BbmAct	Aktivität des Enzyms Aminopeptidase
2	BbmAct	Aktivität des Enzyms <i>Intestinal Alkaline Phosphatase</i>
3	BbmAct	Aktivität des Enzyms Lactase

*Fortsetzung auf der nächsten Seite*

**Tabelle E.7 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
4	BbmAct	Aktivität des Enzyms Maltase
5	BbmAct	Aktivität des Enzyms Saccharase

**Tabelle E.8.: Genexpression des Darmepithels**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	MarkerExpr	Expression des Gens ALPI
2	MarkerExpr	Expression des Gens AMPN
3	MarkerExpr	Expression des Gens CASP3
4	MarkerExpr	Expression des Gens LCT
5	MarkerExpr	Expression des Gens PCNA
6	MarkerExpr	Expression des Gens SI
7	MarkerExpr	Expression des Gens SLC5A1

**Tabelle E.9.: Konzentration der Zytokine im Blut**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	CytoBlo	Konzentration des Proteins IFN- $\gamma$ im Blut
2	CytoBlo	Konzentration des Proteins IL-6 im Blut
3	CytoBlo	Konzentration des Proteins IL-8 im Blut
4	CytoBlo	Konzentration des Proteins IL-13 im Blut
5	CytoBlo	Konzentration des Proteins IL-17 im Blut
6	CytoBlo	Konzentration des Proteins TGF- $\beta$ im Blut

**Tabelle E.10.: Expression immunrelevanter Gene**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	IsExprBlo	Expression des Gens IL10 in Blutzellen
2	IsExprBlo	Expression des Gens IL8 in Blutzellen
3	IsExprBlo	Expression des Gens NOD2 in Blutzellen
4	IsExprIlePP	Expression des Gens CD86 in Peyer-Plaques des Ileums
5	IsExprIlePP	Expression des Gens CTLA4 in Peyer-Plaques des Ileums
6	IsExprIlePP	Expression des Gens IL10 in Peyer-Plaques des Ileums
7	IsExprIlePP	Expression des Gens IL8 in Peyer-Plaques des Ileums
8	IsExprIlePP	Expression des Gens NOD1 in Peyer-Plaques des Ileums
9	IsExprIlePP	Expression des Gens NOD2 in Peyer-Plaques des Ileums
10	IsExprIlePP	Expression des Gens PGLYRP1 in Peyer-Plaques des Ileums
11	IsExprIlePP	Expression des Gens PGLYRP2A in Peyer-Plaques des Ileums
12	IsExprIlePP	Expression des Gens PGLYRP2B in Peyer-Plaques des Ileums
13	IsExprIlePP	Expression des Gens SIRT1 in Peyer-Plaques des Ileums

