



**Metatranskriptom-Analysen der mikrobiellen  
Lebensgemeinschaften des menschlichen Ohres und der  
Achselhöhle**

**Dissertation**

zur Erlangung des akademischen Grades  
Doktor der Naturwissenschaften (Dr. rer. nat.)  
der Fakultät für Biologie der Universität Bielefeld

vorgelegt von

**Eugenie Fredrich-Vahle**

Oktober 2016

*Für meinen lieben Opa.*

# INHALTSVERZEICHNIS

|  |            |
|--|------------|
| <b>INHALTSVERZEICHNIS.....</b>   | <b>I</b>   |
| <b>ABBILDUNGSVERZEICHNIS.....</b>  | <b>VII</b> |
| <b>TABELLENVERZEICHNIS.....</b>  | <b>X</b>   |
| <b>ABKÜRZUNGSVERZEICHNIS .....</b>   | <b>XII</b> |
| <br>   |            |
| <b>I. Zusammenfassung .....</b>  | <b>1</b>   |
| <br>   |            |
| <b>II. Einleitung .....</b>  | <b>4</b>   |
| <br>   |            |
| <b>1 Die Mikrobiota des Menschen .....</b>   | <b>4</b>   |
| 1.1 Beschreibung der humanen Mikrobiota mittels Hochdurchsatzsequenzierung.....                  | 5          |
| 1.1.1 16S-rDNA-Amplikon-Sequenzierung.....   | 6          |
| 1.1.2 Metagenomik.....   | 8          |
| 1.1.3 Metatranskriptomik.....  | 9          |
| <br>   |            |
| <b>2 Die mikrobielle Lebensgemeinschaft des menschlichen Ohres und Gehörgangsentzündung.....</b> | <b>10</b>  |
| 2.1 Gehörgangsentzündung – Krankheitsbild, Entstehung und Behandlung.....                        | 10         |
| 2.2 Die Mikrobiologie der Gehörgangsentzündung.....  | 12         |
| <br>   |            |
| <b>3 Die Mikrobiota der menschlichen Achselhöhle und Körpergeruchsentstehung .....</b>           | <b>13</b>  |
| 3.1 Kultivierungsbasierte Beschreibungen der Achselhöhlen-Mikrobiota.....                        | 16         |



|          |  |           |
|----------|--|-----------|
| 3.2      | Molekulare Mechanismen der Körpergeruchsentstehung.....  | 17        |
| 3.2.1    | Steroide .....   | 17        |
| 3.2.2    | Kurzkettige Fettsäuren.....  | 19        |
| 3.2.3    | Sulfanylalkanole .....   | 23        |
| 3.2.4    | Beteiligung menschlicher Komponenten an der Körpergeruchsentstehung ...                                  | 26        |
| 3.3      | Kultivierungsunabhängige Beschreibungen der Achselhöhlen-Mikrobiota .....                                | 28        |
| 3.4      | Übermäßiger Körpergeruch, seine Bekämpfung und die Auswirkungen auf die Mikrobiota der Achselhöhle ..... | 33        |
| 3.4.1    | Übermäßiges Schwitzen und Körpergeruch .....   | 33        |
| 3.4.2    | Topische Behandlung von Schweißbildung und Körpergeruch durch Deodorants und Antiperspirantien.....      | 34        |
| 3.4.3    | Einfluss von Deodorants und Antiperspirantien auf die Mikrobiota der Achselhöhle .....                   | 36        |
| <b>4</b> | <b>Zielsetzung .....</b>   | <b>37</b> |

**III. Material und Methoden..... 39**

|          |   |           |
|----------|---|-----------|
| <b>1</b> | <b>Probenmaterial .....</b>                     | <b>39</b> |
| 1.1      | Probe aus dem menschlichen Ohr .....            | 39        |
| 1.2      | Proben der Achselhöhlen-Mikrobiota .....        | 39        |
| 1.2.1    | Vorbereitung der Probanden.....                 | 39        |
| 1.2.2    | Deo-Studie.....                                 | 40        |
| 1.2.3    | Bestimmung des <i>Malodor Score</i> .....       | 40        |
| 1.2.4    | Probennahme.....                                | 41        |
| 1.2.5    | Bestimmung der koloniebildenden Einheiten ..... | 41        |
| <b>2</b> | <b>Labormaterialien und Geräte .....</b>        | <b>42</b> |
| 2.1      | Chemikalien, Reagenzien und Medienzusätze ..... | 42        |
| 2.1.1    | Chemikalien.....                                | 42        |

|          |  |           |
|----------|--|-----------|
| 2.1.2    | Oligonukleotide.....   | 42        |
| 2.1.3    | Kits.....  | 43        |
| 2.1.4    | Enzyme und Marker .....  | 43        |
| 2.2      | Verbrauchsaterialien.....  | 43        |
| 2.3      | Geräte .....   | 44        |
| 2.4      | Software und Bioinformatik-Tools .....                                   | 45        |
| <b>3</b> | <b>Medien, Puffer und Lösungen.....</b>                                  | <b>45</b> |
| <b>4</b> | <b>Allgemeine molekularbiologische Arbeiten .....</b>                    | <b>46</b> |
| 4.1      | Isolierung von Gesamt-DNA aus <i>C. jeikeium</i> K411.....               | 46        |
| 4.2      | Polymerase-Kettenreaktion (PCR) .....                                    | 47        |
| 4.3      | Agarose-Gelelektrophorese .....  | 48        |
| 4.4      | Konzentrationsbestimmung und Qualitätskontrolle von DNA und RNA.....     | 48        |
| <b>5</b> | <b>Metatranskriptomik .....</b>  | <b>49</b> |
| 5.1      | RNA-Isolierung .....   | 49        |
| 5.2      | PCR zur Überprüfung der RNA auf Verunreinigung durch genomische DNA..... | 51        |
| 5.3      | DNase-Behandlung (Roche) .....   | 52        |
| 5.4      | RNA-Aufreinigung mit DNase-Verdau.....                                   | 52        |
| 5.5      | cDNA-Synthese und Amplifikation.....                                     | 53        |
| 5.6      | Säulenbasierte-Aufreinigung der cDNA .....                               | 54        |
| 5.7      | cDNA-Fragmentierung und Aufreinigung .....                               | 54        |
| 5.8      | Synthese und Aufreinigung von cDNA-Sequenzierbibliotheken .....          | 55        |
| 5.9      | Sequenzierung.....   | 55        |
| <b>6</b> | <b>Bioinformatik .....</b>   | <b>56</b> |
| 6.1      | Bearbeiten der Sequenzierdaten mit SEQPREP .....                         | 56        |
| 6.2      | Qualitätsanalyse der Sequenzier-Reads mit dem FASTX-Toolkit.....         | 56        |
| 6.3      | Metatranskriptom-Analysen mit METRANS.....                               | 57        |

6.4 Metatranskriptom-Analysen mit MG-Rast .....59

    6.4.1 Taxonomische Analyse von Metatranskriptomik-Datensätzen in MG-RAST.....62

    6.4.2 Funktionale Analyse von Metatranskriptomik-Datensätzen in MG-RAST .....64

6.5 Taxonomische Analyse der Metatranskriptomik-Datensätze auf Spezies-Ebene mit der *GenBank*-Datenbank.....69

6.6 Durchführung von Hauptkomponentenanalysen (PCAs) .....69

6.7 Analyse von Funktionen mit Korrelation zum Körpergeruch .....70

**IV. Ergebnisse ..... 73**

**1 Metatranskriptomik des menschlichen Ohrs im Zusammenhang mit Außenohrentzündung..... 74**

1.1 Laborverfahren der Metatranskriptomik des menschlichen Ohrs .....74

1.2 Metatranskriptomik-Analyse der Ohr-Probe mithilfe von MG-RAST .....76

    1.2.1 Bioinformatik-basierte Beschreibung der Datenanalyse mittels MG-RAST.....76

    1.2.2 Taxonomische Analyse mit MG-RAST .....78

    1.2.3 Funktionale Analyse des Metatranskriptomik-Datensatzes des menschlichen Ohrs mit MG-RAST .....84

1.3 Metatranskriptomik-Analyse des menschlichen Ohrs mithilfe von METRANS.....91

    1.3.1 Bioinformatik-basierte Beschreibung der Datenanalyse mittels METRANS .....91

    1.3.2 16S-rRNA-basierte Taxonomie mittels METRANS .....93

    1.3.3 Funktionale Taxonomie des Metatranskriptomik-Datensatzes aus dem Ohr mittels METRANS.....95

    1.3.4 Funktionale Beschreibung des Metatranskriptoms des menschlichen Ohrs in METRANS 97

**2 Metatranskriptomik der menschlichen Achselhöhlen-Mikrobiota .... 100**

2.1 Etablierung einer Labormethode für die Metatranskriptomik-Analyse von Probenmaterial aus der menschlichen Achselhöhle .....100

|       |  |     |
|-------|--|-----|
| 2.1.1 | Ergebnisse der Deo-Studie.....   | 100 |
| 2.1.2 | Laborprotokoll für die Metatranskriptomik-Analyse der Achselhöhlen-Mikrobiota .....  | 102 |
| 2.2   | Bioinformatik-Analyse der Sequenzdaten mittels MG-RAST .....   | 106 |
| 2.3   | Taxonomische Analyse des Metatranskriptoms der Achselhöhlen-Mikrobiota .....   | 112 |
| 2.3.1 | Taxonomische Beschreibung der Achselhöhlen-Mikrobiota auf Genus-Ebene<br>112   |     |
| 2.3.2 | Hauptkomponentenanalyse der rRNA-basierten taxonomischen<br>Zusammensetzung der Achselhöhlen-Mikrobiota auf Genus-Ebene..... | 116 |
| 2.3.3 | Taxonomische Beschreibung der Achselhöhlen-Mikrobiota auf Spezies-Ebene<br>119   |     |
| 2.3.4 | 16S-rRNA-basierte Darstellung der Achselhöhlen-Mikrobiota auf Grundlage<br>der <i>GenBank</i> -Datenbank .....               | 127 |
| 2.4   | Funktionale Analyse des Metatranskriptoms der Achselhöhlen-Mikrobiota.....   | 129 |
| 2.4.1 | Allgemeine funktionale Analyse der Achselhöhlen-Mikrobiota .....   | 129 |
| 2.4.2 | Analyse von Transkripten aus <i>Corynebacterium striatum</i> im Zusammenhang<br>mit Körpergeruchsentstehung.....             | 133 |
| 2.4.3 | Analyse von Aminoacylase-Transkripten .....  | 135 |
| 2.4.4 | Analyse von Cystathionin- $\beta$ -Lyase-Transkripten .....  | 140 |
| 2.4.5 | Analyse von Dipeptidase-Transkripten .....   | 148 |
| 2.4.6 | Analyse anderer im Zusammenhang mit Körpergeruchsentstehung<br>beschriebener Funktionen.....                                 | 151 |
| 2.4.7 | Beschreibung von Funktionen mit Korrelation zum Körpergeruch .....   | 155 |

**V. Diskussion ..... 167**

**1 Methode zur Beschreibung der menschlichen Mikrobiota ..... 167**

|     |  |     |
|-----|--|-----|
| 1.1 | Etablierung einer Labormethode zur Beschreibung der mikrobiellen<br>Lebensgemeinschaften des Ohrs und der Achselhöhle..... | 168 |
|-----|--|-----|

|              |  |            |
|--------------|--|------------|
| 1.2          | Bioinformatik-Analyse von Metatranskriptomik-Datensätzen der menschlichen Mikrobiota .....                           | 172        |
| <b>2</b>     | <b>Metatranskriptomik der Mikrobiota des menschlichen Ohrs im Zusammenhang mit Gehörgangsentzündung.....</b>         | <b>177</b> |
| 2.1          | Taxonomische Betrachtung der aktiven Lebensgemeinschaft.....   | 177        |
| 2.2          | Funktionale Betrachtung der Mikrobiota des menschlichen Ohrs.....  | 179        |
| <b>3</b>     | <b>Metatranskriptomik der menschlichen Achselhöhlen-Mikrobiota im Zusammenhang mit Körpergeruchsentstehung .....</b> | <b>181</b> |
| 3.1          | Taxonomische Betrachtung der Achselhöhlen-Mikrobiota.....  | 181        |
| 3.2          | Funktionale Betrachtung der Achselhöhlen-Mikrobiota .....  | 185        |
| <b>VI.</b>   | <b>Ausblick .....</b>  | <b>195</b> |
| <b>VII.</b>  | <b>Literaturverzeichnis .....</b>  | <b>197</b> |
| <b>VIII.</b> | <b>Anhang .....</b>  | <b>214</b> |
| 1            | Referenzsequenzen zu COG1168 (MalY) aus Staphylokokken .....   | 214        |
| 2            | Referenzsequenzen zur TpdA-Dipeptidase .....   | 215        |
| 3            | Funktionen mit Korrelation zum Körpergeruch .....  | 216        |

## ABBILDUNGSVERZEICHNIS

|  |     |
|--|-----|
| <b>Abbildung 1:</b> Übersicht bekannter molekularer Mechanismen der Körpergeruchsentstehung durch die Aktivität der Achselhöhlen-Mikrobiota.....   | 18  |
| <b>Abbildung 2:</b> Schematische Darstellung der Labormethode zur Metatranskriptomik-Analyse der Ohrprobe.....   | 75  |
| <b>Abbildung 3:</b> <i>Quality-Plot</i> des Metatranskriptomik-Datensatzes aus dem menschlichen Ohr.....   | 76  |
| <b>Abbildung 4:</b> Übersicht der Sequenzdatenanalyse des Metatranskriptomik-Datensatzes aus dem Ohr in MG-RAST.....   | 77  |
| <b>Abbildung 5:</b> rRNA-basierte Taxonomie des Metatranskriptomik-Datensatzes aus dem Ohr. ....   | 79  |
| <b>Abbildung 6:</b> Funktionale Taxonomie des Metatranskriptomik-Datensatzes aus dem Ohr. ....   | 81  |
| <b>Abbildung 7:</b> Taxonomische Darstellung des Metatranskriptomik-Datensatzes aus dem Ohr auf Spezies-Ebene.....   | 83  |
| <b>Abbildung 8:</b> Funktionale Klassifizierung des Metatranskriptomik-Datensatzes aus der Ohr-Probe auf Grundlage der COG-Klasseneinteilung. ....   | 85  |
| <b>Abbildung 9:</b> Übersicht der Sequenzdatenanalyse des Metatranskriptomik-Datensatzes aus dem Ohr mittels METTRANS. ....  | 92  |
| <b>Abbildung 10:</b> 16S-rRNA-basierte taxonomische Zuordnung des Metatranskriptomik-Datensatzes der Ohr-Probe mittels METTRANS unter Verwendung des RDP-CLASSIFIER bei einem Konfidenz-Schwellenwert von 0,8..... | 94  |
| <b>Abbildung 11:</b> Funktionale Taxonomie der normalisierten Sequenzabundanzen mit Homologie zu proteinkodierenden Genen in METTRANS.....   | 96  |
| <b>Abbildung 12:</b> Fließdiagramm der Metatranskriptom-Analyse von Proben der Achselhöhlen-Mikrobiota.....  | 103 |
| <b>Abbildung 13:</b> <i>Quality-Plot</i> des Metatranskriptomik-Datensatzes von Probe 48n.....   | 108 |
| <b>Abbildung 14:</b> Übersicht der Sequenzdatenanalyse der Metatranskriptomik-Datensätze aus der Achselhöhle mittels MG-RAST.....  | 109 |
| <b>Abbildung 15:</b> Taxonomische Darstellung der aktiven Achselhöhlen-Mikrobiota auf Genus-Ebene.....   | 114 |
| <b>Abbildung 16:</b> Hauptkomponentenanalyse (PCA) der rRNA-basierten taxonomischen Zusammensetzung der Achselhöhlen-Mikrobiota auf Genus-Ebene. ....  | 117 |
| <b>Abbildung 17:</b> Diversität, Artenzahl und Äquität der aktiven Lebensgemeinschaften der Achselhöhle.....   | 121 |
| <b>Abbildung 18:</b> Zusammensetzung der Achselhöhlen-Mikrobiota auf Spezies-Ebene basierend auf der <i>M5RNA</i> -Datenbank.....  | 123 |

|  |     |
|--|-----|
| <b>Abbildung 19:</b> Zusammensetzung der Achselhöhlen-Mikrobiota des Probanden 18 auf Spezies-Ebene basierend auf der <i>M5RNA</i> -Datenbank.....   | 124 |
| <b>Abbildung 20:</b> Zusammensetzung der Achselhöhlen-Mikrobiota des Probanden 19 auf Spezies-Ebene basierend auf der <i>M5RNA</i> -Datenbank.....   | 125 |
| <b>Abbildung 21:</b> 16S-rRNA-basierte taxonomische Darstellung der Metatranskriptomik-Datensätze der Achselhöhle auf Genus-Ebene unter Verwendung der <i>GenBank</i> -Datenbank.....  | 128 |
| <b>Abbildung 22:</b> Funktionale Profile der Metatranskriptomik-Datensätze der Achselhöhle nach der COG-Klasseneinteilung..  | 130 |
| <b>Abbildung 23:</b> Hauptkomponentenanalyse (PCA) der funktionalen Zuordnung der Metatranskriptomik-Datensätze auf Grundlage der COG-Klasseneinteilung.....   | 132 |
| <b>Abbildung 24:</b> Funktionales taxonomisches Profil der Gruppe COG1473 (Metall-abhängige Amidase/Aminoacylase/Carboxypeptidase) auf Genus-Ebene. ....   | 136 |
| <b>Abbildung 25:</b> Funktionale taxonomische Profile der GenBank-Annotationen N <sup>α</sup> -Acylglutamin-Aminoacylase, Aminoacylase und mögliche Aminoacylase auf Spezies-Ebene. ....   | 137 |
| <b>Abbildung 26:</b> Ausschnitt des <i>Multiple Sequence Alignment</i> der Sequenzen mit der <i>GenBank</i> -Annotation zur N <sup>α</sup> -Acylglutamin-Aminoacylase in MG-RAST sowie der Referenzsequenzen aus der händischen Abfrage in der <i>GenBank</i> -Datenbank.. | 138 |
| <b>Abbildung 27:</b> Kladogramm der in MG-RAST identifizierten Sequenzen mit Homologie zur N <sup>α</sup> -Acylglutamin-Aminoacylase und der Referenzsequenzen aus der händischen Abfrage in der <i>GenBank</i> -Datenbank. ....   | 139 |
| <b>Abbildung 28:</b> Funktionales taxonomisches Profil der Annotationen zur Cystathionin-β-Lyase auf Spezies-Ebene.....  | 141 |
| <b>Abbildung 29:</b> Taxonomisches Profil der Sequenzen mit Homologie zur Gruppe der MalY-Proteine (COG1168). ....   | 144 |
| <b>Abbildung 30:</b> Funktionales taxonomisches Profil der Annotationen zur <i>Thiol Precursor Dipeptidase</i> (TpdA) auf Spezies-Ebene. ....  | 148 |
| <b>Abbildung 31:</b> Funktionales taxonomisches Profil der Referenzsequenzen mit Annotation zur TpdA Dipeptidase in der Online-Version der <i>GenBank</i> -Datenbank.....  | 150 |
| <b>Abbildung 32:</b> Funktionales taxonomisches Profil der Annotationen mit Beteiligung am Fettsäureabbau auf Genus-Ebene mithilfe der <i>Best Hit Classification</i> in MG-RAST.....  | 152 |
| <b>Abbildung 33:</b> Funktionales taxonomisches Profil der Annotationen zu BCAA-Aminotransferasen (EC 2.6.1.42) auf Genus-Ebene mithilfe der <i>Best-Hit-Classification</i> in MG-RAST. ....   | 154 |
| <b>Abbildung 34:</b> Funktionales taxonomisches Profil der Annotationen zu fermentativen Stoffwechselwegen auf Genus-Ebene mithilfe der <i>Best Hit Classification</i> in MG-RAST.....   | 155 |

|   |     |
|---|-----|
| <b>Abbildung 35:</b> Darstellung positiv, negativ und nicht mit dem Körpergeruch korrelierter Enzymfunktionen im Stoffwechselweg des Leucin-, Isoleucin- und Valin-Abbaus auf Grundlage der KO-Ontologie mithilfe des KEGG-MAPPER SEARCH&COLOR PATHWAY-Tools. ....                              | 159 |
| <b>Abbildung 36:</b> Taxonomische Profile relevanter Funktionen aus dem Stoffwechselweg des Leucin-, Isoleucin- und Valin-Abbaus auf Grundlage der KO-Ontologie. ....   | 160 |
| <b>Abbildung 37:</b> Darstellung positiv, negativ und nicht korrelierter Enzymfunktionen im Glutathion-Stoffwechsel auf Grundlage der KO-Ontologie mithilfe des KEGG-MAPPER SEARCH&COLOR PATHWAY-Tools. ....  | 163 |
| <b>Abbildung 38:</b> Taxonomische Profile relevanter Funktionen aus dem Glutathion-Metabolismus auf Grundlage der KO-Ontologie. ....  | 164 |
| <b>Abbildung 39:</b> Darstellung der positiv (grün) korrelierten Enzymfunktionen Phenylacetaldehyd-Dehydrogenase (EC 1.2.1.39) und Phenylacetat-Ligase (EC 6.2.1.30) aus dem Phenylalanin-Stoffwechsel auf Grundlage der KO-Ontologie mithilfe des KEGG-MAPPER SEARCH&COLOR PATHWAY-Tools. .... | 165 |
| <b>Abbildung 40:</b> Taxonomische Profile der Annotationen zur Phenylacetaldehyd Dehydrogenase (EC 1.2.1.39) und Phenylacetat-Ligase (EC 6.2.1.30) aus dem KO-Annotationssystem auf Genus-Ebene. ....   | 166 |
| <b>Abbildung 41:</b> Postulierte Methode zur Metatranskriptom-Analyse kleiner menschlicher Proben und die Umsetzung für die Probe des Außenohres und die Proben aus der Achselhöhle in der vorliegenden Arbeit. ....  | 170 |
| <b>Abbildung 42:</b> Schema der Pathogen-Wirt-Beziehungen der Ohrprobe im Zusammenhang mit Gehörgangsentzündung. ....   | 179 |
| <b>Abbildung 43:</b> <i>Multiple Sequence Alignment</i> der mittels BLASTX identifizierten Referenzsequenzen aus der Genbank-Datenbank zu cDNA-Sequenzen mit Homologie zu COG1168 (MaY). ....   | 214 |
| <b>Abbildung 44:</b> <i>Multiple Sequence Alignment</i> der mittels BLASTX identifizierten Referenzsequenzen aus der Genbank-Datenbank zu cDNA-Sequenzen mit Homologie zur TpdA-Dipeptidase. ....   | 215 |
| <b>Abbildung 45:</b> Funktionale Zuordnungen innerhalb der Metatranskriptom-Datensätze der Achselhöhlen-Proben mit Korrelation zum Körpergeruch auf Grundlage des KO-Annotationssystems. ....   | 216 |



# TABELLENVERZEICHNIS

|   |     |
|---|-----|
| <b>Tabelle 1:</b> Übersicht der kultivierungsunabhängigen Studien zur Zusammensetzung der Achselhöhlen-Mikrobiota.....  | 30  |
| <b>Tabelle 2:</b> Liste der verwendeten Chemikalien und ihrer Bezugsquellen. ....   | 42  |
| <b>Tabelle 3:</b> Liste der verwendeten Kits und ihrer Bezugsquellen .....  | 43  |
| <b>Tabelle 4:</b> Liste der verwendeten Enzyme und Marker und ihrer Bezugsquellen.....  | 43  |
| <b>Tabelle 5:</b> Liste der verwendeten Verbrauchsmaterialien und ihrer Bezugsquellen .....   | 43  |
| <b>Tabelle 6:</b> Liste der verwendeten Geräte und Apparaturen und ihrer Bezugsquelle .....   | 44  |
| <b>Tabelle 7:</b> Liste der verwendeten Software und Bioinformatik-Tools und ihrer Bezugsquellen/Referenzen .....   | 45  |
| <b>Tabelle 8:</b> Schwellenwerte der Analyseparameter für die Erstellung von COG-Profilen in MG-RAST.....   | 65  |
| <b>Tabelle 9:</b> Beschreibung der Virulenzfaktoren von <i>Staphylococcus aureus</i> aus dem Metatranskriptomik-Datensatz des menschlichen Ohres in MG-RAST auf Grundlage des SwissProt-Annotationssystems. ....            | 87  |
| <b>Tabelle 10:</b> : Beschreibung der Virulenzfaktoren von <i>Staphylococcus aureus</i> aus dem Metatranskriptomik-Datensatz des menschlichen Ohres in METRANS auf Grundlage der Treffer gegen die SwissProt-Datenbank..... | 98  |
| <b>Tabelle 11:</b> Übersicht der im Rahmen der Deo-Studie beobachteten Ergebnisse. ....   | 101 |
| <b>Tabelle 12:</b> Übersicht des für die Metatranskriptomik verwendeten Probenmaterials der Achselhöhle. ....   | 104 |
| <b>Tabelle 13:</b> Übersicht der Daten zum Laborprotokoll der Metatranskriptomik der Achselhöhlen-Mikrobiota.....   | 106 |
| <b>Tabelle 14:</b> Übersicht der Sequenzdatenbearbeitung mit SEQPREP. ....  | 107 |
| <b>Tabelle 15:</b> Übersicht der Sequenz-Datensätze der Achselhöhlen-Mikrobiota nach Bearbeitung in MG-RAST.....  | 111 |
| <b>Tabelle 16:</b> Korrelation einzelner Bakterienspezies mit dem Körpergeruch.....   | 126 |
| <b>Tabelle 17:</b> Analyse der für <i>Corynebacterium striatum</i> im Zusammenhang mit Geruchsentstehung beschriebenen Funktionen in den Metatranskriptomik-Datensätzen der Achselhöhle.....                                | 134 |
| <b>Tabelle 18:</b> Identifizierung von Sequenzen mit Homologie zu Sequenzen der Cystathionin- $\beta$ -Lyase in der GenBank-Datenbank. ....   | 142 |
| <b>Tabelle 19:</b> Referenzsequenzen der MalY-Sequenzen in der GenBank-Datenbank.....   | 145 |

|   |     |
|---|-----|
| <b>Tabelle 20:</b> Referenzsequenzen der in MG-RAST identifizierten Sequenzen mit Homologie zur TpdA Dipeptidase. ....  | 149 |
| <b>Tabelle 21:</b> Funktionen mit den stärksten positiven bzw. negativen Korrelationen zum Körpergeruch und die dominanten Genera innerhalb der jeweiligen funktionalen taxonomischen Profile ..... | 156 |