*Fortsetzung auf der nächsten Seite*

**Tabelle E.10 – Fortsetzung**

Nr.	Merkmalsgruppe	Beschreibung
14	IsExprJejLN	Expression des Gens CD86 in Lymphknoten des Jejunums
15	IsExprJejLN	Expression des Gens CTLA4 in Lymphknoten des Jejunums
16	IsExprJejLN	Expression des Gens IL10 in Lymphknoten des Jejunums
17	IsExprJejLN	Expression des Gens IL8 in Lymphknoten des Jejunums
18	IsExprJejLN	Expression des Gens PGLYRP1 in Lymphknoten des Jejunums
19	IsExprJejLN	Expression des Gens PGLYRP2B in Lymphknoten des Jejunums
20	IsExprJejLN	Expression des Gens SIRT1 in Lymphknoten des Jejunums
21	IsExprJejPP	Expression des Gens CD86 in Peyer-Plaques des Jejunums
22	IsExprJejPP	Expression des Gens CTLA4 in Peyer-Plaques des Jejunums
23	IsExprJejPP	Expression des Gens IL10 in Peyer-Plaques des Jejunums
24	IsExprJejPP	Expression des Gens IL8 in Peyer-Plaques des Jejunums
25	IsExprJejPP	Expression des Gens NOD1 in Peyer-Plaques des Jejunums
26	IsExprJejPP	Expression des Gens NOD2 in Peyer-Plaques des Jejunums
27	IsExprJejPP	Expression des Gens PGLYRP1 in Peyer-Plaques des Jejunums
28	IsExprJejPP	Expression des Gens PGLYRP2A in Peyer-Plaques des Jejunums
29	IsExprJejPP	Expression des Gens PGLYRP2B in Peyer-Plaques des Jejunums
30	IsExprJejPP	Expression des Gens SIRT1 in Peyer-Plaques des Jejunums
31	IsExprLiv	Expression des Gens PGLYRP2A in der Leber
32	IsExprLiv	Expression des Gens PGLYRP2B in der Leber
33	IsExprSpln	Expression des Gens CD86 in der Milz
34	IsExprSpln	Expression des Gens CTLA4 in der Milz
35	IsExprSpln	Expression des Gens IL10 in der Milz
36	IsExprSpln	Expression des Gens IL8 in der Milz
37	IsExprSpln	Expression des Gens NOD1 in der Milz
38	IsExprSpln	Expression des Gens NOD2 in der Milz
39	IsExprSpln	Expression des Gens PGLYRP2B in der Milz

**Tabelle E.11.: Zusammensetzung der Immunzellpopulationen**

Nr.	Merkmalsgruppe	Beschreibung
1	LymphoBlo	Vorkommen von CD4 <sup>-</sup> CD8a <sup>-</sup> Zellen im Blut
2	LymphoBlo	Vorkommen von CD4 <sup>-</sup> CD8a <sup>med</sup> Zellen im Blut
3	LymphoBlo	Vorkommen von CD4 <sup>-</sup> CD8a <sup>+</sup> Zellen im Blut
4	LymphoBlo	Vorkommen von CD4 <sup>+</sup> Zellen im Blut
5	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD25 <sup>high</sup> Zellen im Blut
6	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD25 <sup>-</sup> Zellen im Blut
7	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD25 <sup>med</sup> Zellen im Blut
8	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD8a <sup>-</sup> Zellen im Blut
9	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD8a <sup>med</sup> Zellen im Blut
10	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD8a <sup>+</sup> Zellen im Blut
11	LymphoBlo	Vorkommen von $\gamma\delta$ TCR <sup>high</sup> Zellen im Blut
12	LymphoBlo	Vorkommen von $\gamma\delta$ TCR <sup>-</sup> Zellen im Blut
13	LymphoBlo	Vorkommen von $\gamma\delta$ TCR <sup>med</sup> Zellen im Blut
14	LymphoBlo	Vorkommen von IgM <sup>high</sup> Zellen im Blut
15	LymphoBlo	Vorkommen von IgM <sup>-</sup> Zellen im Blut