## ABKÜRZUNGSVERZEICHNIS

|                                 |  |
|---------------------------------|--|
| % (v/v)                         | Volumenprozent   |
| % (w/v)                         | Gewichtsprozent  |
| °C                              | Grad Celsius   |
| µg                              | Mikrogramm   |
| µl                              | Mikroliter   |
| A                               | Adenin   |
| A.                              | <i>Anaerococcus</i> , [gefolgt von Artnamen]                         |
| ADRs                            | <i>Artificial Duplicate Reads</i> (artifizielle Read-Duplikate)      |
| bp                              | Basenpaar(e)   |
| bzw.                            | beziehungsweise  |
| C                               | Cytosin  |
| C.                              | <i>Corynebacterium</i> , [gefolgt von Artnamen]                      |
| ca.                             | circa  |
| <i>CeBiTec</i>                  | <i>Centrum für Biotechnologie der Universität Bielefeld</i>          |
| dATP                            | Desoxyadenosintriphosphat  |
| dCTP                            | Desoxycytidintriphosphat   |
| Deo                             | Deodorant  |
| dGTP                            | Desoxyguanosintriphosphat  |
| DNA                             | Desoxyribonukleinsäure   |
| dNTP                            | Desoxyribonukleosidtriphosphat                                       |
| dTTP                            | Desoxythymidintriphosphat  |
| <i>E.coli</i>                   | <i>Escherichia coli</i>  |
| EC-Nummer                       | <i>Enzyme-Commission-Nummer</i>                                      |
| EDTA                            | Ethylendiamintetraessigsäure   |
| <i>et al.</i>                   | <i>et alii</i> (und andere)  |
| F.                              | <i>Fingoldia</i> , [gefolgt von Artnamen]                            |
| g                               | Gramm  |
| G                               | Guanin   |
| GC-Gehalt                       | Guanin-Cytosin-Gehalt  |
| GC-MS                           | Gaschromatographie-Massenspektrometrie                               |
| ggf.                            | gegebenenfalls   |
| GPAC                            | <i>gram positive anaerobic cocci</i> (Gram-positive anaerobe Kokken) |
| h                               | Stunde(n)  |
| H <sub>2</sub> O                | Wasser   |
| i. d. R.                        | in der Regel   |
| K <sub>2</sub> HPO <sub>4</sub> | Dikaliumhydrogenphosphat   |

|   |   |
|---|---|
| KBE   | koloniebildende Einheiten   |
| kDa   | Kilodalton  |
| KH <sub>2</sub> PO <sub>4</sub>               | Kaliumdihydrogenphosphat  |
| l   | Liter   |
| LSU   | <i>large subunit</i> (große ribosomale Untereinheit)  |
| M   | Molar   |
| MgCl <sub>2</sub>                             | Magnesiumchlorid  |
| min   | Minute(n)   |
| ml  | Milliliter  |
| mM  | Millimolar  |
| mRNA  | <i>messenger RNA</i>  |
| Na <sub>2</sub> S <sub>2</sub> O <sub>3</sub> | Natriumthiosulfat   |
| NaCl  | Natriumchlorid  |
| N-AGA   | N <sup>α</sup> -Acylglutamin-Aminoacylase   |
| NGS   | <i>next generation sequencing</i> (bezeichnet Hochdurchsatzsequenziertechniken)             |
| nm  | Nanometer   |
| nM  | Nanomolar   |
| nt  | Nukleotide  |
| OTC   | <i>over the counter</i> (bezeichnet nicht verschreibungspflichtige Präparate)               |
| OTU   | <i>operational taxonomic unit</i> (Klasseneinteilung zur Beschreibung ähnlicher Organismen) |
| P.  | <i>Propionibacterium</i> , [gefolgt von Artnamen]   |
| PCA   | <i>Principal Component Analysis</i> (Hauptkomponentenanalyse)                               |
| PCR   | <i>polymerase chain reaction</i> (Polymerase-Kettenreaktion)                                |
| pH  | pH-Wert   |
| R.  | <i>Rickettsia</i> , [gefolgt von Artnamen]  |
| rcf   | <i>relative centrifugal force</i> , Zentrifugalbeschleunigung                               |
| rDNA  | Ribosomale DNA  |
| RIN   | <i>RNA Integrity Number</i>   |
| RNA   | Ribonukleinsäure  |
| rRNA  | ribosomale RNA  |
| RT  | Raumtemperatur  |
| s   | Sekunde(n)  |
| S.  | <i>Staphylococcus</i> , [gefolgt von Artnamen]  |
| SSU   | <i>small subunit</i> (kleine ribosomale Untereinheit)                                       |
| T   | Thymin  |
| TAE   | Tris-Acetat-EDTA-(Puffer)   |
| Tris  | Tris(hydroxymethyl)-aminomethan   |

|       |                       |
|-------|-----------------------|
| u. a. | unter anderem         |
| ü/N   | über Nacht            |
| V     | Volt                  |
| vs.   | <i>versus</i> , gegen |
| z. B. | Zum Beispiel          |
| z. T. | zum Teil              |

# I. Zusammenfassung

Der menschliche Körper beherbergt diverse mikrobielle Lebensgemeinschaften, welche eine bedeutende Rolle in der Physiologie oder Pathogenität spielen. Mittels kultivierungsbasierter Techniken konnte in der Vergangenheit jedoch lediglich ein Bruchteil dieser Lebensgemeinschaften und ihrer Funktionen beschrieben werden. Moderne Techniken der Hochdurchsatzsequenzierung ermöglichen heute die kultivierungsunabhängige Beschreibung der humanen Mikrobiota. Die Metatranskriptomik ermöglicht die Darstellung der Zusammensetzung und Funktion der aktiven Lebensgemeinschaften und damit die Beschreibung relevanter molekularer Mechanismen. Aufgrund der in der Regel sehr geringen Probenmengen sind nur wenige Lebensgemeinschaften der menschlichen Mikrobiota auf der Ebene des Metatranskriptoms beschrieben.

In der vorliegenden Arbeit wurden eine Probe aus dem menschlichen Außenohr im Zusammenhang mit Außenohrentzündung (Otitis externa) und Proben aus der Achselhöhle im Zusammenhang mit Körpergeruchsentstehung analysiert. Dazu wurde ein Laborprotokoll zur Herstellung von Sequenzierbibliotheken aus sehr geringen Probevolumina etabliert und anschließend wurden die Metatranskriptome der Proben mithilfe von Hochdurchsatzsequenzierungen beschrieben. Insgesamt konnten für die Probe des Außenohrs etwa 144 Millionen Sequenzen und für die Proben der Achselhöhle etwa 84 Millionen Sequenzpaare generiert werden. Die Datenanalyse erfolgte unter Verwendung der Bioinformatik-Pipelines MG-RAST und METRANS.

Die Analyse des Datensatzes aus dem Außenohr zeigte ein für Otitis externa typisches taxonomisches Profil, welches von *Staphylococcus aureus* dominiert wurde. Jedoch konnten auch *Candida*-Spezies beschrieben werden, welche ebenfalls als sekundäre Pathogene im Zusammenhang mit Außenohrentzündung bekannt sind. Im Rahmen der funktionalen Analyse wurden spezifische Pathogenitätsfaktoren aus *S. aureus* sowie mögliche Pathogenitätsfaktoren aus *Candida* spp. beschrieben. Funktionen, die im Zusammenhang mit einer Kooperation zwischen *Candida* und *S. aureus* stehen, konnten nicht erfasst werden. Faktoren des menschlichen Wirtsorganismus, die im Zusammenhang mit der Erkennung von Pathogenen stehen, konnten ebenfalls dargestellt werden.

Das taxonomische Profil der aktiven Mikrobiota der Achselhöhle auf Genus-Ebene bestätigte die Beschreibungen aus kultivierungsabhängigen Daten in der Literatur und zeigte, dass die Lebensgemeinschaften entweder von Staphylokokken oder von Corynebakterien dominiert

werden. Dabei weisen letztere höhere Diversitäten und Äquitäten auf. Auf Spezies-Ebene konnte eine positive Korrelation des Körpergeruches zu Corynebakterien und eine negative zu Staphylokokken beschrieben werden. Darüber hinaus wurden inter- und intraindividuelle Unterschiede sowie zeitliche Variabilitäten der Zusammensetzung der Lebensgemeinschaften und Änderungen in Abhängigkeit von Deodorant- und Wasseranwendung dargestellt. Dabei waren staphylokokkendominierte Lebensgemeinschaften weniger variabel und somit stabiler im zeitlichen Verlauf und in Abhängigkeit von den Umgebungsbedingungen als Lebensgemeinschaften, die von Corynebakterien dominiert wurden.

Die allgemeine funktionale Analyse der Achselhöhlen-Mikrobiota zeigte im Gegensatz zur taxonomischen weniger inter- und intraindividuelle Variabilitäten sowie Variabilitäten in Abhängigkeit von der Zeit und den Umgebungsbedingungen. Es konnten keine 100%igen Sequenzidentitäten zu den aus der Literatur bekannten Funktionen im Zusammenhang mit Körpergeruchsentstehung aus *Corynebacterium striatum* beschrieben werden. Dies deutet darauf hin, dass andere Funktionen aus z. T. unkultivierten Organismen mit der Körpergeruchsentstehung im Zusammenhang stehen. Es konnten Sequenzen mit Homologie zur N<sup>α</sup>-Acylglutamin-Aminoacylase aus *C. striatum* (N-AGA) beschrieben werden. Diese stammen jedoch aus bislang unbekanntem Organismen und weisen Sequenzhomologien von unter 80 % zur N-AGA auf. Diese Sequenzen könnten somit Transkripte körpergeruchsverursachender Aminoacylasen aus unkultivierten Corynebakterien darstellen und zum *Screening* nach neuen Wirksubstanzen gegen die Körpergeruchsentstehung eingesetzt werden. Darüber hinaus wurden MalY-Sequenzen aus Corynebakterien und Staphylokokken beschrieben, welche mögliche Kandidaten für Cystathionin-β-Lyase-Aktivität in der Achselhöhle darstellen. Sequenzen mit Homologie zu corynebakteriellen TpdA-Dipeptidasen wurden ebenfalls erfasst.

Der partielle Abbau von Fettsäuren durch den Stoffwechselweg der β-Oxidation ist wahrscheinlich nicht an der Körpergeruchsentstehung beteiligt. Auch Sequenzhomologien zu Funktionen mit Beteiligung an der Entstehung geruchsbildender Steroide wurden nicht in den Datensätzen erfasst. Sequenzhomologien zu BCAA-Aminotransferasen sowie L-Laktat-Dehydrogenasen und Acetat-Kinasen konnten in den vorliegenden Datensätzen beschrieben werden, sodass eine Beteiligung dieser Funktionen an der Körpergeruchsentstehung möglich ist. Insgesamt werden Funktionen, die positiv mit dem Körpergeruch korrelieren, überwiegend Corynebakterien und Funktionen, die negativ mit dem Körpergeruch korrelieren, überwiegend Staphylokokken zugeordnet. Mit dem Körpergeruch positiv korrelierte Funktionen, die im Zusammenhang mit geruchsbildenden Metaboliten stehen, wurden im Leucin-Abbauweg und im Phenylalanin-Stoffwechsel identifiziert.

Die vorliegende Arbeit stellt die erste Beschreibung der Lebensgemeinschaften des Außenohres und der Achselhöhle auf Metatranskriptom-Ebene dar. Es werden bekannte und bislang unbekannte Organismen und Funktionen im Zusammenhang mit Außenohrentzündung und Körpergeruchsentstehung und neue Erkenntnisse zu diesen Phänomenen beschrieben. Dies verdeutlicht das Potential der Metatranskriptomik zur Beschreibung physiologischer Vorgänge und als Grundlage zur Entwicklung neuer kosmetischer oder pharmazeutischer Wirksubstanzen.



## II. Einleitung

### 1 Die Mikrobiota des Menschen

Der menschliche Körper ist in seinem Inneren steril, während die physiologisch nach außen gerichteten Körperoberflächen, wie die Haut, die Bindehaut des Auges, der Verdauungs-, Urogenital- und Atemtrakt, von Mikroorganismen besiedelt sind. Schätzungen zufolge liegt die Anzahl dieser mikrobiellen Zellen im Bereich zwischen  $10 \times 10^{18}$  und  $100 \times 10^{18}$  und übersteigt damit die Anzahl somatischer Zellen etwa um einen Faktor 10 (Turnbaugh et al., 2007). Im Mutterleib ist der embryonale Körper gänzlich steril und es kommt erst nach der Geburt im Laufe der ersten Lebensjahre zur Besiedlung und zur Ausbildung der mikrobiellen Lebensgemeinschaften (Morgan und Huttenhower, 2012). Die Gesamtheit aller mikrobiellen Bewohner des menschlichen Körpers wird als Mikrobiota und die Gesamtheit ihrer Gene als Mikrobiom bezeichnet (Lederberg und McCray, 2001). Somit kann der menschliche Körper als ein Supraorganismus verstanden werden, der sich aus den menschlichen Zellen und aus den Zellen der Mikrobiota zusammensetzt (Turnbaugh et al., 2007).

Die menschliche Mikrobiota besteht hauptsächlich aus Mitgliedern der Bakterien, Archaeobakterien, Pilze und Viren (Morgan und Huttenhower, 2012). Eukaryotische Einzeller und mikroskopisch kleine Tiere können in Abhängigkeit von der jeweiligen Körperregion vorhanden sein, werden in der Fachliteratur jedoch nicht tiefgehend betrachtet. Die größte und diverseste Gruppe stellen die Bakterien dar, welche in der Literatur am besten beschrieben sind. Die Beziehungen zwischen den Mitgliedern der Lebensgemeinschaften des menschlichen Organismus können parasitisch, symbiotisch oder neutral sein. Entsprechend handelt es sich bei den einzelnen Mikroorganismen um Pathogene, Symbionten oder Kommensale. Ihre Rolle in der Physiologie des Menschen und in der Entstehung von Erkrankungen ist enorm (Fredricks, 2001; Turnbaugh et al., 2007). So helfen beispielsweise die Mikroorganismen des Darmes bei der Verdauung, dem Abbau von Toxinen und bei der Abwehr von Pathogenen (National Research Council (US) Committee on Metagenomics: Challenges and Functional Applications, 2007).

Jede Körperregion bildet aufgrund ihrer spezifischen Eigenschaften, wie Feuchtigkeit, pH-Wert, Nährstoffverfügbarkeit oder Mechanismen der Immunabwehr, eine distinkte ökologische Nische und definiert damit die Zusammensetzung der Mikrobiota an der jeweiligen Körperstelle (Grice und Segre, 2011). Die Mitglieder der Mikrobiota sind entweder resident oder aufgrund von Interaktionen mit anderen Körperstellen, Individuen oder Gegenständen transient (Cogen et al., 2008). Auch intra- und interspezifische Wechselwirkungen innerhalb der Mikrobiota existieren.

Die Zusammensetzung der Lebensgemeinschaften differenziert sich somit einerseits regional oder temporal innerhalb eines Individuums intraspezifisch, andererseits zwischen einzelnen Individuen, d. h. interspezifisch (Grice und Segre, 2011).

## **1.1 Beschreibung der humanen Mikrobiota mittels Hochdurchsatzsequenzierung**

In der Vergangenheit wurde die menschliche Mikrobiota überwiegend durch die Anwendung kultivierungsbasierter Techniken, wie Färbungen oder Bestimmungen metabolischer Eigenschaften beschrieben. Jedoch ist mit einem geschätzten Anteil von etwa 99 % der Großteil der bakteriellen Spezies unter Laborbedingungen nicht kultivierbar (Lasken und McLean, 2014; Pace et al., 1986). Dies hat dazu geführt, dass nur wenige, leicht kultivierbare Vertreter der Mikrobiota bisher beschrieben wurden und als Modellorganismen dienen, während die meisten anderen Bakterien und ihre Funktionen am menschlichen Organismus bisher unbekannt sind (Morgan und Huttenhower, 2012). Erst seit den 1980er Jahren im Zuge der Entwicklung von Sequenzieretechniken und ihrer Anwendung für phylogenetische Studien ist es möglich, die Kultivierung einzelner Mikroorganismen zu umgehen und die Mitglieder der menschlichen Mikrobiota direkt aus einzelnen Proben heraus auf Sequenzebene zu beschreiben (Morgan und Huttenhower, 2012). Bei der damals vorherrschenden Sequenzieretechnik handelte es sich um die Sanger-Sequenzierung.

Die erstmalige Sequenzierung des menschlichen Genoms zu Beginn des neuen Jahrtausends wird als Wendepunkt bezeichnet, der zu weiteren Innovationen in der Sequenzieretechnologie und der Entwicklung des sogenannten *Next Generation Sequencing* (NGS) führte (Nikolaki und Tsiamis, 2013). Derzeit stehen vier Sequenzierplattformen der sogenannten *2<sup>nd</sup> Generation* zur Verfügung (Box 1) (Nikolaki und Tsiamis, 2013). Im Vergleich zur Sanger-Sequenzierung ist diesen Sequenzieretechniken die parallele Sequenzierung sehr großer Anzahlen einzelner DNA-Moleküle gemeinsam (Nikolaki und Tsiamis, 2013; Pareek et al., 2011). Dabei werden einzelne Moleküle zunächst positionsspezifisch auf einer Oberfläche immobilisiert und amplifiziert. Anschließend erfolgt das Auslesen spezifischer Signale in Form von Licht oder einer Spannungsänderung während des Einbaus von Nukleotiden in den Folgestrang. Der Sequenzdurchsatz der neuen Techniken ist mindestens um den Faktor 100 größer als bei der Sanger-Sequenzierung jedoch sind die Leseweiten geringer (Mardis, 2008; Metzker, 2010). Die Technologien entwickeln sich sehr schnell weiter und es stehen bereits sogenannte *3<sup>rd</sup> Generation Sequencing* Technologien zur Verfügung, bei denen einzelne DNA-Moleküle in Echtzeit sequenziert werden (Eid et al., 2009).

Jede Sequenzierplattform besitzt Vor- und Nachteile im Hinblick auf Leseweiten, Dauer, Sequenzqualität und Kosten, die für bestimmte Anwendungen gegeneinander aufgewogen werden müssen, um die passende Sequenziertechnik zu wählen. In den folgenden Abschnitten werden die Techniken und Studien zur Beschreibung der menschlichen Mikrobiota unter Verwendung von NGS-Technologien umrissen.

### 1.1.1 16S-rDNA-Amplikon-Sequenzierung

Die Gensequenz der 16S-rRNA, d. h. der RNA der kleinen ribosomalen Untereinheit, stellt einen sehr guten Biomarker für phylogenetische Analysen von Prokaryoten dar und ermöglicht die kultivierungsunabhängige Beschreibung einzelner Organismen und ganzer Lebensgemeinschaften durch direkte Amplifikation aus Probenmaterial Mithilfe PCR-basierter molekularer Techniken (Pace et al., 1986). Sowohl DNA-Fingerprint-Techniken wie Gradientengelelektrophorese, Restriktionsfragmentlängenpolymorphismus als auch spezifischere Verfahren wie FISH (*Fluorescence in Situ Hybridization*) wurden auf Grundlage der 16S-rDNA etabliert und eignen sich für die Beschreibung mikrobieller Lebensgemeinschaften. Jedoch stellt die Sequenzierung des 16S-rDNA-Gens oder einzelner Ausschnitte die bedeutendste und für die taxonomische Zuordnung am besten geeignete Methode dar (Nikolaki und

## Hochdurchsatz-Sequenziertechniken

### 454-Plattform:

- Amplifikation: Emulsions-PCR
- Sequenzierung und Detektion: Pyrosequenzierung (indirekte Detektion von Pyrophosphat durch Licht aus der Luciferase-Reaktion)
- Leseweite: 200 - 1000 Basen
- Durchsatz pro Lauf: 70 Mb – 700 Mb ( $10^5$  -  $10^6$  Reads)
- Dauer pro Lauf: 18 h – 23 h

### Illumina-Plattform:

- Amplifikation: Polony-Amplifikation
- Sequenzierung und Detektion: Reversible Termination (Nukleotide mit 3'-O-Fluorophor)
- Leseweite: 2 x 150 Basen – 2 x 300 Basen
- Durchsatz pro Lauf: 7,5 Gb – 18 Gb ( $25 \times 10^6$  -  $6 \times 10^9$  Reads)
- Dauer pro Lauf: 4 h – 3 Tage

### SOLiD-Plattform:

- Amplifikation: Emulsions-PCR
- Sequenzierung (Detektion): Ligation und reversible Termination (fluoreszenzmarkierte Nukleotide)
- Leseweite: 35 - 50 Basen
- Durchsatz pro Lauf: 120 Gb
- Dauer pro Lauf: 7 Tage – 8 Tage

### Ion Torrent

- Amplifikation: Emulsions-PCR
- Sequenzierung und Detektion: Detektion von Protonen beim Einbau von Nukleotiden
- Leseweite: ~ 200 Basen
- Durchsatz pro Lauf: 30 Mb – 2 Gb ( $4 \times 10^5$  -  $5,5 \times 10^6$  Reads)
- Dauer pro Lauf: 2,3 h – 7,3 h

**Box 1: Übersicht der Sequenziermechanismen und Leitungskennzahlen der vier NGS Plattformen.**

Tsiamis, 2013). Die Grundlage für die Eignung der 16S-rDNA als phylogenetischer Biomarker liefert ihre Sequenz. Sie besteht aus konservierten und variablen Abschnitten, wobei insgesamt 8 hypervariable Bereiche (V1-V8) mit einer Länge von jeweils 50 bp – 100 bp beschrieben sind. Die Variabilität dieser Sequenzabschnitte steht im Zusammenhang mit der phylogenetischen Verwandtschaft einzelner Spezies. Die hypervariablen Bereiche werden jeweils von konservierten Bereichen flankiert, was das Design universaler PCR-Primer und damit die Amplifikation hypervariabler Bereiche aus vielen verschiedenen Spezies mit nur einem Primerpaar ermöglicht (Baker und Banfield, 2003; Katsaveli et al., 2012; Nossa et al., 2010). Aufgrund der Amplifikation mittels PCR-Techniken können auch relativ kleine Proben analysiert werden.

Die Sequenz des 16S-rRNA-Gens wurde häufig zur Beschreibung des menschlichen Mikrobioms verwendet (Bouslimani et al., 2015; Costello et al., 2009; Grice et al., 2009; Human Microbiome Project Consortium, 2012; Oh et al., 2013). Grice et al. beschrieben 2009 mittels Sanger-Sequenzierung klonierter 16S-rRNA-Gene die Zusammensetzung der bakteriellen Lebensgemeinschaften der Haut an 20 unterschiedlichen Körperstellen bei 10 Probanden zu zwei Zeitpunkten (Grice et al., 2009). Insgesamt wurden 112.283 Sequenzen analysiert, die nahezu die gesamte Länge des 16S-rDNA-Gens abdeckten. Mithilfe von NGS-Techniken können jedoch größere Anzahlen von Amplikons direkt (ohne Klonierung) sequenziert werden. Costello et al. nutzten die Pyrosequenzierung zur Beschreibung der Mikrobiota von 9 Probanden an jeweils 27 Körperstellen zu 4 Zeitpunkten und erzeugten dafür mehr als 1.070.000 Reads aus Amplikons der V2-Region (Costello et al., 2009). Besonders umfangreich sind die Studien im Rahmen des *Human Microbiome Project* (HMP), finanziert durch die *US National Institutes of Health* (Human Microbiome Project Consortium, 2012; Turnbaugh et al., 2007). Insgesamt wurden 242 Probanden an jeweils 15 (Männer) bzw. 18 (Frauen) Körperstellen zu zwei Zeitpunkten beprobt und es wurden zusätzlich Stuhlproben untersucht, welche repräsentativ für die Mikrobiota des Verdauungstraktes sind. Mittels Pyrosequenzierung wurden im Mittel 5.408 Reads pro Probe generiert. Insgesamt wurden die mikrobiellen Lebensgemeinschaften sehr vieler verschiedener Körperstellen des Hals-Nasen-Rachen-Raumes, Darmes, Urogenitalbereiches und sehr viele unterschiedliche Stellen der Haut beschrieben. Diese Lebensgemeinschaften unterscheiden sich hinsichtlich der  $\alpha$ -Diversität (innerhalb einer Lebensgemeinschaft) und  $\beta$ -Diversität (zwischen verschiedenen Lebensgemeinschaften). Dabei scheint die vaginale Mikrobiota eine sehr geringe  $\alpha$ -Diversität zu besitzen, während Speichel eine besonders hohe aufweist (Human Microbiome Project Consortium, 2012; Li et al., 2013). Die interpersonale Variabilität einzelner Lebensgemeinschaften ist in der Regel höher als die intrapersonale Variabilität zu unterschiedlichen Zeitpunkten (Costello et al., 2009; Grice et al.,

2009). Auch Variabilitäten der Lebensgemeinschaften in Abhängigkeit von Erkrankungen oder Hygienegewohnheiten wurden beschrieben (Bousslimani et al., 2015; Oh et al., 2013).

### 1.1.2 Metagenomik

Die Beschreibung von Lebensgemeinschaften mittels 16S-rDNA-Sequenzierung stellt eine bedeutende Technik in der Erforschung des humanen Mikrobioms dar. Jedoch handelt es sich um eine markerbasierte Methode, die zwar Aussagen über die taxonomische Zusammensetzung der Lebensgemeinschaften zulässt, jedoch keine Informationen über die physiologischen und metabolischen Funktionen liefert. Es wird angenommen, dass die genetische Information der menschlichen Mikrobiota die des menschlichen Körpers um einen Faktor 100 übersteigt (Petrosino et al., 2009). Die Beschreibung der Funktionen der menschlichen Mikrobiota ist daher sehr interessant vor allem im Hinblick auf die Entstehung von Krankheiten (Kuczynski et al., 2012). Die Gesamtheit der Gene innerhalb einer Lebensgemeinschaft wird als Metagenom bezeichnet, und die Metagenomik stellt eine molekulare Technologie zur Untersuchung und Beschreibung des Metagenoms dar. Grundsätzlich wird unterschieden zwischen gezielten Methoden, bei denen spezifische DNA-Fragmente amplifiziert werden – weshalb 16S-rDNA-Amplicon-Sequenzierung z. T. auch als Metagenomik-Ansatz betrachtet wird – und ungerichteten, sogenannten *Shotgun*-Verfahren, bei denen stochastisch DNA-Abschnitte sequenziert werden (Kuczynski et al., 2012). Der entscheidende Vorteil von *Shotgun*-Metagenomik-Ansätzen Mithilfe von NGS-Techniken ist die Möglichkeit der Beschreibung bislang unbekannter Gene, da keine bekannten Zielsequenzen benötigt werden. Ein Nachteil ist, dass relativ große Mengen an DNA benötigt werden, die aus Habitaten wie der Haut nicht isoliert werden können (Petrosino et al., 2009). Die *Whole Genome Amplifikation* (WGA) bietet eine Möglichkeit, Restriktionen durch geringe DNA-Mengen zu umgehen, birgt jedoch die Gefahr, Verzerrungen durch die DNA-Amplifikation herbeizuführen.

Im Rahmen des *Human Microbiome Project* wurden 681 Proben im *Shotgun*-Verfahren auf der Illumina-Plattform sequenziert und Stoffwechselwege rekonstruiert (Human Microbiome Project Consortium, 2012). Auch im Rahmen des europäischen Projektes *MetaHIT* wurde das intestinale Metagenom von 124 Individuen untersucht und die kodierten Funktionen wurden beschrieben (Qin et al., 2010). Die beiden Studien zeigten sehr große interindividuelle Ähnlichkeiten hinsichtlich der funktionalen Zusammensetzung der Mikrobiota einzelner Habitate. In einigen wenigen Habitaten in der Mundhöhle waren einzelne funktionale Gruppen im Vergleich zu anderen Habitaten überrepräsentiert. Jedoch setzten sich die meisten Metagenome aus Funktionen des Primärstoffwechsels, wie Translation, Nukleotid- und ATP-Synthese und Glykolyse zusammen. Es ist daher zu vermuten, dass uncharakterisierte

Funktionen mit geringer *Read*-Abundanz in den Datensätzen die tatsächlich beobachtbaren Unterschiede zwischen den Funktionen der Lebensgemeinschaften beeinflussen (Human Microbiome Project Consortium, 2012).

### 1.1.3 Metatranskriptomik

Während die Metagenomik die Gesamtheit der Gene und somit das funktionale Potential einer Lebensgemeinschaft beschreibt, bleibt weiterhin offen, welche dieser Funktionen zum gegebenen Zeitpunkt aktiv sind. Nur die aktiven Mitglieder der Mikrobiota und ihre spezifischen Funktionen charakterisieren eine Lebensgemeinschaft und beeinflussen tatsächlich ihr Habitat, bzw. ihren Wirt. Analog zur Entwicklung von der Genomik zur Transkriptomik vollzieht sich aktuell – gestützt von NGS-Technologien - die Entwicklung von der Metagenomik zur Metatranskriptomik. In einigen Fällen werden sogar sogenannte Multiomik-Ansätze verfolgt, die zusätzlich das Metaproteom und das Metametabolom beschreiben (Bouslimani et al., 2015; Pérez-Cobas et al., 2013). Im Vergleich zu etablierten Methoden der Transkriptomik, wie beispielsweise qPCR oder Microarrays, bietet die Metatranskriptomik die Möglichkeit, theoretisch unbegrenzte Anzahlen von u. U. unbekanntem Genen zu beschreiben, da keine Zielsequenzen benötigt werden (Moran, 2009). Komparative Ansätze erlauben Rückschlüsse auf mikrobielle Funktionen, die im Zusammenhang mit bestimmten physiologischen Eigenschaften des Wirtsorganismus oder mit Erkrankungen im Zusammenhang stehen.

Metatranskriptomik-Analysen der menschlichen Mikrobiota umfassen die Lebensgemeinschaften des Darmes und der Vagina (Gosalbes et al., 2011; Twin et al., 2013). Twin et al. untersuchten eine vaginale Probe einer Patientin mit bakterieller Vaginose und fanden eine Dominanz der Bakterienspezies *Prevotella amnii* (Twin et al., 2013). Die Korrelation dieser Spezies mit bakterieller Vaginose konnte anschließend in Analysen von Proben mehrerer Probanden mittels qPCR bestätigt werden. Darüber hinaus wurden abundante Funktionen der vaginalen Mikrobiota, welche im Wesentlichen aus Funktionen des Protein-, Kohlenhydrat- und Lipidstoffwechsels bestanden, beschrieben. Gosalbes et al. beschrieben die Zusammensetzung und Aktivität der aktiven intestinalen Mikrobiota von zehn Probanden Mithilfe der 454-Sequenzierung (Gosalbes et al., 2011). Transkripte der 16S-rRNA-Gene wurden zur komparativen taxonomischen Beschreibung der Lebensgemeinschaften auf interindividueller Ebene genutzt. Die interindividuelle Beschreibung der mRNA-Sequenzen zeigte ebenfalls ein sehr uniformes Muster des Primärstoffwechsels.

Eine der Herausforderungen in der praktischen Umsetzung von Metatranskriptomik-Ansätzen der menschlichen Mikrobiota stellt die Tatsache dar, dass 90 % - 99 % der gesamten RNA mikrobieller Lebensgemeinschaften durch rRNA beschrieben wird, die keine funktionalen

Informationen liefert, bzw. die Sequenziertiefe von mRNAs herabsetzt. Mögliche Lösungen sind die aktive Abreicherung von rRNA oder das Nutzen der rRNA-Sequenzen zur taxonomischen Analyse der aktiven Mikrobiota. Weiterhin stellt die Isolierung hinreichend großer Mengen an RNA eine große Hürde zur Untersuchung der aktiven Mikrobiota des Menschen dar. Daher ist die Metatranskriptomik des Menschen bisher insgesamt wenig untersucht.

## **2 Die mikrobielle Lebensgemeinschaft des menschlichen Ohres und Gehörgangsentzündung**

### **2.1 Gehörgangsentzündung – Krankheitsbild, Entstehung und Behandlung**

Bei Gehörgangsentzündung (*otitis externa*) handelt es sich um eine diffuse, in der Regel bakterielle Entzündung des äußeren Gehörganges. Sie wird häufig auch als Außenohrentzündung, Schwimmerohr oder als tropisches Ohr bezeichnet (Lee et al., 2013; Rosenfeld et al., 2014; Schaefer und Baugh, 2012). Die akute Form der Erkrankung zeichnet sich durch eine hohe Morbidität aus. Die jährliche Prävalenz in den USA liegt bei 1,2 % bis 1,3 %, wobei Kinder im Alter zwischen 7 und 12 Jahren und Erwachsene im Alter zwischen 65 und 77 Jahren am häufigsten betroffen sind (Beers und Abramo, 2004; Hajioff und Mackeith, 2010; Rowlands et al., 2001). Im Verlauf des Jahres ist die Häufigkeit der Erkrankung im Spätsommer am höchsten und im Winter am niedrigsten (Rosenfeld et al., 2014). Schätzungen zufolge liegen die in den USA unmittelbar durch die Erkrankung verursachten Kosten bei etwa 600 Milliarden Dollar (Centers for Disease Control und Prevention (CDC), 2011).

Der äußere Gehörgang besitzt eine Länge von etwa 2,5 cm und einen Durchmesser von 7 mm bis 9 mm (Guss und Ruckenstein, 2010). Das äußere (distale) Drittel des Gehörganges ist knorpelig und beherbergt apokrine Drüsen, Talgdrüsen und Haarfollikel. Dahingegen sind die inneren (proximalen) zwei Drittel knöchig, besitzen kein Unterhautgewebe und beherbergen eine dichte Anordnung von Nervenbahnen. Der Übergang zwischen den beiden Bereichen stellt in der Regel die engste Stelle im Gehörgang dar. Die Unebenmäßigkeit des Gehörganges begünstigt die Anhäufung von Ablagerungen während einer Infektion (Lee et al., 2013). Das Cerumen oder Ohrenschmalz setzt sich zusammen aus Sekreten apokriner Drüsen und epitheliale Schorf. Es bildet einen hydrophoben, leicht sauren Schutzfilm im Gehörgang, der die Aufweichung des Gewebes bei Kontakt mit Wasser verhindert und bakterielles Wachstum inhibiert (Francis, 2010).

Die Entstehung von Gehörgangsentzündungen wird häufig durch das Eindringen von Wasser in den Gehörgang oder durch kleine Verletzungen des Gehörganges verursacht und zeichnet sich durch ein schnelles Einsetzen (innerhalb von 48 h) aus. Die Prädispositionsfaktoren für Gehörgangsentzündung umfassen diverse Faktoren (Peterkin, 1974; Rosenfeld et al., 2014; Schaefer und Baugh, 2012). Genetische Faktoren umfassen einen engen Gehörgangskanal, der auch durch Auswüchse oder Cysten verengt sein kann, sowie einen stark behaarten Gehörgang, Störungen bei der Cerumenproduktion und die Blutgruppe A. Die Umweltfaktoren umfassen das Schwimmen sowie längeren Wasserkontakt, Schwitzen, hohe Luftfeuchtigkeit und Wärme aber auch das Eindringen von Seife in den Gehörgang und Stress im Allgemeinen. Traumata führen zur Zerstörung der schützenden Cerumen-Schicht und werden in der Regel ausgelöst durch Wattestäbchen, Gehörhilfen, Ohrstöpsel und andere Fremdkörper. Auch dermatologische Grunderkrankungen wie Ekzem, Schuppenflechte (Psoriasis), seborrhoisches Ekzem und andere entzündliche Dermatosen begünstigen die Entstehung von Gehörgangsentzündung. Auch eitriger Ausfluss, der im Zusammenhang mit einer Außenohrentzündung auftritt, kann zur Entstehung einer Außenohrentzündung beitragen.

Die Symptome sind starke Schmerzen, Juckreiz, Schwellung und / oder Rötung im Gehörgang sowie Druckschmerz der Ohrmuschel und des Tragus. Darüber hinaus kann es zu Ohrenfluss (Otorrhoe), Rötung des Trommelfells, Entzündung des Unterhautgewebes (Cellulitis) der Ohrmuschel, Lymphknotenschwellung, Schmerzen im Kiefer und Hörverlust kommen (Rosenfeld et al., 2014). Die Diagnose wird in der Regel im Rahmen einer ärztlichen Untersuchung Mithilfe eines Otoskops oder mikroskopischer Verfahren anhand der Symptome gestellt. Im Falle eines perforierten Trommelfells kann die Abgrenzung zur Außenohrentzündung jedoch erschwert sein (Rosenfeld et al., 2014).

Bei der Behandlung von Gehörgangsentzündungen spielt die Schmerzlinderung eine bedeutende Rolle. Bei leichten Schmerzen wird Paracetamol verabreicht und bei starken Schmerzen kommen Opioid-Analgetika zum Einsatz (Lee et al., 2013). Gegebenenfalls wird auch eine Reinigung des Gehörganges durchgeführt. Die Behandlung der Entzündung erfolgt meist durch die Anwendung topischer antimikrobieller Substanzen, häufig in Kombination mit Corticosteroiden (Kaushik et al., 2010). Die Anwendung topischer Antibiotika wird gegenüber der systemischen Gabe aufgrund der geringeren Nebenwirkungen und der Vermeidung der Selektion resistenter Organismen bevorzugt (Rosenfeld et al., 2014). Verwendete Antibiotika umfassen Aminoglykoside, Polymyxin B und Quinolone (Schaefer und Baugh, 2012). Auch Acetat wird, z. T. in Kombination mit anderen Substanzen in der alternativen Therapie, erfolgreich zur Behandlung von Gehörgangsentzündung angewendet (Lee et al., 2013; Schaefer und Baugh,



2012). Bei erfolgreicher Therapie werden die Symptome innerhalb von einem Tag deutlich gelindert. Die Anwendung topischer Antibiotika erfolgt in der Regel für eine Dauer von 7 bis 10 Tagen (Schaefer und Baugh, 2012).

## 2.2 Die Mikrobiologie der Gehörgangsentzündung

Analog zu anderen Körperregionen beherbergt der äußere Gehörgang eine mikrobielle Lebensgemeinschaft mit einer charakteristischen Zusammensetzung. Die Mikrobiota des gesunden Gehörganges wurde sowohl kultivierungsbasiert als auch kultivierungsunabhängig Mithilfe von Hochdurchsatzsequenzierung beschrieben und ähnelt hinsichtlich ihrer Zusammensetzung den Lebensgemeinschaften talghaltiger Hautpartien (Costello et al., 2009; Stroman et al., 2001). Die kultivierungsbasierte Beschreibung der Lebensgemeinschaften des Gehörganges und Cerumen von 164 Probanden von Stroman et al. 2009 zeigte, dass die Lebensgemeinschaft überwiegend (zu 99 %) aus Gram-positiven Bakterien besteht. Darunter sind Staphylokokken mit einer mittleren relativen Abundanz von 66 % die häufigsten Vertreter. Darüber hinaus wurden coryneforme Bakterien beschrieben, wobei *Turicella otitidis* die dominanteste Spezies darstellte. Unter den Streptokokken-ähnlichen Bakterien war *Alloiococcus otitis* der dominanteste Vertreter. Unter den wenigen Gram-negativen Vertretern befanden sich vier Stämme von *Pseudomonas aeruginosa*. Später zeigten Costello et al. Mithilfe kultivierungsunabhängiger Untersuchungen, dass sich die Mikrobiota des Außenohres aus *Propionibacterineae*, Staphylokokken, *Corynebacterineae*, *Micrococcineae*, Streptokokken und anderen Bakterien zusammensetzt. Die Kern-Mikrobiota, die alle untersuchten Probanden beherbergen, wird dabei von der Familie der *Propionibacterineae* und der Gattung *Staphylococcus* gebildet.

Otitis externa stellt in der Regel eine bakterielle Entzündung dar, wobei *Pseudomonas aeruginosa* und *Staphylococcus aureus* am häufigsten isoliert werden (Lee et al., 2013; Rosenfeld et al., 2014; Schaefer und Baugh, 2012). Etwa 33 % bis 50 % der Erkrankungen sind jedoch polymikrobiell (Lee et al., 2013; Rosenfeld et al., 2014). Weitere Isolate im Zusammenhang mit Gehörgangsentzündung stellen im Wesentlichen aerobe Bakterienarten wie *Staphylococcus epidermidis*, *Mycobacterium otitidis*, *Mycobacterium alconae*, *Staphylococcus capitis* und *Staphylococcus haemolyticus* dar (Clark et al., 1997; Klein, 2009; Lee et al., 2013). Weitere seltenere Isolate umfassen *Staphylococcus caprae*, *Staphylococcus auricularis*, *Enterococcus faecalis*, *Enterobacter cloacae*, *Acinetobacter calcoaceticus*, *Proteus mirabilis*, *Bacteroides fragilis*, *Peptostreptococcus magnus* und verschiedene Streptokokken-Stämme (Clark et al., 1997; Ninkovic et al., 2008; Roland und Stroman, 2002). Selten wurden auch anaerobe Bakterien isoliert (Clark et al., 1997). In tropischen oder subtropischen Lebensräumen

oder in Folge einer Behandlung von Gehörgangsentzündung mit Antibiotika kann es zu einer Sekundärinfektion mit Pilzen der Gattungen *Candida* und *Aspergillus* kommen (Ahmad et al., 2007; Martin et al., 2005; McDowall, 1974; Pontes et al., 2009). Martin et al. beschreiben positive Pilzisolat aus 166 von 1242 Pädiatrie-Patienten und eine signifikant erhöhte Inzidenz im Zeitraum nach der Vermarktung fluoroquinolonhaltiger Ohrentropfen im Vergleich zum Zeitraum vor der Vermarktung (Martin et al., 2005). Die häufigsten Isolate waren *Candida albicans*, *Candida parapsilosis* und *Aspergillus fumigatus*, sowie weitere *Candida*- und *Aspergillus*-Spezies in sehr geringer Abundanz. Zum Zeitpunkt des Verfassens dieser Arbeit lag keine kultivierungsunabhängige Beschreibung der mikrobiellen Lebensgemeinschaften im Zusammenhang mit Otitis externa vor.

Sowohl bakterielle als auch pilzliche Pathogene der Gehörgangsentzündung konnten auf Hörapparaten, auch nach einer Reinigung mit einer 70%igen Alkohollösung, und in den Gehörgängen von Menschen mit Hörapparaten nachgewiesen werden (Ahmad et al., 2007; Karaca et al., 2013). Die Benutzung von Hörapparaten kann somit die Mikrobiota des äußeren Gehörganges verändern und zu Gehörgangsentzündung führen.

### **3 Die Mikrobiota der menschlichen Achselhöhle und Körpergeruchsentstehung**

Die Haut ist mit etwa 1,8 m<sup>2</sup> das flächenmäßig größte Organ des Menschen und wird charakterisiert vor allem durch den leicht sauren pH-Wert von etwa 5,5, einen hohen Lipidgehalt, dem konstanten Verlust von Korneozyten des *stratum corneum*, Einbuchtungen in der Regel in Form von Drüsen und dem Abwehrsystem des Wirtsorganismus in Form von Proteasen, Lysozym und antimikrobiellen Peptiden (Cogen et al., 2008; Kong, 2011; Percival et al., 2012). In Analogie zu anderen natürlich vorkommenden Lebensräumen kann die Hautoberfläche in unterschiedliche topographische Regionen unterteilt werden (Grice und Segre, 2011). Diese Regionen unterscheiden sich hinsichtlich der Parameter Feuchtigkeit, Nährstoffverfügbarkeit und dem Vorkommen antimikrobieller Peptide (Fredricks, 2001; Percival et al., 2012). Grundsätzlich wird in der Literatur unterschieden zwischen trockenen Bereichen der Haut, wie beispielsweise dem Schienbein oder dem Unterarm, talghaltigen Bereichen, wie dem Kopf und dem oberen Rückenbereich, und feuchten verdeckten Bereichen, wie der Leistengegend oder der Achselhöhle (Grice et al., 2009; Grice und Segre, 2011).

Die mikrobielle Lebensgemeinschaft der Haut besteht aus Bakterien, Archaeobakterien, Pilzen der Gattung *Malassezia*, Viren der Gattung *Microvirus* und Milben der Gattung *Demodex*, wobei die Bakterien die größte und diverseste taxonomische Gruppe darstellen (Chiller et al., 2001; Cogen et al., 2008; Fredricks, 2001; Marples, 1956; Noble, 1984; Roth und James, 1988, 1989). Die vier Bakterientaxa *Actinobacteria*, *Firmicutes*, *Bacteroidetes* und *Proteobacteria* wurden als häufigste identifiziert (Grice und Segre, 2011). Dabei werden auf der Genus-Ebene talghaltige Hautbereiche von Propionibakterien und feuchte Hautbereiche von Staphylokokken und Corynebakterien dominiert, während trockene Bereiche der Haut eine sehr diverse Zusammensetzung aufweisen (Grice und Segre, 2011). Einige Organismen konnten mit bestimmten Erkrankungen der Haut in Zusammenhang gebracht werden. So steht die Gattung *Malassezia* im Zusammenhang mit dem Krankheitsbild der seborrhoischen Dermatitis, *Staphylococcus aureus* mit der atopischen Dermatitis und *Propionibacterium acnes* mit der Akne (Grice und Segre, 2011; Kong, 2011). Auch die Entstehung von Körpergeruch, der in der Regel vor allem ein kosmetisches Problem darstellt, steht mit der Achselhöhlen-Mikrobiota im Zusammenhang.

Die Achselhöhle stellt einen zeitweise verdeckten und aufgrund einer hohen Dichte an Schweißdrüsen feuchten Lebensraum dar. Grundsätzlich wird in der Literatur zwischen

## Schweißdrüsen der Achselhöhle:

### Ekkrine Schweißdrüsen:

- Funktional von Geburt an
- Vorkommen: an nahezu allen Körperstellen
- Physiologische Funktion: Thermoregulation
- Innervierung: cholinerg
- Öffnung mündet in die Hautoberfläche
- Schweißbildung durch Transport über die Zellmembranen im Ductus
- Zusammensetzung ekkriner Schweiß:
  - Wasser
  - Elektrolyte ( $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Ca}^{2+}$ ,  $\text{Mg}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{HCO}_3^-$ ,  $\text{Cu}^{2+}$ ,  $\text{Fe}^{3+}$ )
  - Kleine organische Verbindungen (Acetat, Ammoniak, Ascorbinsäure, Buttersäure, Capronsäure, Caprylsäure, Citrat, Formylsäure, Laktat, Propionat, Harnstoff)

### Apokrine Schweißdrüsen:

- Funktional ab der Pubertät
- Vorkommen: Achselhöhle, Leistengegend, Perineum, Bauchnabel, Brustwarzenhof
- Physiologische Funktion: unklar
- Innervierung: adrenerg
- Öffnung mündet in einen Haarschaft
- Schweißbildung durch Ausschüttung von Zytosol apikaler Zellpartien in den Ductus
- Zusammensetzung apokriner Schweiß: Elektrolyte, Steroide, Proteine, Vitamine, Fette

### Apoekkrine Schweißdrüsen:

- Entwickeln sich ab der Pubertät aus ekkrinen Schweißdrüsen (Hypothese)
- Vorkommen: Achselhöhle, Perineum, Genitalregion
- Physiologische Funktion: Thermoregulation
- Innervierung: cholinerg und adrenerg
- Zusammensetzung apoekkriner Schweiß: ähnlich zu ekkrinem Schweiß

**Box 2: Übersicht der drei Typen von Schweißdrüsen in der Achselhöhle.**

drei unterschiedlichen Formen von Schweißdrüsen unterschieden, wobei alle drei Formen in der Achselhöhle zu finden sind (s. Box 2). Ekkrine Schweißdrüsen sind bereits ab der Geburt funktional und ihre Hauptaufgabe besteht in der Verdunstungskühlung zum Zwecke der Thermoregulation des Körpers (Sato et al., 1989; Wilke et al., 2007). Sie kommen an nahezu jeder Körperstelle vor, wobei ihre Dichte in den Handinnenflächen, Fußsohlen und in der Achselhöhle mit bis zu 600 pro  $\text{cm}^2$  besonders hoch ist (Hölzle, 2002; Johnson und Proppe, 1996). Ihre Öffnung mündet direkt an der Hautoberfläche und sie sekretieren neben Wasser viele Elektrolyte wie  $\text{Na}^+$ ,  $\text{K}^+$ ,  $\text{Mg}^{2+}$ ,  $\text{Cl}^-$ ,  $\text{HCO}_3^-$ , Laktat, Harnstoff und Ammoniak (Harvey et al., 2010; Noël et al., 2012; Wilke et al., 2007). Ekkrine Schweißdrüsen werden cholinerg innerviert. Dahingegen sind apokrine Schweißdrüsen zwar bereits bei der Geburt angelegt, werden jedoch erst mit dem Einsetzen der Pubertät aktiv. Sie liegen immer haarassoziiert vor, da ihre Öffnung in das Innere eines Haarschaftes in definierten Körperregionen wie z. B. der Achselhöhle oder dem Prineum mündet. In der Achselhöhle dominieren die apokrinen Schweißdrüsen im Vergleich zu den ekkrinen mit einem Verhältnis von 10:1 (Sato et al., 1987). Das Sekret dieser Schweißdrüsen ist geruchslos mit einer milchigen bis öligen Konsistenz bestehend aus Elektrolyten, Steroiden, Proteinen, Vitaminen und Fetten (Labows et al., 1979; Leyden et al., 1981; Noël et al., 2012; Wilke et al., 2007). Apokrine Schweißdrüsen werden adrenerg innerviert. Darüber hinaus kommt in der Achselhöhle eine Mischform der haarassoziierten Drüse vor. Die apoekkrinen Schweißdrüsen entwickeln sich aus ekkrinen Drüsen und dienen ebenfalls hauptsächlich der Thermoregulation (Sato et al., 1987; Wilke et al., 2007). Die Nährstoffverfügbarkeit für die Mitglieder der Mikrobiota ist aufgrund der Sekrete der Schweißdrüsen sehr hoch. Dies führt zu einer Bakteriendichte von etwa  $10^6$  Zellen pro  $\text{cm}^2$  Haut (Leyden et al., 1981). Der charakteristische Schweißgeruch der Achselhöhle entsteht durch eine Biotransformation der anfänglich geruchslosen Komponenten des apokrinen Schweißes durch bakterielle Mitglieder der Achselhöhlen-Mikrobiota. Die Entstehung von Körpergeruch ist somit abhängig von der Zusammensetzung und Aktivität der bakteriellen Lebensgemeinschaft in der Achselhöhle.