*Fortsetzung auf der nächsten Seite*



**Tabelle E.11 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
16	LymphoBlo	Vorkommen von IgM <sup>med</sup> Zellen im Blut
17	LympholleLN	Vorkommen von CD4 <sup>-</sup> CD8a <sup>-</sup> Zellen in Lymphknoten (Ileum)
18	LympholleLN	Vorkommen von CD4 <sup>-</sup> CD8a <sup>med</sup> Zellen in Lymphknoten (Ileum)
19	LympholleLN	Vorkommen von CD4 <sup>-</sup> CD8a <sup>+</sup> Zellen in Lymphknoten (Ileum)
20	LympholleLN	Vorkommen von CD4 <sup>+</sup> CD25 <sup>high</sup> Zellen in Lymphknoten (Ileum)
21	LympholleLN	Vorkommen von CD4 <sup>+</sup> CD25 <sup>-</sup> Zellen in Lymphknoten (Ileum)
22	LympholleLN	Vorkommen von CD4 <sup>+</sup> CD25 <sup>med</sup> Zellen in Lymphknoten (Ileum)
23	LympholleLN	Vorkommen von CD4 <sup>+</sup> CD8a <sup>-</sup> Zellen in Lymphknoten (Ileum)
24	LympholleLN	Vorkommen von CD4 <sup>+</sup> CD8a <sup>med</sup> Zellen in Lymphknoten (Ileum)
25	LympholleLN	Vorkommen von CD4 <sup>+</sup> CD8a <sup>+</sup> Zellen in Lymphknoten (Ileum)
26	LympholleLN	Vorkommen von CD4 <sup>+</sup> Zellen in Lymphknoten (Ileum)
27	LympholleLN	Vorkommen von $\gamma\delta$ TCR <sup>high</sup> Zellen in Lymphknoten (Ileum)
28	LympholleLN	Vorkommen von $\gamma\delta$ TCR <sup>-</sup> Zellen in Lymphknoten (Ileum)
29	LympholleLN	Vorkommen von $\gamma\delta$ TCR <sup>med</sup> Zellen in Lymphknoten (Ileum)
30	LympholleLN	Vorkommen von IgM <sup>high</sup> Zellen in Lymphknoten (Ileum)
31	LympholleLN	Vorkommen von IgM <sup>-</sup> Zellen in Lymphknoten (Ileum)
32	LympholleLN	Vorkommen von IgM <sup>med</sup> Zellen in Lymphknoten (Ileum)
33	LymphollePP	Vorkommen von CD4 <sup>-</sup> CD8a <sup>-</sup> Zellen in Peyers Patches (Ileum)
34	LymphollePP	Vorkommen von CD4 <sup>-</sup> CD8a <sup>med</sup> Zellen in Peyers Patches (Ileum)
35	LymphollePP	Vorkommen von CD4 <sup>-</sup> CD8a <sup>+</sup> Zellen in Peyers Patches (Ileum)
36	LymphollePP	Vorkommen von CD4 <sup>+</sup> CD25 <sup>high</sup> Zellen in Peyers Patches (Ileum)
37	LymphollePP	Vorkommen von CD4 <sup>+</sup> CD25 <sup>-</sup> Zellen in Peyers Patches (Ileum)
38	LymphollePP	Vorkommen von CD4 <sup>+</sup> CD25 <sup>med</sup> Zellen in Peyers Patches (Ileum)
39	LymphollePP	Vorkommen von CD4 <sup>+</sup> CD8a <sup>-</sup> Zellen in Peyers Patches (Ileum)
40	LymphollePP	Vorkommen von CD4 <sup>+</sup> CD8a <sup>med</sup> Zellen in Peyers Patches (Ileum)
41	LymphollePP	Vorkommen von CD4 <sup>+</sup> CD8a <sup>+</sup> Zellen in Peyers Patches (Ileum)
42	LymphollePP	Vorkommen von CD4 <sup>+</sup> Zellen in Peyers Patches (Ileum)
43	LymphollePP	Vorkommen von $\gamma\delta$ TCR <sup>high</sup> Zellen in Peyers Patches (Ileum)
44	LymphollePP	Vorkommen von $\gamma\delta$ TCR <sup>-</sup> Zellen in Peyers Patches (Ileum)
45	LymphollePP	Vorkommen von $\gamma\delta$ TCR <sup>med</sup> Zellen in Peyers Patches (Ileum)
46	LymphollePP	Vorkommen von IgM <sup>high</sup> Zellen in Peyers Patches (Ileum)
47	LymphollePP	Vorkommen von IgM <sup>-</sup> Zellen in Peyers Patches (Ileum)
48	LymphollePP	Vorkommen von IgM <sup>med</sup> Zellen in Peyers Patches (Ileum)
49	LymphoJejLN	Vorkommen von CD4 <sup>-</sup> CD8a <sup>-</sup> Zellen in Lymphknoten (Jejunum)
50	LymphoJejLN	Vorkommen von CD4 <sup>-</sup> CD8a <sup>med</sup> Zellen in Lymphknoten (Jejunum)
51	LymphoJejLN	Vorkommen von CD4 <sup>-</sup> CD8a <sup>+</sup> Zellen in Lymphknoten (Jejunum)
52	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> CD25 <sup>high</sup> Zellen in Lymphknoten (Jejunum)
53	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> CD25 <sup>-</sup> Zellen in Lymphknoten (Jejunum)
54	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> CD25 <sup>med</sup> Zellen in Lymphknoten (Jejunum)
55	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> CD8a <sup>-</sup> Zellen in Lymphknoten (Jejunum)
56	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> CD8a <sup>med</sup> Zellen in Lymphknoten (Jejunum)
57	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> CD8a <sup>+</sup> Zellen in Lymphknoten (Jejunum)
58	LymphoJejLN	Vorkommen von CD4 <sup>+</sup> Zellen in Lymphknoten (Jejunum)
59	LymphoJejLN	Vorkommen von $\gamma\delta$ TCR <sup>high</sup> Zellen in Lymphknoten (Jejunum)
60	LymphoJejLN	Vorkommen von $\gamma\delta$ TCR <sup>-</sup> Zellen in Lymphknoten (Jejunum)
61	LymphoJejLN	Vorkommen von $\gamma\delta$ TCR <sup>med</sup> Zellen in Lymphknoten (Jejunum)
62	LymphoBlo	Vorkommen von $\gamma\delta$ <sup>+</sup> CD5 <sup>-</sup> Zellen im Blut