Bereits in den 1950er Jahren wurde durch das Konzept der Biotransformation die bedeutende Rolle der Hautbakterien bei der Körpergeruchsentstehung in der Achselhöhle beschrieben (Shelley et al., 1953). Heutzutage wird die Forschung zur Entstehung von Schweißgeruch im Wesentlichen von der forschenden Kosmetikindustrie betrieben. Zur Beschreibung der mikrobiellen Lebensgemeinschaften der Achselhöhle und dem Zusammenhang zwischen einzelnen Organismengattungen und Körpergeruch standen lange Zeit nur kultivierungsbasierte Methoden zur Verfügung. Auch molekulare Mechanismen, die der Körpergeruchsentstehung zugrunde liegen, wurden teilweise oder ganz für einzelne Organismen beschrieben. Die

Verfügbarkeit von Hochdurchsatz-Sequenzieretechniken und geeigneten Amplifikationsmethoden für Nukleinsäuren ermöglichte schließlich kultivierungsunabhängige Beschreibungen der Achselhöhlen-Mikrobiota. Im Folgenden werden die Erkenntnisse aus der Literatur zu der Mikrobiota der Achselhöhle und den molekularen Mechanismen der Körpergeruchsentstehung beschrieben.

### **3.1 Kultivierungsbasierte Beschreibungen der Achselhöhlen-Mikrobiota**

Die erste kultivierungsbasierte Beschreibung der Achselhöhlen-Mikrobiota im Zusammenhang mit Körpergeruchsentstehung wurde 1981 von der Arbeitsgruppe um James Leyden et al. veröffentlicht (Leyden et al., 1981). Dabei wurden 229 männliche und weibliche Probanden beprobt und die Proben wurden sowohl unter aeroben als auch unter anaeroben Bedingungen kultiviert. Dabei wurden Mitglieder der Bakterienfamilie der *Micrococcaceae*, zu denen nach der damaligen Taxonomie unter anderem die Gattungen *Micrococcus* und *Staphylococcus* gehörten, in allen untersuchten Proben gefunden. Auf Speziesebene wurde die Verteilung innerhalb dieser Gruppe durch die Spezies *S. epidermidis*, *S. saprophyticus*, *S. aureus* und *Micrococcus* spp. beschrieben. Vertreter der Bakteriengattungen *Corynebacterium* und *Propionibacterium* konnten in 85 % bzw. in 70 % aller Proben nachgewiesen werden. Desweiteren wurden im Rahmen dieser Studie zwei Typen von Körpergeruch beschrieben: ein stechender apokriner Geruch, der mit der Dominanz von Corynebakterien in den entsprechenden Proben einherging, sowie ein schwacher säuerlicher Geruch, der mit Isovaleriansäure und einer durch die *Micrococcaceae* dominierten Mikrobiota assoziiert wurde. In einer Koinkubation von Achselhöhlenisolaten mit apokrinem Schweiß auf dem Unterarm entstand der stechende apokrine Geruch nur in Anwesenheit von Corynebakterien bei einer Zelldichte von  $10^4$  KBE /  $\text{cm}^2$ . Staphylokokken generierten einen anderen Geruch.

In einer späteren Studie bestätigten Taylor et al. im Wesentlichen die Erkenntnisse von Leyden et al. bezüglich der Zusammensetzung der Achselhöhlen-Mikrobiota (Taylor et al., 2003). Es konnte auch hier in kultivierungsabhängigen Ansätzen unter Verwendung selektiver Medien gezeigt werden, dass es sich bei den am häufigsten vorkommenden Bakteriengattungen um Staphylokokken (98 %), Corynebakterien (93 %) und Propionibakterien (88 %) handelt. Dabei wurden etwa 2/3 der untersuchten Lebensgemeinschaften zahlenmäßig von Staphylokokken und 1/3 von Corynebakterien dominiert. Weiterhin konnte im Rahmen einer semiquantitativen Beschreibung des Körpergeruches durch den *Malodor Score* eine Korrelation zwischen dem Körpergeruch und den Keimzahlen aerober Corynebakterien in der entsprechenden Probe gezeigt werden. Keimzahlen anderer Genera korrelierten nicht oder nur schwach. Im Rahmen

einer Charakterisierung corynebakterieller Achselhöhlen-Isolate Mithilfe der 16S-rDNA-Sequenzierung konnten Sequenzhomologien zu den Spezies *Corynebacterium G-2* und *C. mucifaciens* beschrieben werden.

Der im Rahmen der kultivierungsbasierten Beschreibungen gezeigte Zusammenhang zwischen Corynebakterien und unangenehem Körpergeruch steht im Einklang mit den späteren Studien der Arbeitsgruppen um Natsch et al. und James et al., in denen corynebakterielle molekulare Mechanismen der Körpergeruchsentstehung beschrieben werden (s. Kapitel 3.2). Eine Korrelation zwischen Staphylokokken und Körpergeruch konnte bisher nicht beschrieben werden. In einer Untersuchung mit insgesamt 260 Staphylokokken- und Corynebakterien-Stämmen zeigten Kwazewska et al., dass Corynebakterien eine geringere Anpassungskapazität gegenüber den Abwehrmechanismen der Haut, insbesondere dem niedrigen pH-Wert besitzen (Kwazewska et al., 2014). Dies könnte eine Erklärung für die relativ hohe Anzahl Staphylokokken-dominiertes Lebensgemeinschaften in der Achselhöhle liefern.

## **3.2 Molekulare Mechanismen der Körpergeruchsentstehung**

Die bekannten mikrobiellen Biotransformationsrouten in der Achselhöhle, die zur Entstehung von Körpergeruch führen, sind ganz oder zum Teil in der Literatur beschrieben und zusammengefasst (Barzantny et al., 2012a; Fredrich et al., 2013; James et al., 2013; Natsch, 2015). Bei den zugrunde liegenden molekularen Mechanismen handelt es sich im Wesentlichen entweder um (i) katabolische Stoffwechselprozesse, in denen eine geruchslose Ausgangssubstanz partiell abgebaut wird und in eine Substanz mit einem Geruch umgewandelt wird, oder (ii) um eine spezifische Spaltung eines Vorläufermoleküls, i.d.R. eines Konjugates, bei der eine Geruchskomponente freigesetzt wird. Bei den geruchsbildenden Komponenten handelt es sich um kleine volatile Moleküle. Die Beschreibung dieser Substanzen wurde in den vergangenen 21 Jahren durch biochemische und zunehmend durch massenspektrometrische Methoden ermöglicht. Besonders hervorzuheben ist dabei die *Sniff-Port*-Technologie, die die Fraktionierung von Schweiß und die sensorische Untersuchung der einzelnen Fraktionen ermöglicht (Natsch et al., 2003).

### **3.2.1 Steroide**

Die Beteiligung von Steroiden an der Entstehung von Moschus-ähnlichem Körpergeruch wird bereits seit den 1980er Jahren untersucht. Insbesondere die geruchsbildenden Steroid-Derivate  $5\alpha$ -Androst-16-n-3 $\alpha$ -ol und  $5\alpha$ -Androst-16-n-3 $\alpha$ -on besitzen sehr niedrige Wahrnehmungsschwellen und wurden mittels GC-MS in apokrinem Schweiß männlicher

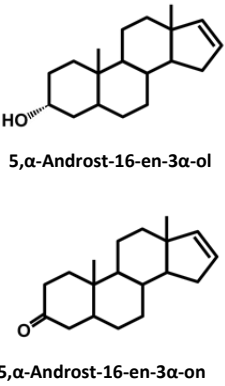
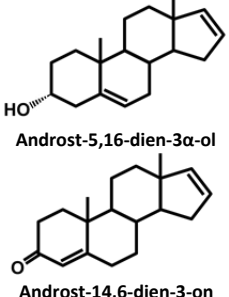
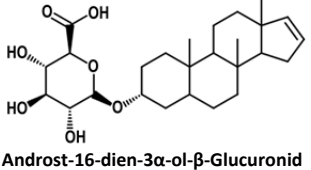
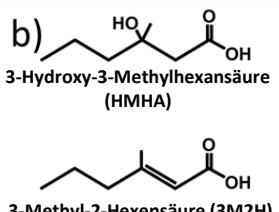
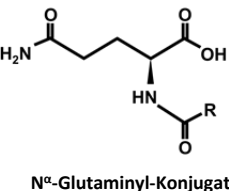
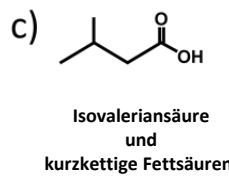

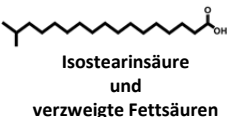
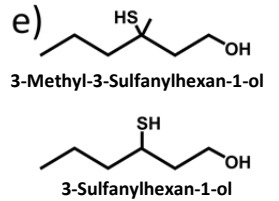
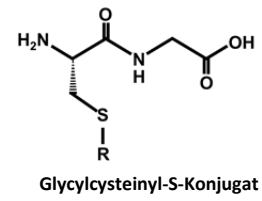
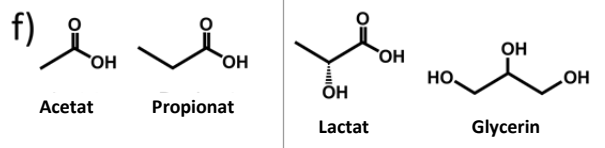
| Geruchskomponente  | Vorläufer-Molekül  | Enzym   | Organismus   |
|--|--|---|--|
| <p>a)</p>  <p>5,α-Androst-16-en-3α-ol</p> <p>5,α-Androst-16-en-3α-on</p>                  |  <p>Androst-5,16-dien-3α-ol</p> <p>Androst-14,6-dien-3-on</p>  <p>Androst-16-dien-3α-ol-β-Glucuronid</p> | <p>4,5-Reduktase<br/>5α-Reduktase</p> <p>5α(β)-Sterol-Dehydrogenase</p> <p>Steroid-4,5-Isomerase</p> <p>β-Glucuronidase</p> | <p><i>Corynebacterium</i> spp.<br/><i>Micrococcus</i> spp.</p> <p><i>Staphylococcus</i> spp.<br/><i>Propionibacterium</i> spp.</p> |
| <p>b)</p>  <p>3-Hydroxy-3-Methylhexansäure (HMHA)</p> <p>3-Methyl-2-Hexensäure (3M2H)</p> |  <p>N<sup>α</sup>-Glutaminyl-Konjugat</p>   | <p>N<sup>α</sup>-Acylglutamin-aminoacylase (N-AGA)</p>  | <p><i>Corynebacterium</i> spp.</p>   |
| <p>c)</p>  <p>Isovaleriansäure und kurzkettinge Fettsäuren</p>                          |  <p>Leucin und verzweigte Fettsäuren</p>  | <p>BCAA Aminotransferase<br/>BCAA Dehydrogenase</p>   | <p><i>Staphylococcus</i> spp.<br/><i>Corynebacterium</i> spp.</p>  |
| <p>d)</p> <p>kurzkettinge Fettsäuren (VFA)</p>   |  <p>Isostearinsäure und verzweigte Fettsäuren</p>   | <p>Enzyme der β-Oxidation (FadD, FadE, FadB, FadA)</p>  | <p><i>Corynebacterium</i> spp.</p>   |
| <p>e)</p>  <p>3-Methyl-3-Sulfanylhexan-1-ol</p> <p>3-Sulfanylhexan-1-ol</p>             |  <p>Glycylcysteinyl-S-Konjugat</p>  | <p>TpdA Dipeptidase<br/>AecD C-S Lyase</p>  | <p><i>Corynebacterium</i> spp.<br/><i>Staphylococcus</i> spp.</p>  |
| <p>f)</p>  <p>Acetat      Propionat</p> <p>Lactat      Glycerin</p>                     |  | <p>Enzyme alternativer Fermentationswege (LDH, AckA)</p>  | <p>fakultativ anaerobe<br/><i>Staphylococcus</i> spp.<br/>mikroaerophile<br/><i>Propionibacterium</i> spp.</p>                     |

Abbildung 1: Übersicht bekannter molekularer Mechanismen der Körpergeruchsentstehung durch die Aktivität der Achselhöhlen-Mikrobiota. Für die bekannten Geruchskomponenten Steroide (a), kurzkettinge Fettsäuren (b-d), Sulfanylalkane (e) und Acetat und Propionat (f) sind die jeweiligen Vorläufermoleküle, die Enzymfunktionen sowie die Organismen der Achselhöhlen-Mikrobiota dargestellt. Aus (Fredrich et al., 2013), modifiziert.

Probanden nachgewiesen (Gower et al., 1994). Erst einige Zeit später wurde die Synthese diesergeruchsbildenden Steroide aus den Vorläufermolekülen Androsta-5,16-dien-3 $\alpha$ -ol und Androst-14,6-dien-3-on durch Mitglieder der Achselhöhlen-Mikrobiota gezeigt (Decréau et al., 2003). Dabei setzten Corynebakterien diese Reaktion besonders effizient um (s. Abbildung 1 a). Austin und Ellis konnten ebenfalls zeigen, dass die geruchsbildenden 16-Androsten-Steroide von gemischten Kulturen von Corynebakterien synthetisiert werden konnten, was für eine kooperative Zusammenarbeit dieser Bakterien bei der Entstehung von Körpergeruch spricht (Austin und Ellis, 2003). Jedoch konnte keines der corynebakteriellen Isolate die vollständige Umwandlung durchführen. Weiterhin waren insgesamt nur sehr wenige Isolate aus einer geringen Anzahl an Probanden in der Lage, die Reaktionen durchzuführen. Trotz Bemühungen bei der Erstellung von Stoffwechselkarten bleibt die genaue Herkunft der Vorläufermoleküle sowie die genauen Stoffwechselwege ungeklärt.

Im Rahmen eines neuartigen Ansatzes stellten Starkenmann et al. die Hypothese auf, dass die geruchsbildenden Steroide in einer wasserlöslichen Form als Konjugate in Form von Sulfat- oder Glucuronid-Derivaten von den apokrinen Schweißdrüsen sekretiert werden (Starkenmann et al., 2013). Mithilfe moderner LC- und MS-Techniken konnte das  $\beta$ -Glucuronid von 5 $\alpha$ -Androst-16-n-3 $\alpha$ -ol in Schweißproben bestehend aus ekkrinen und apokrinen Sekreten nachgewiesen werden (s. Abbildung 1 a). Die gemessenen Konzentrationen betragen durchschnittlich 79 ng/ml in weiblichen und 241 ng/ml in männlichen Proben. In der selben Studie konnte während der Inkubation steriler Schweißproben mit einer Glucuronidase das entsprechende  $\alpha$ -Androstenol freigesetzt werden. Aufgrund der Tatsache, dass für einige Hautbakterien, wie *Staphylococcus haemolyticus*, *Staphylococcus warneri* und *Propionibacterium acnes*  $\beta$ -Glucuronidase-Aktivität beschrieben worden ist, stellen die Autoren die Hypothese auf, dass es sich bei dem Vorläufermolekül von  $\alpha$ -Androstenol um das entsprechende  $\beta$ -Glucuronid handelt.

### 3.2.2 Kurzkettige Fettsäuren

Die Fette der Haut sind in der Literatur als sehr divers beschrieben. Dabei handelt es sich im Wesentlichen um Cholesterin, Sterolester, Di- und Triglyceride, Fettsäuren, Wachsester und Squalen (Leyden et al., 1981). Besonders divers sind dabei freie, d. h. unveresterte Fettsäuren, die häufig ungewöhnlich lang, ungeradzahlig, ungesättigt und verzweigt sind (Nicolaidis, 1974). Auch flüchtige kurz- und mittelkettige Fettsäuren (VFAs) der Hautoberfläche, welche zum individuellen Körpergeruch beitragen, wurden sehr umfangreich beschrieben und weisen eine hohe Variabilität in ihrer Struktur auf (Gallagher et al., 2008). Die Herkunft dieser geruchsbildenden Komponenten ist jedoch nur teilweise aufgeklärt.



Im Rahmen einer organoleptischen Beurteilung chromatographischer Eluate von Achselschweiß-Extrakten und nachgelagerter Analyse dieser Eluate mittels GC-MS konnten Zeng et al. diverse geradkettige, ungesättigte und verzweigte Fettsäuren mit einer Kettenlänge von sechs bis elf Kohlenstoffatomen identifizieren, welche mit Körpergeruch korrelierten (Zeng et al., 1991). Dabei stellte (E)-3-Methyl-2-Hexensäure (3M2H) die dominanteste Geruchskomponente dar (Abbildung 1 b). Später führten Spielman et al. Studien durch, im Rahmen derer die Proteine aus apokrinen Sekreten aufgetrennt, isoliert, hydrolysiert und schließlich mittels GS/MS analysiert wurden (Spielman et al., 1995). Dabei konnte gezeigt werden, dass 3M2H aus zwei unterschiedlichen Proteinen mit apparenten Massen von 26 kDa und 46 kDa herausgelöst wurde. Es wurde postuliert, dass 3M2H von diesen beiden Proteinen, die ASOB1 und ASOB2 (*Apocrine Secretion Odor-Binding Protein 1 und 2*) genannt wurden, durch die apokrinen Drüsen zur Hautoberfläche transportiert wird. Später wurde in Sequenzanalysen gezeigt, dass es sich bei ASOB2 um Apolipoprotein D (ApoD), ein Lipocalin der Superfamilie der  $\alpha_2\mu$ -Mikroglobuline, handelt (Zeng et al., 1996). Weiterhin konnte in der gleichen Studie bei Hybridisierungsexperimenten die mRNA von ASOB2 spezifisch in apokrinem Drüsengewebe lokalisiert werden. Aufgrund dieser Erkenntnisse wurde zunächst vermutet, dass analog zu tierischen menschliche Geruchsstoffe als Signalmoleküle in spezifische Gewebe transportiert und dort ausgeschieden werden (Cavaggioni und Mucignat-Caretta, 2000; James et al., 2013).

Die überzeugendste und detaillierteste Beschreibung der Beteiligung kurzkettiger Fettsäuren an der Körpergeruchsentstehung lieferten Natsch et al. (Natsch et al., 2003). In NaOH-Hydrolysaten organischer Extrakte von Achselhöhlensekreten wurden 3M2H und die strukturell verwandte 3-Hydroxy-3-Methylhexansäure (HMHA) im Zusammenhang mit Körpergeruch identifiziert (s. Abbildung 1 b). In einem Fraktionierungsexperiment wurden anschließend die nicht-hydrolysierten N<sup>α</sup>-Glutaminkonjugate als Vorläufermoleküle beider Fettsäuren mittels LC-MS identifiziert. Anschließend wurden im Rahmen von Koinkubationen mit bakteriellen Achselhöhlenisolaten vier corynebakterielle Stämme identifiziert, die die geruchsbildenden Fettsäuren von den Vorläufermolekülen abspalten konnten. Im Rahmen von *in-vitro*-Experimenten mit Zelllysaten des Isolates *Corynebacterium striatum* Ax 20 wurde eine N<sup>α</sup>-Acylglutamin-Aminoacylase-Funktion (N-AGA, EC 3.5.1.14) abgeleitet. Die Substratspezifität lag dabei deutlich bei dem Glutaminrest des Vorläufers. Der Acylrest konnte substituiert werden. Nach massenspektrometrischen Proteinanalysen konnte unter Verwendung degenerierter Oligonukleotide ein hochkonserviertes Protein mit einer Länge von 399 Aminosäuren und einem Molekulargewicht von 43.365 Da heterolog exprimiert, aufgereinigt und kinetisch charakterisiert werden. Die höchste Affinität konnte dabei für das Substrat N<sup>α</sup>-Decanoylglutamin beschrieben werden. Die Affinität zu den physiologischen Substraten war

deutlich geringer. Später konnten von der selben Arbeitsgruppe insgesamt 27 unterschiedliche organische Säuren nach Inkubation von N-AGA mit Achselhöhlensekreten beschrieben werden, wobei HMHA und 3M2H die höchsten Konzentrationen aufwiesen (Natsch et al., 2006). Weiterhin wurden andere 3-Hydroxysäuren, methylierte ungesättigte Fettsäuren, Abbauprodukte von Aminosäuren sowie verschiedenste andere Säuren identifiziert.

In einem Versuch die beiden scheinbar widersprüchlichen Forschungsergebnisse von Spielman et al. und Zeng et al. einerseits und Natsch et al. andererseits zu vereinen führten Akiba et al. massenspektrometrische Analysen von ApoD durch (Akiba et al., 2011). Dabei konnte ein *Peak* identifiziert werden, der einer kovalenten Bindung von HMHA an den N-Terminus des prozessierten Apolipoproteins nach Abspaltung einer N-terminalen Signalsequenz entspricht. Ein entsprechender *Peak* für 3M2H und den unmodifizierten N-Terminus konnte nicht identifiziert werden. Eine weitere Untersuchung zeigte, dass es sich bei der N-terminalen Aminosäure des prozessierten Proteins um Glutamin handelt (James et al., 2013). Daher besteht die Möglichkeit, dass der Ursprung des Glutaminkonjugates in ApoD liegt. In einer Studie mit 49 Probanden über einen Zeitraum von drei Jahren detektierten Troccaz et al. das Vorläufermolekül N<sup>α</sup>-3-Hydroxy-3-Methylhexanoyl-L-Glutamin in Schweißproben in Konzentrationen zwischen 9 mg/l und 398 mg/l (Troccaz et al., 2009). Dieser Befund spricht gegen ApoD als Transportprotein, da ein Mengenverhältnis von 1:1 zwischen ApoD und 3M2H einer unphysiologischen ApoD-Konzentration im Bereich von einigen g/l entspräche (James et al., 2013). Somit bleibt die Herkunft der N<sup>α</sup>-Acylglutamin-Vorläufermoleküle ungeklärt.

In Experimenten analog zu denen von Natsch et al. konnten James et al. zehn unterschiedliche N<sup>α</sup>-Acylglutamin-Aminoacylase-positive Stämme aus Achselhöhlenproben identifizieren, die in der Lage waren, sowohl synthetisch hergestelltes N<sup>α</sup>-3-Methyl-2-Hexenyl-L-Glutamin als auch das künstliche Substrat Carbobenzyloxy-L-Glutamin zu spalten (James et al., 2013). Davon gehörten acht Stämme zur bakteriellen Gattung *Corynebacterium* und jeweils ein Stamm zu den Gattungen *Staphylococcus* und *Micrococcus*. Innerhalb der identifizierten Aminoacylase-positiven Stämme konnte lediglich ein weiterer corynebakterieller Stamm beschrieben werden, der eine zu dem von Natsch et al. beschriebenen Isolat vergleichbare Aktivität aufwies. Insgesamt konnte keine der gemessenen Aminoacylase-Aktivitäten eine Erklärung für die hohen Konzentrationen von 3M2H liefern, die mittels GC-MS von der selben Arbeitsgruppe gemessen wurden. Daher wird angenommen, dass die Bakterienspezies mit der höchsten Aminoacylase-Aktivität bisher nicht beschrieben wurden. Kürzlich demonstrierten Fujii et al. die effiziente Abspaltung von HMHA von dem entsprechenden Glutamin-Konjugat durch ein Achselhöhlenisolat (A20) der Gattung *Anaerococcus* (Fujii et al., 2014).

Eine alternative Hypothese zur Entstehung geruchsbildender Fettsäuren in der Achselhöhle wurde in der Vergangenheit ebenfalls von James et al. verfolgt (James et al., 2004a). Die Gruppe vermutete den Ursprung von VFAs im unvollständigen Abbau langer, ungeradzahliger, verzweigter Fettsäuren über den Stoffwechselweg der  $\beta$ -Oxidation (s. Abbildung 1 d). Der Ursprung dieser Fettsäuren wurde in Triglyceriden der Hautoberfläche vermutet, die durch bakterielle Lipase-Aktivität in der Achselhöhle freigesetzt werden. Im Rahmen von Kultivierungen Gram-positiver aerober Achselhöhlenisolate in Medien mit supplementierten Fettsäuren identifizierten sie eine Gruppe von Corynebakterien (*Corynebacteria A*), die die Fähigkeit besitzen, geradzahlige gesättigte und ungesättigte Fettsäuren, sowie Fettsäuren der Länge C15 abzubauen. Jedoch konnte gezeigt werden, dass diese Organismen nicht in der Lage sind, strukturell ungewöhnliche, d. h. verzweigte Fettsäuren, wie beispielsweise Isostearinsäure (16-Methyl-Heptadecanonsäure), Isopalmitinsäure (14-Methyl-Pentadecanonsäure) und 4-Methylcapronsäure (4-Methyl-Octanonsäure) vollständig abzubauen. Daher wurde angenommen, dass Lipid-auxothrophe („lipophile“) Corynebakterien über den Fettsäureabbau an der Entstehung von Körpergeruch beteiligt sind. Im Zuge von Genomsequenzierungen human-assoziiertes Corynebakterien konnten Gene des Fettsäureabbaus in unterschiedlichen corynebakteriellen Spezies beschrieben werden (Tauch et al., 2005). Besonders gut wurde dabei das Lipid-auxotrophe Achselhöhlenisolat *Corynebakterium jeikeium* K411 beschrieben, welches keine Gene der Fettsäure-Synthese besitzt, jedoch insgesamt über 31 Gene des Fettsäureabbaus verfügt (Tauch et al., 2005). Es wurde vermutet, dass die relativ hohe Anzahl paraloger Gene des Fettsäureabbaus mit unterschiedlichen Fettsäurespezifitäten der einzelnen Genprodukte zusammenhängt (Barzantny et al., 2012a). Im Rahmen eines bioinformatischen Ansatzes wurde das transkriptionelle Netzwerk von *C. jeikeium* K411 im Zusammenhang mit Fettsäureabbau beschrieben (Barzantny et al., 2012b). Anschließend wurden im Rahmen von Transkriptionsstudien Transkriptmengen von Fettsäureabbaugenen beschrieben und Mithilfe von DNA-Affinitätsstudien und EMSA der MarR-ähnliche transkriptionelle Regulator der Fettsäureabbaugene FamR identifiziert (Barzantny et al., 2013). Ein direkter Zusammenhang zwischen dem Fettsäureabbau durch *C. jeikeium* K411 und Körpergeruchsentstehung wurde experimentell jedoch nicht gezeigt. Aufgrund der mangelnden Evidenz für die Hypothese des unvollständigen Fettsäureabbaus durch Lipid-auxothrophe Corynebakterien sowie auf Basis unveröffentlichter Studien und einer Reinterpretation der Fachliteratur zogen James et al. ihre anfängliche Vermutung zurück (James et al., 2013). Derzeit ist die Gruppe der Ansicht, dass die Entstehung kurzkettiger Fettsäuren in der Achselhöhle auf den Abbau aliphatischer Aminosäuren durch Staphylokokken zurückzuführen ist und dass die Umwandlung von L-Leucin zu Isovaleriansäure (3-Methylbutansäure) die Hauptroute darstellt (s. Abbildung 1 c). Bereits im

Jahre 2004 veröffentlichte diese Gruppe Ergebnisse aus Biotransformationsstudien, in denen sie die Freisetzung von VFAs durch Propionibakterien und Staphylokokken beschrieben (James et al., 2004b). Dabei konnte gezeigt werden, dass Propionibakterien und Staphylokokken in der Lage sind, VFAs durch die Fermentation von Laktat und Glycerin herzustellen. Dahingegen waren einige *Staphylococcus*-Isolate in der Lage, die Aminosäuren Valin, Leucin und Isoleucin zu Isobuttersäure (2-Methylpropansäure), Isovaleriansäure und 2-Methylbuttersäure umzuwandeln. Zum damaligen Zeitpunkt wurde die Beteiligung dieser Stoffwechselwege an der Körpergeruchsentstehung jedoch als gering eingeschätzt. Ihre jetzige Position begründet die Gruppe damit, dass zum einen der Anteil an Corynebakterien der Gruppe A in Isolaten der Achselhöhle sehr gering sei (< 10 %) und zum anderen der Anteil Iso-Methyl-verzweigter Fettsäuren in Talg sehr gering sei (James et al., 2013). Aktuelle experimentelle Daten zur Umwandlung von Leucin in Isovaleriansäure wurden jedoch bisher nicht publiziert.

### 3.2.3 Sulfanylalkanole

Sulfanylalkanole sind als Geruchskomponenten insbesondere im Zusammenhang mit Weinen der Sorte *Sauvignon blanc* beschrieben (Styger et al., 2011; Swiegers und Pretorius, 2007). Als Geruchskomponenten der menschlichen Achselhöhle wurden sie jedoch erst im Laufe der vergangenen 11 Jahre beschrieben. Hasegawa et al. identifizierten 3-Sulfanylalkan-1-ole im Achselweiß von Männern (Hasegawa et al., 2004). Dabei wurden besonders hohe Mengen an (S)-3-Methyl-3-Sulfanylhexan-1-ol gemessen, welches einen typischen Schweißgeruch mit einer fleischigen und fruchtigen Note aufweist. In darauffolgenden Untersuchungen wurden durch andere Arbeitsgruppen weitere geruchsbildende Sulfanylalkanole mit geringen Wahrnehmungsschwellen im Bereich von wenigen pg/l Luft beschrieben (James et al., 2013). Aufgrund der geringen Wahrnehmungsschwellen und der im Vergleich zu geruchsbildenden Fettsäuren geringeren Konzentrationen wird die Beteiligung von Sulfanylalkanolen an der Körpergeruchsentstehung als sehr hoch eingeschätzt. Die weiterführende Untersuchung von geruchsbildenden Sulfanylalkanolen der Achselhöhle wurde im Wesentlichen von drei Arbeitsgruppen vorangetrieben, die unterschiedliche Vermutungen bezüglich ihrer Herkunft verfolgten.

Troccaz et al. demonstrierten in Biotransformations-Assays und *Sniff-Port*-Analysen die Entstehung von typischem schwefeligen Schweißgeruch nach Inkubation von frischem Schweiß mit dem Achselhöhlenisolat *Staphylococcus haemolyticus* Ax20 (Troccaz et al., 2004). Eine darauffolgende Analyse mittels GC-MS ermöglichte die Identifizierung von 3-Methyl-3-Sulfanylhexan-1-ol, welches in einer typischen Konzentration von 15 nM im Schweiß vorlag. Dabei wies das (S)-Enantiomer den typischen Geruch auf, der mit den Deskriptoren Schweiß,

Zwiebel und Tier in Verbindung gebracht wird. Die (R)-Form hingegen wurde als fruchtig, schwefelig und grapefruitartig beschrieben. Wenig später beschrieb die Arbeitsgruppe darüber hinaus die Schwefelverbindungen 2-Methyl-3-Sulfanylpentan-1-ol und 3-Sulfanylhexan-1-ol nach Biotransformation von Schweiß durch *S. haemolyticus* (Starkenmann et al., 2005). Darüber hinaus identifizierten sie das Cysteinyl-Glycyl-(S)-Konjugat von 3-Methyl-3-Sulfanylhexan-1-ol als Vorläufermolekül in frischem Schweiß (s. Abbildung 1 e). Auf Grundlage der Annahme, dass dieses Vorläufermolekül aus dem Glutathion-Stoffwechsel stammt, wurde das entsprechende Glutathion-Konjugat synthetisiert. Dieses konnte jedoch weder in Biotransformationsstudien umgesetzt, noch in frischem Schweiß detektiert werden. Das Cystein-(S)-Konjugat konnte hingegen in geringen Mengen und das Cysteinyl-Glycyl-(S)-Konjugat in höheren Mengen umgesetzt werden. Analog zu dem molekularen Mechanismus in *Sauvignon blanc* postulierte die Arbeitsgruppe die aufeinanderfolgenden Abbauschritte des Glutathion-Konjugates durch eine  $\gamma$ -Glutamyl-Transpeptidase, eine Carboxypeptidase und eine C-S- $\beta$ -Lyase zur Abspaltung des geruchsbildenden Sulfanylalkanols vom Cysteinrest. Da das Glutathion-Konjugat von keinem der identifizierten Achselhöhlenisolate umgesetzt wurde, folgerten die Autoren, dass die  $\gamma$ -Glutamyl-Transpeptidase-Reaktion zu einem früheren Zeitpunkt an einem anderen Ort stattfindet. Später charakterisierten sie die heterolog exprimierte MetC Cystathionin- $\beta$ -Lyase des Achselhöhlenisolates *S. haemolyticus* Ax3 (Troccaz et al., 2008). Das rekombinante Protein zeigte durch die Spaltung von Cystathionin zu Homocystein und die Umsetzung von Methionin zu Methanliol zwar die Enzymfunktionen der Cystathionin- $\beta$ -Lyase, konnte jedoch bei der Inkubation mit frischem Schweiß keinen Geruch generieren.

Parallel zu den Arbeiten von Troccaz et al. identifizierten Natsch et al. ebenfalls mittels GC-MS- und *Sniff-Port*-Analysen die geruchsbildenden Sulfanylalkanole 3-Methyl-3-Sulfanylhexan-1-ol, 3-Sulfanylpentan-1-ol, 2-Methyl-3-Sulfanylbutan-1-ol und 3-Sulfanylhexan-1-ol nach Hydrolyse von frischem Schweiß mit NaOH (Natsch et al., 2004). Synthetische Cystein-Konjugate dieser Geruchsträger konnten in Biotransformationsstudien ausschließlich durch corynebakterielle Achselhöhlenisolate gespalten werden. Mithilfe einer *Shotgun*-Klonierung einer chromosomalen Bibliothek von *C. striatum* Ax20 konnte ein Cystathionin- $\beta$ -Lyase-Gen identifiziert werden, das nur sehr geringe Sequenzidentität zu anderen corynebakteriellen MetC-Genen aufwies. Nach heterologer Expression in *E. coli* konnten S-substituierte Cystein-Konjugate sowohl durch die Zellextrakte als auch durch das aufgereinigte Enzym gespalten werden. Ähnlich zu dem Enzym N-AGA aus dem gleichen Organismus konnte auch für die  $\beta$ -Lyase lediglich eine geringe Umsatzrate mit einem  $K_M$ -Wert im Bereich von 1,8 mM – 2,4 mM beschrieben werden. In einer späteren Arbeit bestätigte die Arbeitsgruppe das Glycyl-Cysteinyl-S-Konjugat von 3-Methyl-3-Sulfanylhexan-1-ol als Vorläufer von 3-Methyl-3-Sulfanylhexan-1-ol in frischem

Schweiß mit einer höheren Konzentration als das entsprechende Cystein-Konjugat und die Umsetzung durch *Corynebacterium striatum* Ax20 (Emter und Natsch, 2008). Eine darauffolgende Aufreinigung, Beschreibung, Klonierung und heterologe Expression in *E. coli* zeigte, dass es sich bei dem Protein um eine neuartige Dipeptidase (*thiol precursor dipeptidase* A, TpdA) handelt, die eine hohe Affinität zu Glycyl-Cysteinyl-3-Methyl-3-Sulfanylhexan-1-ol aufweist. In Koinkubations-Experimenten mit TpdA und der Cystathionin- $\beta$ -Lyase aus *C. striatum* Ax20 konnte die Freisetzung von 3-Methyl-3-Sulfanylhexan-1-ol sowohl aus dem synthetischen Gly-Cys-Konjugat als auch aus frischem Schweiß demonstriert werden. Diese Beobachtungen ließen schließen, dass die Freisetzung des geruchsbildenden Sulfanylalkanols aus dem Vorläufer durch die aufeinander folgende Aktivität dieser beiden Enzyme katalysiert wird (s. Abbildung 1 e).

Auf Grundlage unveröffentlichter Studiendaten bestätigten James et al. später einige der Beobachtungen zu der Freisetzung geruchsbildender Sulfanylalkanole (James et al., 2013). Sie identifizierten 3-Sulfanylhexan-1-ol und 3-Methyl-3-Sulfanylbutan-1-ol in Achselhöhlen-Proben sowie den Cystein-konjugierten Vorläufer von 2-Methyl-3-Sulfanylbutan-1-ol. Die Spaltung eines synthetischen Vorläufermoleküls konnte in Biotransformationsstudien durch Achselhöhlenisolate der Gattungen *Corynebacterium* und *Staphylococcus* demonstriert werden, wobei erstere häufiger die C-S-Lyase-Aktivität aufwiesen. Die höchste Aktivität konnte für *C. jeikeium* NCIMB 40918 beschrieben werden. Auf Grundlage bioinformatischer Analysen wurden die beiden Proteingruppen der MetC- und MalY-C-S- $\beta$ -Lyasen als wahrscheinliche Kandidaten für die Enzymaktivität in diesem Organismus identifiziert. Neben den MetC- Cystathionin- $\beta$ -Lyasen sind in *E. coli* die bifunktionalen MalY-Typen bekannt, welche sowohl eine Cystathionin- $\beta$ -Lyase-Aktivität besitzen als auch an der Aufnahme von Maltose beteiligt sind (Zdych et al., 1995). James et al. postulierten, dass Cystathionin-  $\beta$ -Lyasen aus Staphylokokken vom MetC-Typ sind, während die corynebakteriellen Gene vom MalY-Typ sind, was auch auf das von Natsch et al. beschriebene Protein aus *C. striatum* Ax20 zutrifft. Die Arbeitsgruppe um James et al. klonierte daraufhin das Gen aus *C. jeikeium* NCIMB 40918 und demonstrierte die Spaltung eines Cystein-konjugierten Vorläufers durch das rekombinante Protein (James et al., 2013). Es zeigte sich, dass die Sequenz dieses Proteins identisch mit der Sequenz der Cystathionin- $\beta$ -Lyase aus dem Achselhöhlen-Isolat *C. jeikeium* K411 war. In einer weiterführenden Studie wurde McbR als transkriptioneller Regulator des MalY-Gens und anderer an der Synthese schwefelhaltiger Aminosäuren beteiligter Gene beschrieben (Brune et al., 2011).

Trotz der z. T. sehr detailliert beschriebenen Enzymfunktionen ist nicht abschließend geklärt, ob es sich bei den Vorläufermolekülen geruchsbildender Sulfanylalkanole ausschließlich um

Cysteinyl-Glycyl-Konjugate oder zusätzlich um Cystein-Konjugate handelt (James et al., 2013). Auch die physiologische Bedeutung dieser enzymatischen Umsetzungen *in vivo* sowie die zugehörigen Stoffwechselwege sind nicht bekannt.

### **3.2.4 Beteiligung menschlicher Komponenten an der Körpergeruchsentstehung**

In den oben beschriebenen Studien konnte vielfach gezeigt werden, dass die geruchslosen Vorläufer-Moleküle geruchsbildender Komponenten bereits in frischem Schweiß vorhanden sind. Mit der Beschreibung der Rolle von ApoD beim Transport von Fettsäure-Vorläufern in der Achselhöhle wurde erstmals eine humane molekulare Komponente mit Beteiligung an der Körpergeruchsentstehung beschrieben. Weiterhin kann aufgrund der hohen Diversität geruchsbildender molekularer Komponenten in menschlichem Schweiß und der interindividuellen Unterschiede der Geruchszusammensetzung die Beteiligung menschlicher genetischer Faktoren vermutet werden.

Unterschiede in der Konzentration des Fettsäure-Vorläufermoleküles von (R)/(S)-HMHA und des Schwefel-Vorläufers von (R)/(S)-3-Methyl-3-Sulfanylhexan-1-ol in männlichen und weiblichen Schweißproben konnten an Proben von 49 Probanden über zwei Jahre gezeigt werden (Troccaz et al., 2009). Dabei war das Verhältnis zwischen dem Fettsäure-Vorläufer (mit einem käsigen ranzigen Geruch) und dem Schwefel-Vorläufer (mit einem fruchtigen oder zwiebeligen Geruch) in männlichen Proben durchschnittlich um einen Faktor drei höher als in weiblichen Proben, ohne dass eine Korrelation zum Schweißvolumen oder zur Proteinkonzentration im Schweiß vorlag. Frauen haben somit das Potential, mehr schwefelige Geruchskomponenten freizusetzen. Diese Beobachtung konnte auch in einer späteren Studie reproduziert werden (Troccaz et al., 2015).

Bei Mäusen ist ein Paarungsverhalten beschrieben, das aufgrund der Erkennung von Geruchskomponenten in Urin Fortpflanzungspartner mit unterschiedlichen MHC-Allelen begünstigt und wahrscheinlich der Vermeidung von Inzucht und der Diversifizierung der Nachkommenschaft dient (Yamaguchi et al., 1981). Ein analoger Verhaltensmechanismus, der auf HLA-Allelen basiert, wurde ebenfalls beim Menschen untersucht und diverse Studien legen sowohl das Vorhandensein als auch die Abwesenheit von Zusammenhängen zwischen dem HLA-Genotyp und dem Verhalten bei der Partnerwahl basierend auf der Wahrnehmung von Körpergerüchen nahe (Jacob et al., 2002; Roberts et al., 2008; Wedekind et al., 1995). Um die Verbindung zwischen den biochemischen Eigenschaften menschlichen Körpergeruchs und ihrer genetischen Individualität basierend auf Wahrnehmungsstudien herzustellen, untersuchten Natsch et al. die Zusammensetzung der Geruchskomponenten von 12 Zwillingspaaren (Kuhn und Natsch, 2009). Nach Behandlung der Schweißproben mit dem Enzym N-AGA und Analyse

mittels GC-MS zeigten sich die geringsten Differenzen auf der intraindividuellen Ebene bei Betrachtung von Proben eines Probanden am selben Tag oder an unterschiedlichen Tagen. Die deutlich höhere Ähnlichkeit zwischen den Proben eineiiger Zwillinge im Vergleich zu Proben nicht verwandter Probanden demonstriert eine genetisch bestimmte individuelle Zusammensetzung von geruchsbildenden Carboxylsäuren im menschlichen Schweiß. In einer späteren Studie untersuchte die Arbeitsgruppe den Zusammenhang zwischen geruchsbildenden Carboxylsäuren und HLA-Haplotypen in 12 Familien (Natsch et al., 2010). Auch in dieser Studie konnte die höchste Ähnlichkeit zwischen Proben des selben Individuums beschrieben werden. Proben von Geschwistern wiesen zwar auch in dieser Studie eine höhere Ähnlichkeit auf als Proben nicht verwandter Individuen, jedoch konnte kein Zusammenhang zwischen den molekularen Profilen und den entsprechenden HLA-Haplotypen gezeigt werden.

Ein weiterer Zusammenhang zwischen der molekularen Zusammensetzung geruchsbildender Schweißkomponenten und der menschlichen Genetik kann durch das *ABCC11*-Gen beschrieben werden. Das Gen kodiert einen apikalen Membrantransporter der Genfamilie der *MRP8* (*Multidrug Resistance-Related Protein 8*) der an der Ausbildung von Körpergeruch beteiligt ist (Martin et al., 2010). Zuvor konnte durch eine andere Arbeitsgruppe demonstriert werden, dass ein *SNP* (*Single Nucleotide Polymorphism*), der zu einem Austausch des Glycinrestes an Position 538 zu einem Alaninrest führt (538G->A), im Zusammenhang mit dem Auftreten des trockenen Cerumen (Ohrenschmalz)-Typen steht (Yoshiura et al., 2006). Das entsprechende A-Allel kommt in der asiatischen Bevölkerung deutlich häufiger vor als in der kaukasischen oder afrikanischen. Weiterhin wird häufig berichtet, dass starker Körpergeruch unter der afrikanischen und kaukasischen Bevölkerung weiter verbreitet ist als unter der asiatischen Bevölkerung (Martin et al., 2010). Diese Beobachtung ist jedoch nicht hinreichend wissenschaftlich belegt (James et al., 2013). Die Arbeitsgruppe um Martin et al. demonstrierte, dass homozygote Träger des A-Alleles nahezu keine aminosäurekonjugierten Vorläufermoleküle besitzen. Im Gegensatz dazu sind die Mengen dieser Vorläufer in heterozygoten und für das G-Allel homozygoten Individuen deutlich höher. Darüber hinaus zeigte die Arbeitsgruppe die Expression und Lokalisation des Membrantransporters in apokrinen Schweißdrüsen. Daher wurde eine entscheidende Rolle des Transporters bei der Sekretion geruchsbildender Vorläufer-Moleküle vermutet. In einer späteren Arbeit konnte am heterolog exprimierten *ABCC11*-Transporter in *Sf9*-Zellen der Transport des Glutathionyl-Konjugates von 3-Methyl-3-Sulfanylhexanol (3H3SH) gezeigt werden (Baumann et al., 2014). Die Rolle des Enzyms  $\gamma$ -Glutamyltransferase 1 in der Generierung von Cysteinyl-Glycyl-Konjugaten von Sulfanylalkanolen durch Abspaltung des Glutamatrestes wurde ebenfalls in dieser Arbeit postuliert. Diese Arbeit könnte eine Verbindung zu den Arbeiten von



Starkenmann et al. liefern und den fehlenden enzymatischen Schritt im Abbau von Glutathionyl-S-Konjugaten belegen (s. Kapitel 3.2.3).

Eine weitere Hypothese vermutet den Ursprung geruchsbildender Komponenten in der Achselhöhle in Carotinoiden, welche über bestimmte Nahrungsmittel, wie Tomaten, aufgenommen werden (Stewart, 2014). Diese werden durch das Enzym  $\beta$ -Carotin-9,10-Monooxygenase gespalten und von den apokrinen Drüsen sekretiert. Experimentelle Belege für diese Hypothese fehlen jedoch.

### **3.3 Kultivierungsunabhängige Beschreibungen der Achselhöhlen-Mikrobiota**

Kultivierungsunabhängige Verfahren zur Beschreibung der Achselhöhlen-Mikrobiota existieren seit Ende der 2000er Jahre und ermöglichten eine sehr detaillierte Beschreibung der Zusammensetzung der Achselhöhlen-Mikrobiota (s. Tabelle 1). Aufgrund ihrer herausragenden Eigenschaften als Biomarker und aufgrund der guten Amplifizierbarkeit handelte es sich bei dem untersuchten Molekül nahezu immer um 16S-rDNA. Methodisch kamen dabei bisher spezielle PCR-, Elektrophorese- sowie Sequenzieretechniken zum Einsatz. Letztere ermöglichen einen besonders guten Einblick in die taxonomischen Profile bakterieller Lebensgemeinschaften, da auch unbekannte Sequenzen analysiert werden können und keine interne Referenz für die Interpretation der Daten notwendig ist. Aber auch die anderen beiden Techniken ermöglichen qualitative und z. T. quantitative Aussagen über die Mitglieder der Achselhöhlen-Mikrobiota.

Gao et al. wählten einen qPCR-Ansatz zur quantitativen Beschreibung der Bakterienzahlen der Genera *Corynebacterium*, *Propionibacterium*, *Staphylococcus* und *Streptococcus* und der Pilzgattung *Malassezia* (Gao et al., 2010). Zur Untersuchung von Proben von fünf männlichen und weiblichen Probanden wurden gattungsspezifische *Primer* und *TaqMan MGB* Sonden zur Detektion amplifizierter 16S-rDNA- und ITS-Sequenzen verwendet. Es konnte gezeigt werden, dass die vier Bakteriengattungen in vier von fünf Proben die Mehrheit der bakteriellen Lebensgemeinschaften bildeten. Der Gattung *Malassezia* konnten 53 % - 80 % aller ITS-Amplicons zugeordnet werden, wobei die Anzahlen um einen Faktor 10 bis 100 geringer waren als die der bakteriellen Sequenzen.

Die Arbeitsgruppe um Callewaert et al. verwendete die Methode der denaturierenden Gradienten-Gelelektrophorese (DGGE) und die Pyrosequenzierung zur Beschreibung der bakteriellen Achselhöhlen-Mikrobiota (Callewaert et al., 2013, 2014). Bei den dominanten Bakterienstämmen in Proben von 53 männlichen und weiblichen Probanden handelte es sich um *Actinobacteria*, *Firmicutes* und *Proteobacteria* (Callewaert et al., 2013). Dies wurde ebenfalls

durch Pyrosequenzierung bestätigt. Im Einklang mit den Erkenntnissen aus den kultivierungsbasierten Beschreibungen bildeten die Profile der jeweiligen Probanden zwei distinkte Cluster und wurden entweder von *Staphylococcus*-Spezies (geringer GC-Gehalt) oder *Corynebacterium*-Spezies (hoher GC-Gehalt) dominiert. Dabei konnte mit 61 % die Mehrheit der Probanden dem *Staphylococcus*-Cluster zugeordnet werden und 39 % dem *Corynebacterium*-Cluster. Auch clusternten 87 % der weiblichen Probanden im *Staphylococcus*-Cluster während das *Corynebacterium*-Cluster überwiegend durch männliche Probanden beschrieben wurde. Das *Staphylococcus*-Cluster wurde auf Spezies-Ebene dominiert von *S. epidermidis*, *Staphylococcus* spp. und *S. hominis*. Das *Corynebacterium*-Cluster wird auf Spezies-Ebene überwiegend durch *Corynebacterium* spp., Proteobakterien und *Staphylococcus hominis* beschrieben.