*Fortsetzung auf der nächsten Seite*

**Tabelle E.11 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
63	LymphoBlo	Vorkommen von $\gamma\delta^+$ CD8 <sup>-</sup> Zellen im Blut
64	LymphoBlo	Vorkommen von CD14 <sup>+</sup> CD16 <sup>+</sup> Zellen im Blut
65	LymphoBlo	Vorkommen von CD16 <sup>+</sup> CD14 <sup>-</sup> Zellen im Blut
66	LymphoBlo	Vorkommen von CD2 <sup>+</sup> CD5 <sup>-</sup> Zellen im Blut
67	LymphoBlo	Vorkommen von CD2 <sup>+</sup> CD5 <sup>+</sup> Zellen im Blut
68	LymphoBlo	Vorkommen von CD21 <sup>+</sup> MHCII <sup>-</sup> Zellen im Blut
69	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD8 <sup>-</sup> Zellen im Blut
70	LymphoBlo	Vorkommen von CD4 <sup>+</sup> CD8 <sup>+</sup> Zellen im Blut
71	LymphoBlo	Vorkommen von CD5 <sup>+</sup> $\gamma\delta^-$ Zellen im Blut
72	LymphoBlo	Vorkommen von CD5 <sup>+</sup> $\gamma\delta^+$ Zellen im Blut
73	LymphoBlo	Vorkommen von CD5 <sup>+</sup> CD2 <sup>-</sup> Zellen im Blut
74	LymphoBlo	Vorkommen von CD8a <sup>+</sup> $\gamma\delta^-$ Zellen im Blut
75	LymphoBlo	Vorkommen von CD8a <sup>+</sup> $\gamma\delta^+$ Zellen im Blut
76	LymphoBlo	Vorkommen von CD8a <sup>+</sup> CD4 <sup>-</sup> Zellen im Blut
77	LymphoBlo	Vorkommen von CD8b <sup>+</sup> Zellen im Blut
78	LymphoBlo	Vorkommen von MHCII <sup>+</sup> CD21 <sup>-</sup> Zellen im Blut
79	LympholleIEL	Vorkommen von CD1 <sup>+</sup> CD21 <sup>-</sup> IEL (Ileum)
80	LympholleIEL	Vorkommen von CD1 <sup>+</sup> CD21 <sup>+</sup> IEL (Ileum)
81	LympholleIEL	Vorkommen von CD2 <sup>+</sup> CD5 <sup>-</sup> IEL (Ileum)
82	LympholleIEL	Vorkommen von CD2 <sup>+</sup> CD5 <sup>+</sup> IEL (Ileum)
83	LympholleIEL	Vorkommen von CD21 <sup>+</sup> CD1 <sup>-</sup> IEL (Ileum)
84	LympholleIEL	Vorkommen von CD5 <sup>-</sup> $\gamma\delta^+$ IEL (Ileum)
85	LympholleIEL	Vorkommen von CD5 <sup>+</sup> $\gamma\delta^+$ IEL (Ileum)
86	LympholleIEL	Vorkommen von CD5 <sup>+</sup> CD2 <sup>-</sup> IEL (Ileum)
87	LymphoJeJIEL	Vorkommen von $\gamma\delta^-$ CD8 <sup>-</sup> IEL (Jejunum)
88	LymphoJeJIEL	Vorkommen von $\gamma\delta^+$ CD5 <sup>-</sup> IEL (Jejunum)