Die erste Beschreibung der Achselhöhlen-Mikrobiota unter Verwendung der Sanger-Sequenzieretechnik lieferten Grice et al. (Grice et al., 2009). Die generierten Amplicons aus Proben von neun männlichen und weiblichen Probanden deckten nahezu das gesamte rDNA-Gen ab und wurden vor der Sequenzierung in einen Vektor kloniert. Die beschriebenen taxonomischen Profile der Achselhöhlen-Mikrobiota können drei unterschiedlichen Typen zugeordnet werden: Vier Lebensgemeinschaften wurden von Corynebakterien, drei von Staphylokokken und zwei von Betaproteobakterien dominiert. Darüber hinaus wurden die Taxa *Clostridiales*, *Lactobacillales*, *Propionibacterium* und *Streptococcus* beschrieben.

Nur wenig später veröffentlichten Costello et al. eine Beschreibung des taxonomischen Profils der Achselhöhlen-Mikrobiota unter direkter Anwendung der Pyrosequenzierung ohne Zwischenklonierung (Costello et al., 2009). Proben der Achselhöhlen-Mikrobiota mehrerer Probanden wurden gepoolt und mithilfe spezifischer Primer die V2-Region des 16S rDNA-Gens amplifiziert. Das beschriebene taxonomische Profil wurde von Staphylokokken dominiert. Weiterhin wurden unter Anderem Corynebakterien, Propionibakterien sowie Streptophyta in der Achselhöhlen-Mikrobiota beschrieben, wobei Letztere als Kontamination verursacht durch Pollen interpretiert wurden.

Troccaz et al. nutzten die Pyrosequenzierung zur Beschreibung der Achselhöhlen-Mikrobiota von elf Probanden Mithilfe von 16S-rDNA-Amplicons der Regionen V1-V3 (Troccaz et al., 2015). Das taxonomische Profil wurde nahezu vollständig (99,8 % der Sequenzen) durch die Bakterienstämme *Firmicutes*, *Actinobacteria*, *Proteobacteria* und *Bacteroidetes* beschrieben. Dabei wurde im Einklang mit den Erkenntnissen aus kultivierungsbasierten Untersuchungen eine sehr starke negative Korrelation zwischen *Firmicutes* und *Actinobacteria* beobachtet. Insgesamt wurden 68 Genera beschrieben, wobei *Staphylococcus*, *Propionibacterium* und

**Tabelle 1: Übersicht der kultivierungsunabhängigen Studien zur Zusammensetzung der Achselhöhlen-Mikrobiota. Für jede Literaturquelle ist eine Kurzbeschreibung der Studie und der verwendeten Methode sowie der identifizierten bakteriellen Taxa auf Genus- und auf Speziesebene dargestellt.**

| Literaturverweis          | Studie  | Methoden  | Dominante Bakterientaxa   |
|---------------------------|---|---|---|
| (Grice et al., 2009)      | Taxonomische Beschreibung von Lebensgemeinschaften der menschlichen Haut (9 männliche und weibliche Probanden)  | 16S-rDNA-Sequenzierung (Sanger), ~ 1383 bp, V1-V8, mit Zwischenklonierung   | <i>Staphylococcus</i><br><i>Corynebacterium</i><br><i>Betaproteobakterien</i>   |
| (Costello et al., 2009)   | Taxonomische Beschreibung der menschlichen Mikrobiota an unterschiedlichen Körperstellen (7-9 männliche und weibliche Probanden)                                      | 16S-rDNA-Sequenzierung (Pyrosequenzierung), ~ 311 bp, V2,   | <i>Staphylococcus</i><br><i>Corynebacterium</i><br><i>Propionibacterium</i>   |
| (Gao et al., 2010)        | Quantitative Beschreibung der Haut-Mikrobiota an unterschiedlichen Körperstellen (5 männliche und weibliche Probanden)  | qPCR von 16S-rDNA-Genen (Primer spezifisch für definierte Genera)   | <i>Corynebacterium</i><br><i>Streptococcus</i><br><i>Staphylococcus</i><br><i>Propionibacterium</i>                               |
| (Egert et al., 2011)      | Taxonomische Beschreibung der gesamten Achselhöhlen-Mikrobiota und der aktiven Achselhöhlen-Mikrobiota (10 männliche Probanden)                                       | T-RFLP <i>Fingerprinting</i> (10 Probanden) und Sequenzierung von 16S-rDNA- und revers transkribierten 16S-rRNA-Amplicons (3 Probanden) (Sanger); ~ 880 bp, V1 - V5 | <i>Corynebacterium</i><br><i>Staphylococcus</i><br><i>Propionibacterium</i><br><i>Anaerococcus</i><br><i>Peptoniphilus</i>        |
| (Callewaert et al., 2013) | Beschreibung von <i>Staphylococcus</i> - und <i>Corynebacterium</i> -Clustern der Achselhöhlen-Mikrobiota (53 männliche und weibliche Probanden)                      | DGGE (53 Probanden) und Sequenzierung (9 Proben) von 16S-rDNA-Amplicons (Pyrosequenzierung)   | <i>S. epidermidis</i><br><i>S. hominis</i><br><i>Staphylococcus spp.</i><br><i>Corynebacterium spp.</i><br><i>Proteobacterium</i> |
| (Callewaert et al., 2014) | Beschreibung der Veränderung der Achselhöhlen-Mikrobiota im Zusammenhang mit Deo-Anwendung über einen Zeitraum von zwei Monaten (9 männliche und weibliche Probanden) | DGGE von 16S-rDNA-Amplicons   | <i>Staphylococcus</i><br><i>Corynebacterium</i>   |
| (Troccaz et al., 2015)    | Taxonomische Beschreibung der Achselhöhlen-Mikrobiota und Korrelation mit olfaktorischen Kennwerten (24 männliche und weibliche Probanden)                            | 16S-rDNA-Amplicon-Sequenzierung (11 Probanden) (Pyrosequenzierung), V1-V3   | <i>Staphylococcus</i><br><i>Corynebacterium</i><br><i>Propionibacterium</i>   |
| (Urban et al., 2016)      | Änderungen in der Zusammensetzung der Achselhöhlen-Mikrobiota im Zusammenhang mit der Anwendung von Deos und APs (18 männliche und weibliche Probanden)               | 16S-rDNA-Amplicon-Sequenzierung (Pyrosequenzierung); ~ 300 bp; V4   | <i>Staphylococcaceae</i><br><i>Corynebacterium</i><br><i>Fingoldia</i><br><i>Anaerococcus</i>                                     |

*Corynebacterium* mit insgesamt 96 % der Sequenzen die Datensätze dominierten. Dabei korrelierten corynebakterielle OTUs stark positiv miteinander und der Genus *Propionibacterium* korrelierte negativ mit *Staphylococcus* und *Corynebacterium*. Auf Spezies-Ebene konnten Sequenzhomologien zu *S. epidermidis*, *C. tuberculostearicum*, *S. hominis* und *P. acnes* in allen untersuchten Proben mit einem Anteil von 66 % bis 99 % der Sequenzen identifiziert werden. Daher wurden diese vier Spezies als universeller Kern der Achselhöhlen-Mikrobiota bezeichnet. Darüber hinaus wurden im Rahmen dieser Studie einzelne OTUs mit Körpergeruch oder einzelnen Deskriptoren des Körpergeruches korreliert. Insgesamt konnte eine positive Korrelation zwischen dem Körpergeruch und OTUs des Genus *Corynebacterium*, insbesondere *C. tuberculostearicum*, und der Spezies *S. hominis* beobachtet werden. OTUs von *S. hominis*, *C. tuberculostearicum*, *Corynebacterium* und *Anaerococcus* korrelierten dabei positiv mit mindestens einem Deskriptor. Diese Studie beschreibt erstmals die Korrelation von 16S-rDNA-Sequenzdaten mit Körpergeruch.

Da 16S-rDNA-Amplicons die vorhandene Mikrobiota beschreiben, jedoch keine Unterscheidung zwischen aktiven und weniger aktiven Mitgliedern zulassen, untersuchten Egert et al. die Achselhöhlen-Mikrobiota männlicher Probanden auf der Ebene der 16S-rDNA und 16S-rRNA (Egert et al., 2011). Mithilfe von T-RFLP-Analysen (*terminal restriction length polymorphism*) von 16S-rDNA-Amplicons und revers transkribierten 16S-rRNA-Amplicons wurden die taxonomischen Profile verglichen um die aktiven Mitglieder der Achselhöhlen-Mikrobiota identifiziert. Signifikante Ähnlichkeiten beider Profile ließen die Schlussfolgerung zu, dass die dominanten Mitglieder der Achselhöhlengemeinschaft aktiv sind. Die Klonierung und Sequenzierung von je 48 rDNA- und rRNA-Amplicons von drei Probanden zeigten, dass die bakterielle Lebensgemeinschaft der Achselhöhle im Wesentlichen aus den Genera *Corynebacterium*, *Propionibacterium*, *Staphylococcus*, *Anaerococcus* und *Peptoniphilus* besteht, wobei letzterer mit der Händigkeit der Probanden in Zusammenhang stehen könnte.

*Peptoniphilus* Spezies gehören ebenso wie *Anaerococcus*, *Fingoldia* und andere zu der Gruppe der Gram-positiven anaeroben Cocci (GPAC), welche als Kommensalen der menschlichen Mikrobiota beschrieben und häufig aus klinischen Proben polymikrobieller Infektionen isoliert werden (Murphy und Frick, 2013). Die Identifizierung und taxonomische Zuordnung dieser Bakterien war in der Vergangenheit stets limitiert durch ihre schlechte Kultivierbarkeit unter Laborbedingungen. Die Anwendung kultivierungsunabhängiger molekularer Techniken ermöglicht die Beschreibung dieser Organismen innerhalb der Achselhöhlen-Mikrobiota und die Untersuchung ihrer Rolle im Zusammenhang mit Körpergeruchsentstehung. Auch Vertreter der Archaeobakterien wurden Mithilfe der 16S-rDNA-Ampliconsequenzierung als Mitglieder der

Haut-Mikrobiota beschrieben (Probst et al., 2013). Insbesondere Mitglieder des Taxons *Thaumarchaeota*, die über Gene der Ammonium-Monooxygenase als weiteren Biomarker verfügen, wurden sowohl auf der Haut als auch in Reinräumen und Krankenhäusern nachgewiesen. Die potentielle metabolische Rolle der Ammoniumoxidation auf der Haut ist jedoch nicht abschließend geklärt.

Insgesamt stimmen die Erkenntnisse zur Zusammensetzung der Achselhöhlen-Mikrobiota aus kultivierungsunabhängigen Beschreibungen sowohl untereinander als auch mit denen aus den klassischen kultivierungsabhängigen Studien (s. Kapitel 3.1) überein. Dies betrifft insbesondere das Vorhandensein der Genera *Staphylococcus*, *Corynebacterium*, *Proionibacterium* und *Anaerococcus* als auch die Dominanz entweder durch Corynebakterien oder durch Staphylokokken. Darüber hinaus wurden in einigen kultivierungsunabhängigen Studien vergleichende Beschreibungen der Variabilität der Achselhöhlen-Mikrobiota durchgeführt. Grundsätzlich wurde die interindividuelle Variabilität in allen vorgestellten Studien als höher beschrieben als die intraindividuelle, was weiterhin auf eine personenspezifische Zusammensetzung der Lebensgemeinschaft hindeutet. Jedoch wurde auch eine Asymmetrie zwischen linker und rechter Achselhöhle einzelner Probanden beschrieben, die auf eine hohe Ortsspezifität der Diversität schließen lässt (Callewaert et al., 2013; Egert et al., 2011; Gao et al., 2010). Zeitliche Variabilitäten der Achselhöhlen-Mikrobiota sind ebenfalls vereinzelt untersucht worden und scheinen spezifisch für die jeweiligen Individuen und damit geringer als interindividuelle Variabilitäten zu sein (Costello et al., 2009; Gao et al., 2010). Auch geschlechtsspezifische Unterschiede in den taxonomischen Profilen der Achselhöhlen-Mikrobiota sind beschrieben worden. Dabei scheint die Achselhöhlen-Mikrobiota männlicher Probanden häufiger von Corynebakterien dominiert zu sein als die weiblicher Probanden (Callewaert et al., 2013). Darüber hinaus identifizierten Troccaz et al. einen höheren Anteil an OTUs der Genera *Anaerococcus* und *Peptoniphilus* sowie einiger Corynebakterien-Spezies, *Peptoniphilus asaccharolyticus* und *Staphylococcus haemolyticus* in Proben männlicher Probanden im Vergleich zu weiblichen. Korrelationen zwischen taxonomischen Profilen und weiteren Merkmalen wie z. B. Nationalität oder Alter wurden nicht beschrieben.

Unabhängig von den verwendeten Techniken unterschieden sich die in diesem Kapitel vorgestellten Studien hinsichtlich der Anforderungen an die Probanden. Die Anweisungen, die Probanden vor der Probennahme hinsichtlich ihrer Hygienemaßnahmen und Essgewohnheiten erhalten, werden als *Preconditioning* bezeichnet. In einigen Studien wurden die Probanden angewiesen, für eine bestimmte Zeitspanne keine Deodorants oder Antiperspirantien zu verwenden und nicht zu baden (Costello et al., 2009; Egert et al., 2011; Grice et al., 2009).

Andere Studien verzichteten auf ein Preconditioning (Callewaert et al., 2013; Gao et al., 2010) oder wählten die Probanden nach ihren Hygienegewohnheiten aus (Callewaert et al., 2014; Troccaz et al., 2015). Die Wirkung von Deodorants und Antiperspirantien und ihr Einfluss auf die Achselhöhlen-Mikrobiota werden im nächsten Abschnitt beschrieben.

### **3.4 Übermäßiger Körpergeruch, seine Bekämpfung und die Auswirkungen auf die Mikrobiota der Achselhöhle**

Übermäßiger Schweißgeruch stellt in der Regel kein medizinisches, sondern ein kosmetisches Problem dar. Dennoch können Betroffene z. T. sehr stark unter der Situation leiden. Im Rahmen kosmetischer Forschung wurden sehr viele Substanzen entwickelt, die die vermehrte Schweißbildung und den Körpergeruch reduzieren. Die regelmäßige Anwendung dieser Substanzen in Form von Hygieneprodukten hat auch einen Einfluss auf die Mikrobiota der Achselhöhle. Dies kann heute Mithilfe moderner molekularer Methoden beschrieben werden.

#### **3.4.1 Übermäßiges Schwitzen und Körpergeruch**

Die übermäßige Produktion von Schweiß ohne einen Zusammenhang zur Thermoregulation wird als Hyperhidrose bezeichnet (Naumann et al., 2003; Ram et al., 2007). Bei den am stärksten und am häufigsten betroffenen Körperregionen handelt es sich um die Achselhöhle, die Handinnenflächen, die Fußsohlen und die Stirn (Stashak und Brewer, 2014). Während es für die sekundäre Hyperhidrose eine ersichtliche physiologische oder metabolische Ursache, wie beispielsweise eine Störung im endokrinen System oder im Nervensystem, gibt, bleibt die Ursache der primären Hyperhidrose unklar (Slingluff et al., 2013).

Durch die stark erhöhte Schweißmenge wird das Bakterienwachstum an den entsprechenden Hautbereichen erhöht und es kommt zu einem sehr sarken Körpergeruch, was als Osmidrose oder Bromhidrose bezeichnet wird (Piérard et al., 2003). Körpergerüche stellen ein sehr starkes interindividuelles Signal dar und können, wie beispielsweise im Falle von Stress- oder Angstschwitzen auch mit emotionalen Zuständen im Zusammenhang stehen. Es ist anzunehmen, dass die menschliche Wahrnehmung von Körpergerüchen durch kulturelle Faktoren sowie durch die individuelle Entwicklung geformt wird (Kanlayavattanakul und Lourith, 2011). Menschen, die an starkem Körpergeruch und vermehrter Schweißbildung leiden, sind häufig eingeschränkt in ihren sozialen Interaktionen mit anderen Menschen und in Folge dessen wird ihr Selbstbewusstsein gemindert (Kanlayavattanakul und Lourith, 2011).

Die Behandlungsmöglichkeiten der schweren Form der Hyperhidrose und des damit einhergehenden Körpergeruches umfassen diverse meist invasive Techniken (Stashak und Brewer, 2014). Durch lokale Botox-Injektionen wird die Ausschüttung von Acetylcholin aus den

präsynaptischen Neuronen inhibiert und damit die Schweißmenge reduziert. Oxybutynin ist ein anticholinerges Mittel zur oralen Einnahme. Operative Maßnahmen umfassen die Exzision von betroffenem Hautgewebe, die Entfernung von Fettgewebe inklusive Schweißdrüsen mittels Kürettage oder die Entfernung von Nervenbahnen mittels Sympathectomie. Dahingegen werden mildere Formen der Hyperhidrose und des thermoregulatorischen Schwitzens im Rahmen der täglichen Körperhygiene durch das Auftragen von Deodorants und Antiperspirantien gelindert. Die häufig in diesen Produkten verwendeten Substanzen und ihre Wirkweisen werden im folgenden Kapitel beschrieben.

### **3.4.2 Topische Behandlung von Schweißbildung und Körpergeruch durch Deodorants und Antiperspirantien**

Die topische Behandlung übermäßiger Schweißbildung ist mithilfe von Antiperspirantien möglich, während Deodorants der Behandlung von Körpergeruch dienen. Häufig werden diese beiden Substanzgruppen in der Literatur getrennt voneinander beschrieben, was einerseits durch die unterschiedlichen Wirkmechanismen bedingt ist und andererseits dadurch, dass diese beiden Substanzgruppen in den USA aufgrund regulatorischer Vorgaben getrennt voneinander vermarktet werden. Während Deodorants als Kosmetika eingestuft werden und frei verkäuflich sind, werden Antiperspirantien von der FDA (*US Food and Drug Administration*) als OTC (*over the counter*)-Arzneimittel eingestuft und strengen Kriterien zur Prüfung der Wirksamkeit und Sicherheit unterworfen (Piérard et al., 2003). In Europa gibt es diese regulatorische Trennung nicht, weshalb die freiverkäuflichen Kosmetika häufig Formulierungen aus beiden Substanzklassen enthalten und sowohl die Linderung des Schwitzens als auch die Beseitigung des Körpergeruches bewirken können. Sie werden im täglichen Sprachgebrauch als Deodorants oder Deos bezeichnet.

Die Wirksubstanzen und Wirkmechanismen beider Gruppen sind in Box 3 dargestellt (Kanlayavattanakul und Lourith, 2011; Piérard et al., 2003). Die Wirkmechanismen der Antiperspirantien sind im Wesentlichen physikalisch: Aluminiumsalze sind in saurem Milieu löslich und präzipitieren bei einer Erhöhung des pH-Wertes auf der Haut. Dabei bilden sich Pfropfen (*plugs*) im Bereich der Öffnung des Schweißdrüsenduktus und es kommt zu einer Reduktion der Schweißmenge auf der Haut (Labows et al., 1999). Auch Filmbildner und Adstringentien können durch den Verschluss, bzw. die Verengung der Drüsenöffnungen an der Hautoberfläche den Austritt von Schweiß reduzieren (Gordon und Maibach, 1969; Sato und Dobson, 1969). Substanzen, die Feuchtigkeit absorbieren, reduzieren die Schweißmenge nach Austritt (Piérard et al., 2003).

In Deodorants verhindern antimikrobielle Substanzen die Entstehung von Körpergeruch durch die Reduktion der Keimzahl auf der Hautoberfläche und Parfümstoffe überdecken bereits vorhandenen Körpergeruch. Der Einsatz antimikrobieller Substanzen wie beispielsweise Triclosan ist seit vielen Jahren verbreitet und stellt eine sehr effiziente Methode der Körpergeruchskontrolle dar (Kanlayavattanakul und Lourith, 2011). In der Regel wirken antimikrobielle Substanzen unspezifisch auf alle Mitglieder der Mikrobiota. Jedoch konnten Fujii et al. eine sehr effiziente spezifische antimikrobielle und inhibierende Wirkung auf das geruchsbildende Isolat *Anaerococcus* A20 zeigen, während die Anzahl der anderen Mitglieder der Mikrobiota nicht reduziert wurden (Fujii et al., 2014). Die Identifizierung antimikrobieller oder inhibierender Substanzen mit spezifischer Wirkung auf geruchsbildende Bakterien könnte somit eine vielversprechende Perspektive für die Entwicklung neuer Deosubstanzen darstellen. Eine Neutralisierung körperlgeruchsbildender Fettsäuren ist mithilfe von Carbonaten und die Neutralisierung von Sulfanylalkanolen ist mit sulfhydrylgruppenhaltigen Substanzen möglich (Davis, 1998; Labows et al., 1999). Eine relativ unspezifische Methode stellt die Absorption von Schweißflüssigkeit und den darin gelösten Substanzen oder eine Auslöschung des Geruches durch die Ummantelung der Geruchsstoffe dar

## Wirksubstanzen in

### Antiperspirantien und Deodorants

#### Antiperspirantien:

- Aluminiumsalze: Verschluss der Schweißporen, Aluminiumchlorohydrat (ACH), Aluminiumsulfat, Aluminiumchlorid
- Adstringentien: Verengung der Schweißporen Kalialaun (Aluminiumkaliumsulfat), Gerbsäure, Formaldehyd
- Filmbildner: Verschluss der Schweißporen Olefinsäureamide und –polymere
- Absorbierer: Absorption der Feuchtigkeit, Kombination aus Ölsäure und Glyceryllaurat zur Ausbildung lyotroper Flüssigkristalle

#### Deodorants:

- Antimikrobielle Substanzen: Metallsalze, Alkohol, Triclosan, quaternäre Ammoniumverbindungen, Glycerin-Fettsäureester, Pflanzenextrakte
- Überdeckung des Geruches: Parfümstoffe
- Neutralisierung des Geruches: Natriumbicarbonat, Natriumhydrogencarbonat, Zinkcarbonat, Substanzen mit Sulfhydrylgruppen
- Absorption oder Auslöschung des Geruches: Zinksalz der Rizinolsäure, Metallooxide, Cyclodextrine, Silikate, Carbonate
- Enzyminhibitoren:
  - kompetitive Inhibitoren der Aminoacylase und Cystathionin- $\beta$ -Lyase: O-Acyl-Serin und –Threonin
  - Chelatoren: Ethylendiamindibernsteinsäure (EDDS), Diethylentriaminpentaessigsäure (DTPA)
  - Inhibitoren der Exoenzyme Arylsulphatase und  $\beta$ -Glucuronidase
- Ersatzsubstrate: Glutaminkonjugate von Parfümstoffen

**Box 3: Übersicht der Wirksubstanzen und -mechanismen in Antiperspirantien und Deodorants.**



(Kanlayavattanakul und Lourith, 2011). Spezifische Inhibition der mit Körpergeruchsentstehung im Zusammenhang stehenden Enzyme ist durch kompetitive Substrate oder Chelatoren möglich. Egert et al. identifizierten in einem *Screening* basierend auf einem photometrischen *Assay* diverse Substanzen und Substanzgemische pflanzlicher Herkunft, welche die C-S-Lyase-Aktivität in Zellextrakten von *Staphylococcus hominis* inhibierten (Egert et al., 2013). Darüber hinaus wurde eine diverse Gruppe von Substanzen zur Inhibition von Exoenzymen beschrieben (Eigen und Froebe, 1997). Eine innovative Strategie zur Vorbeugung der Körpergeruchsentstehung stellt der Einsatz von Ersatzsubstraten (*surrogate substrates*) dar. Natsch et al. synthetisierten ein Glutaminkonjugat mit einem Duftstoff anstelle des geruchsbildenden Moleküles 3M2H und demonstrierten die Abspaltung des Duftstoffes durch die N- $\alpha$ -Acylglutaminaminoacylase *in vitro* (Natsch et al., 2005).

### **3.4.3 Einfluss von Deodorants und Antiperspirantien auf die Mikrobiota der Achselhöhle**

Zu diesem Zeitpunkt existieren drei veröffentlichte Studien, die den Einfluss von Antiperspirantien und Deodorants auf die Achselhöhle und die Achselhöhlen-Mikrobiota beschreiben. Im Rahmen ihrer Studie zur Korrelation olfaktorischer Daten mit Daten zur Achselhöhlen-Mikrobiota (s. Tabelle 1) führten Troccaz et al. zusätzlich einen Vergleich zwischen Probanden durch, die Antiperspirantien verwenden (*AP-users*) und denen, die keine Antiperspirantien verwenden (*non-AP users*) (Troccaz et al., 2015). Es konnten Koinzidenzen zwischen der Verwendung von Antiperspirantien und physikochemischen, olfaktorischen und mikrobiologischen Eigenschaften der Achselhöhle beschrieben werden. Der pH-Wert, die Intensität des Schweißgeruches sowie die Keimzahl in der Achselhöhle waren in der Gruppe der *AP users* niedriger als in der Gruppe der *non-AP users* und stiegen jeweils zum Abend hin an. Der Unterschied der Keimzahlen (*AP users vs. non-AP users*) lag im Mittel bei einem Faktor von 50. Weiterhin konnten signifikante Unterschiede in der Intensität des Körpergeruches innerhalb der Gruppe der *non-AP users* beobachtet werden, jedoch nicht innerhalb der Gruppe der *AP-users*. Diese Ergebnisse deuten auf einen möglichen Einfluss von Antiperspirantien auf die Mikrobiota der Achselhöhle hin.

Callewaert et al. untersuchten den Einfluss der Verwendung von Deodorants und Antiperspirantien auf die Zusammensetzung der Achselhöhlen-Mikrobiota bei jeweils 4 Probanden über einen Zeitraum von zehn Wochen (s. Tabelle 1) (Callewaert et al., 2014). Dabei konnte ein deutlicher Effekt der Anwendung von Antiperspirantien auf die Zusammensetzung der Achselhöhlen-Mikrobiota beschrieben werden. Bei den Probanden handelte es sich um Menschen, die regelmäßig Deodorants (*Deo users*) oder Antiperspirantien (*AP users*)

verwendeten, die Verwendung für vier Wochen aussetzten und anschließend wieder aufnahmen. Im Falle der *AP users* konnte mittels *DGGE* eine deutliche Verschiebung des taxonomischen Profiles von Mitgliedern der *Actinobacteria* (*Corynebacterium spp.*) hin zu Mitgliedern der *Firmicutes* (*S. epidermidis* und *Staphylococcus spp.*) bei Aussetzen der AP-Anwendung beobachtet werden. Die entgegengesetzte Verschiebung der Lebensgemeinschaft trat bei Wiederaufnahme der AP-Verwendung ein. Die größten intraindividuellen Unterschiede konnten dabei während der Übergangsphasen beobachtet werden. Innerhalb einer Phase mit bzw. ohne AP war das Profil der mikrobiellen Lebensgemeinschaft stabil. Im Falle der *Deo users* war der Effekt weniger eindeutig und es konnte kein einheitlicher Trend innerhalb der beobachteten Probanden beschrieben werden. Auch Urban et al. beschrieben die Änderungen in der Zusammensetzung der Achselhöhlen-Mikrobiota nach Aussetzen der Deo- bzw. AP-Anwendung (Urban et al., 2016). Auch hier kam es zu einer Verschiebung der Lebensgemeinschaften hin zu *Staphylococcaceae*, während die Lebensgemeinschaften von Probanden, die gewohnheitsmäßig keine Kosmetikprodukte anwendeten, von *Corynebacterium* dominiert wurden. In beiden Studien war sowohl die Vielfalt der bakteriellen Taxa als auch die Diversität der Lebensgemeinschaften höher, wenn AP oder Deo verwendet wurden. Diese Untersuchungen zeigen, dass die Anwendung von Hygieneprodukten zur Linderung des Schwitzens und Reduktion des Körpergeruches einen Einfluss auf die Zusammensetzung der Achselhöhlen-Mikrobiota ausüben.

## 4 Zielsetzung

Die Mikrobiota des Menschen spielt eine bedeutende Rolle bei physiologischen Prozessen und der Entstehung von Erkrankungen. Mithilfe klassischer kultivierungsbasierter Techniken kann jedoch nur ein Bruchteil ihrer Mitglieder beschrieben werden. Hochdurchsatzsequenzierung (NGS) bietet eine Möglichkeit zur direkten kultivierungsunabhängigen Beschreibung komplexer mikrobieller Lebensgemeinschaften. Insbesondere die Metatranskriptomik eignet sich zur Untersuchung der aktiven Mitglieder der Mikrobiota und zur Beschreibung ihrer Funktionen. Aufgrund der Limitierung im Hinblick auf die Menge des Probenmaterials, das sich aus menschlichen Proben gewinnen lässt, sind derzeit nur wenige Beschreibungen auf der Ebene der Metatranskriptomik verfügbar.

Das Ziel der vorliegenden Arbeit ist die Beschreibung des Metatranskriptoms von Proben aus dem menschlichen Ohr im Zusammenhang mit Gehörgangsentzündung und aus der Achselhöhle im Zusammenhang mit der Entstehung von Körpergeruch. Zum einen soll eine Methode zur Durchführung von Metatranskriptomik-Analysen an kleinen Proben etabliert werden. Darüber

hinaus soll die Zusammensetzung der aktiven Vertreter der Mikrobiota sowie ihr funktionales Profil beschrieben werden. Schließlich werden Mikroorganismen und spezifische Funktionen beschrieben, die im Zusammenhang mit Gehörgangsentzündung und Schweißgeruch stehen.

Zum Zeitpunkt des Verfassens dieser Arbeit liegt keine Metatranskriptomik-Studie des Ohres im Zusammenhang mit Gehörgangsentzündung vor. Aus kultivierungsabhängigen Beschreibungen ist die polymikrobielle Natur der Gehörgangsentzündung mit *Staphylococcus aureus* und *Pseudomonas aeruginosa* als hauptsächliche Pathogene bekannt. Jedoch ist bisher nicht beschrieben, welche molekularen Mechanismen im Krankheitsverlauf eine Rolle spielen.

Auch aus der Achselhöhle existieren zum Zeitpunkt des Verfassens der Arbeit keine Metatranskriptomik-Studien. Zwar wurden die Lebensgemeinschaften der Achselhöhle kultivierungsabhängig und –unabhängig beschrieben. Funktionen im Zusammenhang mit der Entstehung von Körpergeruch wurden bisher ausschließlich kultivierungsbasiert beschrieben und liefern nach wie vor kein vollumfängliches Bild der Schweißgeruchsentstehung. In dieser Arbeit wird erstmals eine Hochdurchsatz-Sequenzieretechnik zur Beschreibung der Taxonomie und Funktion der Achselhöhlen-Mikrobiota mit großer Sequenzierentiefe verwendet. Dabei sollen vor allem bisher unbekannte Mikroorganismen und Funktionen der Schweißgeruchsentstehung beschrieben werden. Darüber hinaus wird ein komparativer Ansatz gewählt, der durch den Vergleich von Proben mehrerer Probanden mit unterschiedlichem Körpergeruch Aussagen zu Funktionen mit Beteiligung an der Entstehung von Schweißgeruch machen soll.

## III. Material und Methoden

### 1 Probenmaterial

#### 1.1 Probe aus dem menschlichen Ohr

Die Untersuchungen zum Metatranskriptom des menschlichen Ohrs wurden mit einer Probe durchgeführt, die von einem Patienten mit einer Gehörgangsentzündung stammt. Im Rahmen einer Kollaboration mit einem Bielefelder Hals-Nasen-Ohrenarzt wurde eine Probe aus dem rechten Ohr eines 77-jährigen männlichen Patienten zur Verfügung gestellt. Die Probe wurde nach Erhalt bei -80 °C gelagert. Darüber hinaus wurde vom Arzt ein mikrobiologischer Befund zur Verfügung gestellt. Die aerobe Kultivierung zeigte starkes Wachstum von *Staphylococcus aureus*, während anaerobe Bakterien und Sprosspilze nicht nachgewiesen wurden.

#### 1.2 Proben der Achselhöhlen-Mikrobiota

Die Proben für die Metatranskriptom-Analysen der menschlichen Achselhöhlen-Mikrobiota wurden im Rahmen einer industriellen Kooperation mit einem Kosmetikhersteller bereitgestellt. Die Proben wurden in einer Deo-Studie generiert und es wurde jeweils 1 ml Probenmaterial pro Probe für die Untersuchungen zur Verfügung gestellt. Im Folgenden werden die vom Kooperationspartner durchgeführte Studie und die zugehörigen Arbeitsabläufe (Vorbereitung der Probanden, Probennahme, Bestimmung des *Malodor Score* und die Bestimmung der koloniebildenden Einheiten) beschrieben. Die Metatranskriptom-Analysen wurden anschließend am *Centrum für Biotechnologie (CeBiTec)* der Universität Bielefeld durchgeführt.

##### 1.2.1 Vorbereitung der Probanden

Jeweils sieben Tage vor der Probennahme begann eine sogenannte „Auswasch-Phase“, in der die Probanden angewiesen wurden, ausschließlich ein spezielles Deodorant zu verwenden, um eventuelle Kosmetik-Rückstände aus den Hautporen und Schweißdrüsen zu entfernen. Das in der Auswasch-Phase verwendete Deodorant setzt sich wie folgt zusammen (Angaben nach der Internationalen Nomenklatur für kosmetische Inhaltsstoffe, INCI): *Alcohol denat., Aqua, Parfum, Tetramethyl Acetyloctahydronaphthalenes*. Die Probanden wurden angewiesen, jeweils 24 h vor der Probennahme kein Deodorant zu verwenden.

Weiterhin wurde den Probanden für die Auswaschphase ein spezielles Duschgel mit folgender Zusammensetzung zur Verfügung gestellt (Angaben nach der Internationalen Nomenklatur für kosmetische Inhaltsstoffe, INCI): *Aqua, Sodium Laureth Sulfate, Cocamidopropyl Betaine, Sodium Chloride, PEG-7 Glyceryl Cocoate, Citric Acid, Sodium Benzoate, Laureth-2, Propylene*

*Glycol, PEG-55 Propylene Glycol Oleate, Sodium Citrate*. Die Probanden wurden angewiesen, ausschließlich dieses Duschgel während der Auswaschphase zu verwenden und jeweils 24 h vor der Probennahme nicht zu duschen.

Ebenfalls trugen die Probanden ein speziell für die Studie bereitgestelltes Baumwoll-T-Shirt für jeweils 12 h vor der Probennahme. Die Probanden wurden weiterhin angewiesen, 24 h vor der Probennahme keine Vitaminpräparate, Zwiebeln und keinen Knoblauch zu verzehren.

### 1.2.2 Deo-Studie

Die Deo-Studie wurde mit 24 männlichen Probanden durchgeführt, bei denen es sich um Mitarbeiter des Kooperationspartners handelte. Die Probanden wurden zu Beginn der Studie informiert und haben schriftlich der Teilnahme an der Studie zugestimmt. Diese Zustimmung konnte jederzeit widerrufen werden. Die Studie wurde im Einklang mit den *Guidelines* von *Cosmetics Europe* und der Deklaration von Helsinki durchgeführt.

Die Studie erstreckte sich über zwei Beprobungstage im Abstand von 19 Tagen. Die erste Probennahme fand am 16. Mai 2012 und die zweite am 4. Juni 2012 statt. Jeder Proband wurde an jedem Beprobungstag zweimalig beprobt: vor Anwendung ( $t_0$ ) und sechs Stunden nach Anwendung ( $t_6$ ). Die Proben wurden für die Metatranskriptomik und für die Bestimmung der koloniebildenden Einheiten verwendet. Zusätzlich wurde vor jeder Probennahme der *Malodor Score* bestimmt. Die Probanden wurden an jedem Tag zwei Gruppen zugeordnet. Gruppe A (Probanden 1-12 im Mai und Probanden 13-24 im Juni) wurden mit demineralisiertem Wasser behandelt. Gruppe B (Probanden 13-24 im Mai und 1-12 im Juni) wurden mit einem Deodorant, dem Wirkprodukt, behandelt. Somit liegen für jeden Probanden 4 Proben vor:  $t_0$  H<sub>2</sub>O,  $t_6$  H<sub>2</sub>O,  $t_0$  Deo und  $t_6$  Deo. Das Wirkprodukt setzt sich wie folgt zusammen (Angaben nach der Internationalen Nomenklatur für kosmetische Inhaltsstoffe, INCI): *Alcohol denat., Triethyl Citrate, Aqua, 1,2-Hexanediol, Caprylyl Glycol, 2-Benzylheptanol, Cocamidopropyl PG-Dimonium Chloride Phosphate, Phenoxyethanol, Tropolone, Amyl Cinnamal, Heliotropine*.

### 1.2.3 Bestimmung des *Malodor Score*

Der *Malodor Score* ist ein semiquantitatives Maß für die Intensität von unangenehmem Körpergeruch. In dieser Studie wird der *Malodor Score* ebenfalls von Mitarbeitern des Kooperationspartners (*Sniffer*) vor jeder Probennahme durch Riechen an der Achselhöhle der Probanden ermittelt. Die Probanden werden dabei nicht vollständig von den *Sniffern* gesehen, so dass die körperliche Erscheinung der Probanden keinen Einfluss auf die Beurteilung des Körpergeruchs hat. Es stehen insgesamt fünf (Mai), bzw. sechs (Juni) *Sniffer* pro Beprobungstag

zur Verfügung. Jeder Proband wird von vier *Sniffern* abgerochen. Zur Beurteilung des Körpergeruchs wird die folgende Klasseneinteilung verwendet:

- 0: Kein Geruch
- 1: Schwacher Geruch
- 2: Mittlerer Geruch
- 3: Intensiver Geruch
- 4: Sehr intensiver Geruch

Der *Malodor Score* wird schließlich durch Mittelung der vier Einzelwerte errechnet. Der *Malodor Score* wird für jede Probe jeweils unmittelbar vor der Probennahme ermittelt.

#### **1.2.4 Probennahme**

Die Probennahme wurde mithilfe einer etablierten Methode durchgeführt, die vom Kooperationspartner zur Beprobung der Achselhöhlen-Mikrobiota eingesetzt wird (Egert et al., 2011). Dabei wird ein steriler Kunststoff-Zylinder mit einer kreisförmigen Grundfläche von etwa 5 cm<sup>2</sup> fest auf die Haut gedrückt und mit 2,5 ml physiologischer Kochsalzlösung (0,9 % w/v) befüllt. Dann wird die Hautfläche innerhalb des Zylinders mit einem sterilen Glasstab geschabt und die Mikrobiota wird auf diese Weise ausgewaschen. Anschließend wird die Flüssigkeit mit einer sterilen Pipette entnommen und der Vorgang wird wiederholt.

Für die Metatranskriptom-Analysen werden die Probanden an der rechten Achselhöhle beprobt und jeweils 1 ml des gesamten Probenvolumens (etwa 4 ml) wird für diesen Zweck zur Verfügung gestellt. Ein weiterer Milliliter wird zur Bestimmung der koloniebildenden Einheiten verwendet.

#### **1.2.5 Bestimmung der koloniebildenden Einheiten**

Die Bestimmung der Keimzahl bzw. der koloniebildenden Einheiten in den Achselhöhlen-Proben erfolgt durch Kultivierung auf Vollmedium. Jeweils 1 ml der Probe wird mit 9 ml Inaktivator-Lösung (Zusammensetzung s. unten) vermischt. Die Inaktivator-Lösung dient dazu, die antimikrobielle Wirkung von kosmetischen Substanzen abzuschwächen oder zu eliminieren. Das Gemisch wird für 5 min bis maximal 30 min bei RT inkubiert und jeweils 0,5 ml werden auf zwei Agar-Platten *LTH-Trypton Soy Agar* (LTH-TSA) verteilt. Die Inkubation erfolgt für 7 Tage bei 30 °C bis 37 °C. Während der Inkubation teilen sich lebende kultivierbare Zellen und bilden lokal sichtbare Kolonien aus. Anschließend wird durch Auszählen der Kolonien und unter Berücksichtigung des Verdünnungsfaktors die Anzahl der koloniebildenden Einheiten pro ml Probe (KBE/ml) berechnet.

Zusammensetzung der Inaktivator-Lösung

1 g Trypton  
 8,5 g NaCl  
 5 g Na<sub>2</sub>S<sub>2</sub>O<sub>3</sub>  
 0,34 g KH<sub>2</sub>PO<sub>4</sub>  
 0,44 g K<sub>2</sub>HPO<sub>4</sub>

Mit LTH-Wasser (enthält Sojalecithin, Tween 80, Histidin) auf 1 l auffüllen.

## 2 Labormaterialien und Geräte

### 2.1 Chemikalien, Reagenzien und Medienzusätze

#### 2.1.1 Chemikalien

Tabelle 2: Liste der verwendeten Chemikalien und ihrer Bezugsquellen.

| Substanz           | Bezugsquelle   |
|--------------------|----------------|
| Agarose            | Peqlab         |
| Bromphenolblau     | Serva          |
| dNTPs              | Bioline        |
| EDTA               | Amersham       |
| Essigsäure         | Roth           |
| Ethanol            | VWR            |
| Ethidiumbromid     | Serva          |
| Glycerin 87%       | Merck          |
| Hefeextrakt        | Oxoid          |
| Hirn-Herz-Bouillon | Roth           |
| Natriumacetat      | Merck          |
| Salzsäure          | Roth           |
| Select Agar        | Invitrogen     |
| TRIS               | MP Biomedicals |
| Tween 80           | Sigma          |
| Xylencyanol FF     | Sigma          |
| β- Mercaptoethanol | Roth           |

#### 2.1.2 Oligonukleotide

Das *Primer*-Paar 341F und 806R wurde in der Test-PCR verwendet, in der die Kontamination isolierter Gesamt-RNA mit genomischer DNA überprüft wurde. Die *Primer* wurden in einer Konzentration von je 10 µM/µl verwendet. Die *Annealing*-Temperatur beträgt 56 °C (Yu et al., 2005). Die Nukleotid-Sequenzen der *Primer* werden im Folgenden dargestellt:

341F 5' CCTAYGGGRBGCASCAG 3'  
 805R 5'GGACTACNNGGGTATCTAAT 3'

### 2.1.3 Kits

Tabelle 3: Liste der verwendeten Kits und ihrer Bezugsquellen.

| Produkt  | Bezugsquelle         |
|--|----------------------|
| <i>Agilent DNA 1000/7500/High Sensitivity Kits</i> | Agilent Technologies |
| <i>Agilent RNA 600 Pico/Nano Kits</i>              | Agilent Technologies |
| <i>DNase I Kit</i>                                 | Roche                |
| <i>NucleoSpin Extract II Kit</i>                   | Macherey-Nagel       |
| <i>Ovation RNA-Seq System V2</i>                   | NuGEN                |
| <i>Ovation Ultralow DR Multiplex System 1-8</i>    | NuGEN                |
| <i>RNase-free DNase Set</i>                        | Qiagen               |
| <i>RNeasy Micro Kit</i>                            | Qiagen               |

### 2.1.4 Enzyme und Marker

Tabelle 4: Liste der verwendeten Enzyme und Marker und ihrer Bezugsquellen.

| Produkt                                     | Bezugsquelle |
|---|--------------|
| DNA-Marker 250bp-Leiter (P-705)             | MBBL         |
| <i>Phusion High-Fidelity DNA Polymerase</i> | Finnzymes    |

## 2.2 Verbrauchsmaterialien

Tabelle 5: Liste der verwendeten Verbrauchsmaterialien und ihrer Bezugsquellen.

| Produkt                              | Bezugsquelle               |
|--------------------------------------|----------------------------|
| Filter-Pipettenspitzen               | Biozym                     |
| Glaswaren                            | Schott                     |
| Nitrilhandschuhe                     | VWR International          |
| Parafilm                             | Pechiney Plastic Packaging |
| PCR-Stripes/Reaktionsgefäße (0,2 ml) | Sarstedt                   |
| Petrischalen                         | Greiner                    |
| Pipettenspitzen                      | Starlab GmbH               |
| Reaktionsgefäße (1,5 ml)             | Sarstedt                   |
| Reaktionsgefäße (15 ml)              | Greiner                    |
| Reaktionsgefäße (50 ml)              | Greiner                    |
| Ribotubes                            | MP Biomedicals             |



## 2.3 Geräte

Tabelle 6: Liste der verwendeten Geräte und Apparaturen und ihrer Bezugsquellen.

| <b>Gerät/Apparatur</b>                      | <b>Bezugsquelle</b>                   |
|---|---------------------------------------|
| <i>Agilent 2100 Bioanalyzer</i>             | <i>Agilent Technologies</i>           |
| Autoklav                                    | <i>Systec, Biomedis</i>               |
| <i>Bioruptor</i> Ultraschall-Bad            | <i>Diagenode</i>                      |
| Eismaschine                                 | <i>Scotsman Ice Systems</i>           |
| Elektrophorese-Kammern für Agarosegele      | Eigenbau, Universitätswerkstatt       |
| Feinwaage                                   | <i>Satorius</i>                       |
| Fotodrucker                                 | <i>Mitsubishi</i>                     |
| Gefriertruhe                                | <i>Sanyo</i>                          |
| Geldokumentationsanlage                     | <i>Alpha Innotech</i>                 |
| Gelkammer und Käme für Agarosegele          | Eigenbau, Universitätswerkstatt       |
| Glaspipetten                                | <i>Brand</i>                          |
| <i>Illumina MiSeq</i> Sequenziergerät       | <i>Illumina</i>                       |
| Kühlschrank-Gefrierkombination              | <i>Privileg</i>                       |
| Kühltischzentrifuge 5417 R                  | <i>Eppendorf</i>                      |
| Magnetrührer mit Heizplatte                 | <i>IKA</i>                            |
| Magnet-Separator für 1,5-ml-Reaktionsgefäße | Eigenbau, Universitätswerkstatt       |
| Mikrowellengerät                            | <i>Siemens</i>                        |
| Thermomixer                                 | <i>Bioer</i>                          |
| Schüttelinkubator                           | <i>New Brunswick Scientific</i>       |
| <i>Nanodrop</i> Photospektrometer           | <i>Peqlab</i>                         |
| <i>PCR-Thermocycler</i>                     | <i>Biometra</i>                       |
| Peleusball                                  | <i>VWR</i>                            |
| pH-Meter                                    | <i>Knick Elektronische Messgeräte</i> |
| <i>Pipetman</i> Pipetten                    | <i>Eppendorf</i>                      |
| Reinstwasseranlage                          | <i>TKA</i>                            |
| <i>RiboLyser</i>                            | <i>PeqLab</i>                         |
| Spannungsgeber                              | <i>Bio-Rad Laboratories</i>           |
| Tischzentrifuge 5424                        | <i>Eppendorf</i>                      |
| <i>Vortex Genie 2</i>                       | <i>Scientific Industries</i>          |
| Wasserbad                                   | P-D Industriegesellschaft mbH         |

## 2.4 Software und Bioinformatik-Tools

Tabelle 7: Liste der verwendeten Software und Bioinformatik-Tools und ihrer Bezugsquellen/Referenzen.

| Programm/Tool     | Verwendungszweck   | Bezugsquelle/ Referenz  |
|-------------------|--|---|
| ADOBE READER X    | Verfassen der Arbeit   | Adobe Systems   |
| BLAST             | Sequenzvergleich   | (Altschul et al., 1990)   |
| CLUSTALOMEGA      | Erstellen von <i>Multiple Sequence Alignments</i> und eines phylogenetischen Baumes                            | (Goujon et al., 2010; Larkin et al., 2007; McWilliam et al., 2013)  |
| FASTX-Toolkit     | Qualitätsanalyse der Sequenzierreads   | <a href="http://hannonlab.cshl.edu/fastx_toolkit/index.html">http://hannonlab.cshl.edu/fastx_toolkit/index.html</a> |
| GENEDOC           | Visualisierung von <i>Multiple Sequence Alignments</i>   | (Nicholas und Nicholas, 1997)   |
| METRANS           | Metatranskriptom-Sequenzanalyse  | Dr. Christina Ander   |
| MG-RAST           | Metatranskriptom-Sequenzanalyse  | (Meyer et al., 2008)  |
| OFFICE 2010 Paket | Verfassen der Arbeit   | Microsoft   |
| ORIGINPRO 9.1G    | Erstellen von Hauptkomponenten-Analysen  | Origin Lab  |
| SEQPREP           | <i>Quality trimming</i> und <i>merging</i> der <i>Reads</i>  | <a href="https://github.com/jstjohn/SeqPrep">https://github.com/jstjohn/SeqPrep</a>                                 |
| KEGGMAPPER        | Graphische Darstellung von Enzymfunktionen und zugehörigen Stoffwechselwegen, die mit Körpergeruch korrelieren | <a href="http://www.genome.jp/kegg/mapper.html">http://www.genome.jp/kegg/mapper.html</a>                           |

## 3 Medien, Puffer und Lösungen

Soweit nicht anders angegeben, werden alle Puffer in deionisiertem Wasser angesetzt.