89	LymphoJeJIEL	Vorkommen von CD16 <sup>+</sup> IEL (Jejunum)
90	LymphoJeJIEL	Vorkommen von CD2 <sup>+</sup> CD5 <sup>-</sup> IEL (Jejunum)
91	LymphoJeJIEL	Vorkommen von CD2 <sup>+</sup> CD5 <sup>+</sup> IEL (Jejunum)
92	LymphoJeJIEL	Vorkommen von CD3 <sup>+</sup> IEL (Jejunum)
93	LymphoJeJIEL	Vorkommen von CD4 <sup>+</sup> CD8b <sup>-</sup> IEL (Jejunum)
94	LymphoJeJIEL	Vorkommen von CD4 <sup>+</sup> CD8b <sup>+</sup> IEL (Jejunum)
95	LymphoJeJIEL	Vorkommen von CD5 <sup>+</sup> $\gamma\delta^-$ IEL (Jejunum)
96	LymphoJeJIEL	Vorkommen von CD5 <sup>+</sup> $\gamma\delta^-$ IEL (Jejunum)
97	LymphoJeJIEL	Vorkommen von CD5 <sup>+</sup> $\gamma\delta^+$ IEL (Jejunum)
98	LymphoJeJIEL	Vorkommen von CD5 <sup>+</sup> CD2 <sup>-</sup> IEL (Jejunum)
99	LymphoJeJIEL	Vorkommen von CD8 <sup>+</sup> $\gamma\delta^+$ IEL (Jejunum)
100	LymphoJeJIEL	Vorkommen von CD8a <sup>+</sup> $\gamma\delta^-$ IEL (Jejunum)
101	LymphoJeJIEL	Vorkommen von CD8b <sup>+</sup> CD4 <sup>-</sup> IEL (Jejunum)

**Tabelle E.12.: Großes Blutbild**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	BloodCellCounts	Basophile Granulozyten
2	BloodCellCounts	Eosinophile Granulozyten

*Fortsetzung auf der nächsten Seite*

**Tabelle E.12 – Fortsetzung**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
3	BloodCellCounts	Erythrozyten
4	BloodCellCounts	Hämatokrit
5	BloodCellCounts	Hämoglobin
6	BloodCellCounts	Leukozyten
7	BloodCellCounts	Lymphozyten
8	BloodCellCounts	Metamyelozyten
9	BloodCellCounts	Mittlere corpuskuläre Hämoglobinkonzentration
10	BloodCellCounts	Mittleres corpuskuläres Hämoglobin
11	BloodCellCounts	Mittleres corpuskuläres Volumen
12	BloodCellCounts	Monozyten
13	BloodCellCounts	Segmentkernige Granulozyten
14	BloodCellCounts	Stabkernige Granulozyten
15	BloodCellCounts	Thrombozyten

**Tabelle E.13.: Gewicht der Tiere und Größe einzelner Organe**

<b>Nr.</b>	<b>Merkmalsgruppe</b>	<b>Beschreibung</b>
1	BodySize	Gewicht der Bauchspeicheldrüse
2	BodySize	Gewicht der Leber
3	BodySize	Gewicht der Milz
4	BodySize	Gewicht des Thymus
5	BodySize	Körpergewicht
6	BodySize	Länge des Dünndarms



# Wissenschaftliche Veröffentlichungen

**Chai W., Wang Z., Janczyk P., Twardziok S., Blohm U., Osterrieder N., Burwinkel M.** Elevated dietary zinc oxide levels do not have a substantial effect on porcine reproductive and respiratory syndrome virus (PPRSV) vaccination and infection. *Virology Journal*, 2014, 11(1), 140

**Gefeller E. M., Martens H., Aschenbach J. R., Klingspor S., Twardziok S., Wrede P., Pieper R., Lodemann U.** Effects of age and zinc supplementation on transport properties in the jejunum of piglets. *Journal of Animal Physiology and Animal Nutrition*, 2014, DOI:10.1111/jpn.12232

**Chai W., Zakrzewski S. S., Günzel D., Pieper R., Wang Z., Twardziok S., Janczyk P., Osterrieder N., Burwinkel M.** High-dose dietary zinc oxide mitigates infection with transmissible gastroenteritis virus in piglets. *BMC Veterinary Research*, 2014, 10(1), 75

**Twardziok S. O., Pieper R., Aschenbach J.R., Bednorz C., Brockmann G. A., Fromm M., Klingspor S., Kreuzer S., Lodemann U., Martens H., Martin L., Richter J.F., Scharek-Tedin L., Siepert B.F., Starke I.C., Tedin K., Vahjen W., Wieler L.H., Zakrzewski S.S., Zentek J., Wrede P.** Cross-talk between Host, Microbiome and Probiotics: A Systems Biology Approach for Analyzing the Effects of Probiotic *Enterococcus faecium* NCIMB 10415 in Piglets. *Molecular Informatics*, 2014, 33(3), 171-182

**Sachsenröder J., Twardziok S. O., Scheuch M., Johne R.** The general composition of the faecal virome of pigs depends on age, but not on feeding with a probiotic bacterium. *PLOS ONE*, 2014, 9(2), e88888

**Wang Z., Burwinkel M., Chai W., Lange E., Blohm U., Breithaupt A., Hoffmann B., Twardziok S., Rieger J., Janczyk P., Pieper R., Osterrieder N.** Dietary *Enterococcus faecium* NCIMB 10415 and Zinc Oxide Stimulate Immune Reactions to Trivalent Influenza Vaccination in Pigs but Do Not Affect Virological Response upon Challenge Infection. *PLOS ONE*, 2014, 9(1), e87007

**Siepert B., Reinhardt N., Kreuzer S., Bondzio A., Twardziok S., Brockmann G., Nöckler K., Szabó I., Janczyk P., Pieper R., Tedin K.** *Enterococcus faecium* NCIMB 10415 supplementation affects intestinal immune-associated gene expression in post-weaning piglets. *Veterinary Immunology and Immunopathology*, 2014, 157, 65-77

**Rieger J., Twardziok S., Huenigen H., Hirschberg R. M., Plendl J.** Porcine intestinal mast cells. Evaluation of different fixatives for histochemical staining techniques considering tissue shrinkage. *European Journal of Histochemistry*, 2013, 57(3), e21

**Klingspor, S., Martens, H., Çaushi, D., Twardziok, S., Aschenbach, J.R., Lodemann, U.** Characterization of the effects of *Enterococcus faecium* on intestinal epithelial transport properties in piglets. *Journal of Animal Science*, 2013, 91(4), 1707-1718

**Chai W., Burwinkel M., Wang Z., Palissa C., Esch B., Twardziok S., Rieger J., Wrede P., Schmidt M. F. G.** Antiviral effects of a probiotic *Enterococcus faecium* strain against transmissible gastroenteritis coronavirus. *Archives of Virology*, 2013, 158(4), 799-807

**Wang Z., Chai W., Burwinkel M., Twardziok S., Wrede P., Palissa C., Esch B., Schmidt M. F. G.** Inhibitory Influence of *Enterococcus faecium* on the Propagation of Swine Influenza A Virus *In Vitro*. *PLOS ONE*, 2013, 8(1), e53043

**Trojnar E., Sachsenröder J., Twardziok S., Reetz J., Otto P. H., Johne R.** Identification of an avian group A rotavirus containing a novel VP4 gene with a close relationship to those of mammalian rotaviruses. *Journal of General Virology*, 2013, 94(Pt 1), 136-142

**Deiner C., Reiche M., Lassner D., Grienitz D., Twardziok S., Moesch A., Wenning P., Martens H.** Allelic variations in coding regions of the vitamin D receptor gene in dairy cows and potential susceptibility to periparturient hypocalcaemia. *Journal of Dairy Research*, 2012 (Nov), 79(4), 423-428

**Sachsenröder J., Twardziok S., Hammerl J. A., Janczyk P., Wrede P., Hertwig S., Johne R.** Simultaneous Identification of DNA and RNA Viruses Present in Pig Faeces Using Process-Controlled Deep Sequencing. *PLOS ONE*, 2012, 7(4), e34631

**Rai M. F., Graeve T., Twardziok S., Schmidt M. F. G.** Evidence for Regulated Interleukin-4 Expression in Chondrocyte-Scaffolds under *In Vitro* Inflammatory Conditions. *PLOS ONE*, 2011, 6(10), e25749

**Twardziok S., Siebert H., Heyl A.** Stochasticity in reactions: a probabilistic boolean modeling approach. *Proceedings of the 8th International Conference on Computational Methods in Systems Biology*, 2010, 76-85

# Danksagung

Hiermit möchte ich verschiedenen Menschen danken, die mich während meiner Promotion auf unterschiedliche Arten unterstützt haben. Mein Dank gilt:

- **Herrn Professor M. F. G. Schmidt** für die Betreuung meiner Promotion am Fachbereich Veterinärmedizin der Freien Universität Berlin und für viele nützliche Hinweise zur Anfertigung dieser Arbeit.
- **Herrn Professor P. Wrede** für die intensive wissenschaftliche Betreuung meiner Dissertation und für die Möglichkeit diese interessante Thema in der Abreitsgruppe Bioinformatik und Molekularbiologie der Charité-Universitätsmedizin Berlin zu bearbeiten.
- **Herrn PD J. Kleffe** für die wissenschaftliche Betreuung meiner Dissertation und seine Unterstützung und Beratung bei der Durchführung dieser Arbeit.
- allen **Mitgliedern des Sonderforschungsbereichs 852** und hierbei insbesondere den Gruppenleitern **Prof. Dr. T. Alter, Prof. Dr. B. Appel, Prof. Dr. M. Blaut, Prof. Dr. G. Brockmann, Prof. Dr. R. Einspanier, Prof. Dr. M. Fromm, PD Dr. S. Hartmann, PD Dr. R. Johne, Dr. G. Loh, Prof. Dr. H. Martens, Dr. K. Nöckler, Prof. Dr. K. Osterrieder, Prof. Dr. J. Plendl, Prof. Dr. M. F. G. Schmidt, Dr. K. Tedin, Dr. W. Vahjen** und **Prof. Dr. J. Zentek** für die Durchführung der Tierversuche, für spannende Diskussionen und für die Bereitstellung der Forschungsdaten für die Piggeldat-Datenbank.

Mein besonderer Dank gilt sowohl meinem Vater für die Unterstützung meines Studiums als auch meiner Frau Monika, die während meiner Arbeit viel Verständnis und Geduld für mich aufbrachte.





# Selbstständigkeitserklärung

Hiermit bestätige ich, dass ich die vorliegende Arbeit selbständig angefertigt habe. Ich versichere, dass ich ausschließlich die angegebenen Quellen und Hilfen in Anspruch genommen habe.

Berlin, den 17. November 2014

Sven O. Twardziok