### BYT

37 g/l Hirn-Herz-Bouillon

10 g/l Hefeextrakt

15 g/l Select Agar

1 % (v/v) Tween 80

Alle Bestandteile gut in Wasser lösen und autoklavieren

dNTP-Mix

2,5 mM dATP  
2,5 mM dCTP  
2,5 mM dGTP  
2,5 mM dTTP

in autoklaviertem Wasser lösen

TAE-Puffer (50fach)

2 M Tris  
500 mM Natriumacetat  
50 mM EDTA

mit Essigsäure auf pH 7,8 einstellen

Agarose

1 % (w/v) Agarose

in 1x TAE-Puffer durch Aufkochen lösen

DNA-Ladepuffer

1 mM EDTA  
50 % (w/v) Glycerin  
0,1 % (w/v) Bromphenolblau  
0,1 % (w/v) Xylencyanol FF

mit Salzsäure bis zum Umschlagspunkt (dunkelgrün) titrieren

TE-Puffer

10 mM Tris  
1 mM EDTA

mit Salzsäure auf pH 8 einstellen

## 4 Allgemeine molekularbiologische Arbeiten

### 4.1 Isolierung von Gesamt-DNA aus *C. jeikeium* K411

Die gesamt-DNA aus dem Haut-Keim *Corynebacterium jeikeium* K411 wurde in dieser Arbeit als *Template* für die Positivkontrolle für die Test-PCR auf Verunreinigung von RNA durch genomische DNA verwendet. Die Inokulation einer BYT-Festmedium-Platte erfolgte durch Überführen von Zellmaterial von einer bebrüteten Platte auf eine sterile Platte. Anschließend wurde die Platte für 2 Tage bei 37 °C bebrütet.

Für die DNA-Isolierung wurde Zellmaterial mit einer sterilen Pipetten-Spitze von der Platte geschabt und in ein steriles 1,5-ml-Reaktionsgefäß überführt. Nach Zugabe von 1 ml autoklaviertem Wasser wurde das Zellmaterial gründlich resuspendiert und 5 min bei 100 °C im

Thermomixer aufgeköcht. Anschließend erfolgte eine Zentrifugation für 5 min bei  $20238 \times rcf$  zur Entfernung der Zelltrümmer. Der Überstand wurde in ein neues 1,5-ml-Reaktionsgefäß überführt und konnte anschließend als *Template* für die Test-PCR verwendet werden.

## 4.2 Polymerase-Kettenreaktion (PCR)

Mithilfe der Polymerase-Kettenreaktion (PCR) können definierte DNA-Abschnitte amplifiziert werden. Die Grundvoraussetzung hierfür ist jedoch, dass die Sequenzen, die die Zielsequenz flankieren, bekannt sind, damit entsprechende *Primer* generiert werden können. Die Amplifikation der Matrizen-DNA (*Template*) während der PCR erfolgt in mehreren Zyklen, wobei jeder Zyklus aus den folgenden drei Schritten besteht:

1. Denaturierung: Durch Erhitzen werden die Basenpaarungen der *Template*-DNA aufgebrochen, so dass die DNA anschließend einzelsträngig vorliegt und für die Amplifikation zugänglich ist. Dabei werden Temperaturen von über 90 °C verwendet.
2. *Primer-Annealing*: Mittels Komplementärbasenpaarung lagern sich die *Primer* an die flankierenden Sequenzen der Zielsequenz. Dabei lagert sich jeweils der entsprechende *Primer* an das 5'-Ende der Sequenz und definiert somit den Startpunkt der Amplifikation. Die *Annealing*-Temperatur wird durch die *Primer*-Sequenzen definiert.
3. Elongation: Mithilfe einer thermostabilen DNA-Polymerase werden die komplementären DNA-Stränge ausgehend vom 3'-Ende der *Primer* synthetisiert, so dass schließlich die Zielsequenz verdoppelt und doppelsträngig vorliegt. Die Elongationstemperatur wird durch die verwendete DNA-Polymerase definiert.

Da die Anzahl der DNA-Produkte sich ab dem dritten Zyklus mit jedem weiteren Zyklus verdoppelt, erfolgt die Amplifikation exponentiell.

In dieser Arbeit wurde eine PCR zum Nachweis von DNA-Kontaminationen nach der Gesamt-RNA-Isolierung mit der *Fusion High-Fidelity* DNA-Polymerase durchgeführt. Weiterhin erfolgte eine PCR im Rahmen der Herstellung der Sequenzierbibliotheken mit dem *Ovation Ultralow DR Multiplex System 1-8*. Einen Spezialfall der PCR stellt die isothermale Amplifikation dar. Dabei erfolgt die Denaturierung, das *Primer-Annealing* sowie die Elongation bei der gleichen Temperatur. In dieser Arbeit erfolgte die Amplifikation der cDNA mit dem *Ovation RNA-Seq System* isothermal.

### 4.3 Agarose-Gelelektrophorese

Mithilfe der Agarose-Gelelektrophorese kann ein Gemisch aus DNA-Molekülen unterschiedlicher Länge aufgetrennt werden. DNA-Moleküle wandern in einem elektrischen Feld aufgrund der negativen Ladung der Phosphatgruppen zur Anode. Unterschiedlich lange DNA-Moleküle können in einem freien elektrischen Feld nicht aufgetrennt werden, da jedes Nukleotid eine negative Ladung trägt und somit das Verhältnis von Ladung zu Masse für alle DNA-Moleküle nahezu konstant ist. Daher fungiert die Agarose als Größenausschlussmatrix, innerhalb derer kurze DNA-Moleküle schneller und somit innerhalb eines gegebenen Zeitintervalls weiter wandern als lange. Bei hinreichend guter Auftrennung ordnen sich Moleküle gleicher Länge in einer Bande an. Zur Größenbestimmung unbekannter DNA-Längen werden Banden mit dem Bandenmuster eines Längenstandards (Marker) abgeglichen.

In dieser Arbeit wurde eine 1 %-ige Agaroselösung in TAE-Puffer (1fach) angesetzt, in einer Mikrowelle aufgeköcht und auf einem Magnetrührer vollständig gelöst. Die warme Agaroselösung wurde in eine Gelkammer gegossen und der Probenkamm eingesteckt. Nachdem das Gel ausgehärtet war, wurde der Probenkamm gezogen und das Gel wurde mit TAE-Puffer überschichtet. Jeweils 7 µl eines PCR-Ansatzes wurden mit je 2 µl Ladepuffer vermischt und in die Probenaschen gegeben. Zusätzlich wurden 5 µl des Markers in eine separate Probenasche gegeben. Der Gellauf erfolgte 35 min bei 80 V. Anschließend wurde das Gel 5 min in einer Ethidiumbromid-Färbelösung (1µg/ml) gefärbt und mithilfe der Geldokumentationsanlage abfotografiert.

### 4.4 Konzentrationsbestimmung und Qualitätskontrolle von DNA und RNA

In dieser Arbeit wurden zwei unterschiedliche Methoden für die Qualitätskontrolle und Konzentrationsbestimmung von DNA und RNA verwendet.

Spektrophotometrie (*NanoDrop*): Bei der Analyse von Nukleinsäuren mithilfe des *NanoDrop* wird ein Absorptionsspektrum für jede Probe aufgenommen und die Konzentration an DNA bzw. RNA über die Absorption bei 260 nm berechnet. Zur Qualitätskontrolle werden zusätzlich zwei weitere Absorptionswerte herangezogen: Das Verhältnis der Absorption bei 260 nm und bei 280 nm ( $A_{260}/A_{280}$ ) beschreibt die Kontamination der Probe mit Protein. Um eine hinreichende Reinheit der Probe zu gewährleisten, sollte dieser Wert im Falle von RNA etwa 2,0 und im Falle von DNA etwa 1,8 betragen. Das Verhältnis der Absorption bei 260 nm und bei 230

nm weist auf eine Verunreinigung der Probe mit Salzen und anderen Substanzen hin und sollte im Bereich von 2,0 – 2,2 liegen (Thermo Scientific, 2013).

Elektrophorese (*Bioanalyzer*): Die Analyse von Nukleinsäuren in Mikrofluidik-Chips mithilfe des *Bioanalyzer* ist analog zur Gelelektrophorese. Die Proben wandern durch mit einer Gelmatrix gefüllte Kanäle, werden elektrophoretisch aufgetrennt und mittels Fluoreszenzfärbung detektiert. Die DNA- bzw. RNA-Längenverteilung der Probe wird anschließend in einem Elektropherogramm dargestellt. Es wird jeweils nur 1 µl Probenvolumen benötigt und die Detektionsgrenzen liegen in Abhängigkeit vom verwendeten Kit im Nanogramm- oder Picogramm-Bereich. Die Beschreibung der Längenverteilung und der Konzentration erfolgt hierbei über die Marker-Fragmente im Leiter-Elektropherogramm sowie über die Marker-Fragmente, die jeder Probe zugegeben werden. Im Falle von RNA wird zusätzlich das *Peak-Verhältnis* zwischen 16S- und 23S rRNA errechnet und in Form der *RNA Integrity Number (RIN)* angegeben. Die RIN kann einen Wert zwischen null und zehn annehmen, wobei ein hoher Wert auf eine besonders hohe Qualität und Integrität hinweist.

## 5 Metatranskriptomik

Für die Metatranskriptomik-Analyse von Proben aus dem menschlichen Ohr und aus der Achselhöhle wurde im Rahmen dieser Arbeit ein Laborverfahren entwickelt und für die Bearbeitung sehr geringer Probevolumina optimiert. Im Folgenden werden die einzelnen Arbeitsschritte detailliert beschrieben.

### 5.1 RNA-Isolierung

Die RNA-Isolierung wurde säulenbasiert mithilfe des *RNeasy Micro Kits* durchgeführt. Dabei werden Zellen in einem Guanidin-Isothiocyanat-haltigen Puffer gelöst, mechanisch im *Ribolyser* aufgeschlossen und das Zell-Lysat einschließlich der RNA an eine Säule gebunden. Nach mehreren Waschschritten und einem DNase-Verdau kann die RNA schließlich von der Säule eluiert werden. Grundsätzlich wurden bei Arbeiten mit RNA Handschuhe getragen und ausschließlich sterile, 3fach autoklavierte oder RNase-freie Verbrauchsmaterialien und Lösungen verwendet. Die Pipettier-Schritte wurden mit speziellen für RNA-Arbeiten bestimmten Pipetten unter Verwendung von Filter-Pipetten-Spitzen durchgeführt. Weiterhin wurden RNA-Arbeiten stets räumlich von DNA-Arbeiten getrennt um unerwünschte Abbauprozesse zu vermeiden. Die Lagerung der RNA während der Durchführung des Protokolls erfolgt wann immer möglich auf Eis. Abhängig von der Probe variieren die ersten Schritte des Protokolls.

Im Falle der Ohr-Probe wurden 800 µl RLT-Puffer (mit 10 µl β-Mercaptoethanol pro ml RLT) in einem 1,5-ml Reaktionsgefäß vorgelegt, der Q-Tip wurde eingetaucht und das Probenmaterial wurde durch kräftiges Auf- und Abpipettieren mit einer sterilen Pipetten-Spitze herausgelöst. Nachdem der Q-Tip entfernt wurde, wurden erneut 200 µl RLT-Puffer zugegeben. Alle folgenden Arbeitsschritte wurden nach dem untenstehenden Protokoll durchgeführt.

Im Falle der Achselhöhlenproben wurden die Reaktionsgefäße etwa 30 min auf Eis aufgetaut und danach 15 min bei 4 °C und 16595 rcf zentrifugiert. Anschließend wurden 950 µl des Überstandes verworfen und 700 µl RLT-Puffer (mit 10 µl β-Mercaptoethanol pro ml RLT) zugegeben. Alle folgenden Arbeitsschritte wurden nach dem untenstehenden Protokoll durchgeführt.

Protokoll (etabliertes RNA-Isolierungsprotokoll am Centrum für Biotechnologie, Universität Bielefeld mit geringfügigen Abwandlungen):

- Zellsuspension in RLT-Puffer in ein Ribotube überführen
- Proben 30 s im Ribolyser bei einer Geschwindigkeit von 6500 aufschließen
- Proben 5 min auf Eis kühlen und erneut 30 s bei einer Geschwindigkeit von 6500 aufschließen
- Ribotubes 15 min bei 4 °C und 16595 rcf zentrifugieren
- 500 µl Ethanol in einem 1,5-ml Reaktionsgefäß vorlegen, den Überstand des Ribolysats zugeben, durch Pipettieren gut vermischen und das Gemisch auf eine RNeasy Micro Säule geben
- Säule 30 s bei 20238 rcf zentrifugieren und Durchfluss verwerfen
- 700 µl RW1-Puffer auf die Säule geben, 30 s bei 20238 rcf zentrifugieren und Durchfluss verwerfen
- 70 µl RDD-Puffer mit 10 µl DNase I vermischen, auf die Säule geben und 15 min bei RT inkubieren
- 500 µl RW1 zugeben, 30 s bei 20238 rcf bei RT zentrifugieren, Durchfluss verwerfen und die Säule in ein neues Auffanggefäß überführen
- 500 µl RPE-Puffer zugeben, 30 s bei 20238 rcf bei RT zentrifugieren und Durchfluss verwerfen
- 500 µl 80 % Ethanol zugeben, 2 min bei 20238 rcf bei RT zentrifugieren, Durchfluss verwerfen und Säule in ein neues Auffanggefäß überführen
- 5 min bei 20238 rcf und RT zentrifugieren um die Säule zu trocknen
- Säule in ein 1,5-ml Reaktionsgefäß überführen

- RNase-freies Wasser für Elution zugeben, 1 min bei 20238 rcf zentrifugieren und Säule verwerfen

Im Falle der Ohrprobe wurde in 25 µl und im Falle der Achselhöhlenproben wurde in 12 µl Wasser eluiert. Die RNA wurde bei -80 °C gelagert.

## 5.2 PCR zur Überprüfung der RNA auf Verunreinigung durch genomische DNA

Aufgrund der strukturellen Ähnlichkeit zwischen DNA und RNA kann es passieren, dass isolierte RNA noch genomische DNA aus der Probe enthält. Da die DNA ein Metatranskriptomik-Experiment verfälschen würde, muss sie vollständig entfernt werden. Um selbst kleinste Mengen an DNA-Kontaminationen detektieren zu können, wird eine Test-PCR unter Verwendung von 16S-rDNA-Primern durchgeführt. Die *Primer 341F* und *806R* binden an konservierte Bereiche der 16S-rDNA und flankieren die variablen Regionen *V3* und *V4*. Das entsprechende Amplifikat hat eine Größe von etwa 460 bp. Die PCR wurde mit der *Phusion High-Fidelity* DNA-Polymerase durchgeführt. Als *Template* diente je 1 µl der isolierten RNA. Als Positivkontrolle wurde genomische DNA von *C. jeikeium* K411 und als Negativkontrolle RNase-freies Wasser verwendet.

### PCR-Ansatz

|                                  |         |
|----------------------------------|---------|
| RNase-freies Wasser              | 12,0 µl |
| 5 × <i>HF Buffer</i>             | 4,0µl   |
| dNTP-Mix                         | 0,5 µl  |
| <i>Primer 341F</i>               | 0,75 µl |
| <i>Primer 806R</i>               | 0,75 µl |
| MgCl <sub>2</sub>                | 0,5 µl  |
| <i>Template</i>                  | 1,0 µl  |
| <i>Phusion HF</i> DNA-Polymerase | 0,5 µl  |

### PCR-Programm

|    |       |       |        |
|----|-------|-------|--------|
| 1. | 98 °C | 10 s  |        |
| 2. | 98 °C | PAUSE |        |
| 3. | 98 °C | 30 s  |        |
| 4. | 98 °C | 5 s   | } 31 x |
| 5. | 56 °C | 20s   |        |
| 6. | 72 °C | 20 s  |        |
| 7. | 72 °C | 5 min |        |
| 8. | 70 °C | 3 min |        |
| 9. | 4 °C  | PAUSE |        |



Im Anschluss an die PCR wurde jeder Ansatz mittels Agarose-Gelelektrophorese analysiert. Proben, bei denen keine Bande detektiert werden konnte, wurden mittels Bioanalyzer und ggf. mittels *NanoDrop* vermessen. Im Falle des Vorhandenseins einer Bande wurde die entsprechende RNA-Probe mit DNase I (Roche) behandelt. Anschließend wurde eine erneute Aufreinigung mit dem *RNeasy Micro Kit* (Qiagen) durchgeführt.

### 5.3 DNase-Behandlung (Roche)

Die Behandlung mit der DNase I (Roche) erfolgte direkt im Reaktionsgefäß. Auf diese Weise wurde der Großteil der verbleibenden genomischen DNA entfernt.

Protokoll (etabliertes Protokoll am Centrum für Biotechnologie, Universität Bielefeld):

- RNA-Probe mit RNase-freiem Wasser auf 100 µl auffüllen
- 10 µl 10 × Reaktionspuffer zugeben und gut mischen
- 10 µl DNase I (Roche) zugeben und sehr vorsichtig mischen
- 45 min bei 37 °C im Wasserbad inkubieren
- 3 min auf Eis abkühlen lassen

### 5.4 RNA-Aufreinigung mit DNase-Verdau

Nach dem DNase-Verdau wurde der Ansatz unter Verwendung des *RNeasy Micro Kits* aufgereinigt.

Protokoll (etabliertes Protokoll am Centrum für Biotechnologie, Universität Bielefeld):

- Probe mit 350 µl RLT-Puffer gut vermischen
- 250 µl Ethanol in einem 1,5-ml-Reaktionsgefäß vorlegen, RLT-Puffer-Probengemisch mit Ethanol vermischen und den gesamten Ansatz auf die Säule geben
- 30 s bei 20238 rcf bei RT zentrifugieren und Durchfluss verwerfen
- 500 µl RW1-Puffer auf die Säule geben, 30 s bei 20238 rcf bei RT zentrifugieren und Durchfluss verwerfen
- 70 µl RDD-Puffer mit 10 µl DNase I vermischen, auf die Säule geben und 15 min bei RT inkubieren
- 500 µl RW1 zugeben, 30 s bei 20238 rcf bei RT zentrifugieren, Durchfluss verwerfen und die Säule in ein neues Auffanggefäß überführen
- 500 µl RPE-Puffer zugeben, 30 s bei 20238 rcf bei RT zentrifugieren und Durchfluss verwerfen
- 500 µl RPE-Puffer zugeben, 30 s bei 20238 rcf bei RT zentrifugieren und Durchfluss verwerfen

- 1 min bei 20238 rcf und RT zentrifugieren um die Säule zu trocknen
- Säule in ein 1,5-ml Reaktionsgefäß überführen
- RNase-freies Wasser für Elution zugeben, 1 min bei 20238 rcf zentrifugieren und Säule verwerfen

Anschließend wurden die Proben erneut mittels PCR auf DNA-Rückstände getestet. Proben, bei denen keine Bande detektiert wurden, wurden mittels *Bioanalyzer* und ggf. mittels *NanoDrop* vermessen.

## 5.5 cDNA-Synthese und Amplifikation

Aufgrund des geringen Probenvolumens war die Ausbeute an Gesamt-RNA aus Proben der menschlichen Mikrobiota sehr gering. Daher muss nach der cDNA-Synthese eine Amplifikation der cDNA durchgeführt werden, um genügend Material für die Herstellung von Sequenzierbibliotheken zu erhalten.

In dieser Arbeit wurde das *Ovation RNA-Seq System V2 (NuGEN)* verwendet. Bei diesem System wird ausgehend von 500 pg bis 100 ng RNA die Synthese des ersten cDNA-Stranges unter Verwendung von *Random-Primern* und *Poly-T-Primern* durchgeführt. Die *Primer* besitzen am 3'-Ende einen DNA-Abschnitt, der an die komplementäre Region der RNA bindet, und am 5'-Ende einen RNA-Abschnitt, der zunächst einen Überhang bildet. Bei der Synthese des zweiten cDNA-Stranges wird dieser Abschnitt ebenfalls in cDNA umgewandelt und paart mit dem RNA-Abschnitt am 5'-Ende des ersten cDNA-Stranges, so dass an dieser Stelle ein RNA:DNA-Hybrid vorliegt. Anschließend wird die synthetisierte cDNA mit den *Agencourt RNAClean XP Purification Beads (Beckman Coulter, im Kit enthalten)* aufgereinigt. Das Prinzip dieser Aufreinigung liegt in der Bindung der DNA an die *Beads*. Aufgrund ihrer paramagnetischen Eigenschaften können die *Beads* und die gebundene DNA mithilfe eines Magneten von der Lösung getrennt werden. Nach einigen Waschschrritten mit 70%igem Ethanol kann die DNA in Puffer eluiert werden. Schließlich wird die cDNA mit einer isothermalen Methode amplifiziert. Dabei wird das Enzym RNase H verwendet, das die RNA in der RNA:DNA-Hybrid-Region abbaut. Auf diese Weise entsteht eine Stelle, an die erneut ein RNA/DNA-*Primer* binden und die DNA-Synthese mithilfe einer DNA-Polymerase initiieren kann. Während der DNA-Strang synthetisiert wird, wird der bereits vorhandene DNA-Strang verdrängt. Anschließend wird der RNA-Abschnitt im neu entstandenen RNA:DNA-Hybrid abgebaut und der Vorgang wiederholt sich mehrfach.

Das Kit wurde in dieser Arbeit grundsätzlich nach Herstellerangaben verwendet. Ausnahmen werden im Folgenden beschrieben:

- Die Deckeltemperatur des Thermocyclers betrug 99 °C
- Für Proben mit einer RNA-Konzentration < 100 pg/µl wurden jeweils die doppelten Volumina pro Ansatz verwendet
- Der Mastermix für jede Teilreaktion wurde entweder in einem 0,2-ml- oder in einem 1,5-ml-Reaktionsgefäß angesetzt
- Die Aufreinigung der cDNA mithilfe der magnetischen *Beads* wurde in 1,5-ml-Reaktionsgefäßen mit einem in der Universitätswerkstatt angefertigten Magneten durchgeführt. Dies gilt auch für das Entfernen der *Beads* nach der Amplifikation.

## 5.6 Säulenbasierte-Aufreinigung der cDNA

Nach der isothermalen Amplifikation werden die *Beads* mithilfe des Magneten aus der Lösung entfernt und der Ansatz wird mit dem *NucleoSpin Extract II Kit* aufgereinigt. Dabei handelt es sich ebenfalls um ein säulenbasiertes Verfahren, in dem die DNA an die Säulenmatrix bindet und nach mehreren Waschsritten in einem Puffer eluiert wird. In dieser Arbeit wurde das Protokoll für *PCR clean-up* nach Angaben des Herstellers angewendet. Die Elution erfolgte in 22 bis 50 µl *Elution Buffer*. Im Anschluss an die Aufreinigung wurden die Proben im *Bioanalyzer* vermessen und bei -20 °C gelagert.

Die weiteren Arbeitsschritte einschließlich der Sequenzierung wurden im Falle der Ohr-Probe von der *vertis Biotechnologie AG* durchgeführt. Dazu wurde die amplifizierte cDNA mittels Ultraschallbehandlung (30 s bei 4 °C) fragmentiert. Nach einem *End-Repair* wurden *TruSeq*-Sequenzieradapter an die cDNA ligiert und die entsprechenden Fragmente in 13 PCR-Zyklen amplifiziert. Die Sequenzierung wurde mit einem *Illumina-HiSeq-2000*-Gerät durchgeführt, wobei eine *Lane* einer *Flow Cell* für die Probe verwendet wurde. Die Sequenzierung erfolgte im *Single-End*-Modus mit einer Leseweite von 100 Basen. Im Falle der Achselhöhlenproben wurden alle weiteren Arbeitsschritte einschließlich der Sequenzierungen am CeBiTec durchgeführt und werden im Folgenden beschrieben.

## 5.7 cDNA-Fragmentierung und Aufreinigung

Bevor die Sequenzierbibliotheken erstellt werden können, muss die synthetisierte cDNA fragmentiert werden. Im Gegensatz zur Herstellerempfehlung im Protokoll für Sequenzierbibliotheken wird der *Bioruptor* (Diagenode) verwendet. Dabei handelt es sich um ein Ultraschallbad, in dem DNA-Proben bei einer definierten Temperatur im Verlauf mehrerer Ultraschall-Zyklen geschert werden.

In dieser Arbeit wurden 150 ng – 300 ng cDNA in der Fragmentierung eingesetzt. Die entsprechende Menge an cDNA-Lösung wurde in einem 1,5-ml-Reaktionsgefäß vorgelegt und

mit TE-Puffer auf 100 µl aufgefüllt. Der Wassertank der *Bioruptors* wurde auf 3 °C vorgekühlt und das Wasser wurde in das Ultraschall-Wasserbad geleitet. Die Geschwindigkeit des Wasser-Zulaufes wurde über ein Ventil so eingestellt, dass Zu- und Ablauf gleich schnell erfolgten und der Wasserspiegel konstant auf der vorgesehenen Höhe blieb. Durch den Wasseraustausch wurde die Wasser-Temperatur konstant gehalten. Anschließend wurden die Proben in den Rotor gespannt und die Ultraschall-Intensität auf die höchste Stufe gestellt. Für die Ultraschall-Behandlung wurde ein Intervall von 30 s und für die Pause ein Intervall von 2 min eingestellt. Die Fragmentierung erfolgte insgesamt für eine Dauer von 32 min und 30 s, was 13 Ultraschall-Zyklen entspricht. Anschließend wurden die Proben aus dem Ultraschallbad entfernt und mithilfe der *Agencourt RNAClean XP Purification Beads* (*Beckman Coulter*, im *Ovation Ultralow Library Kit* enthalten) nach den Angaben im Protokoll für die Herstellung von Sequenzierbibliotheken aufgereinigt (einschließlich der oben beschriebenen Änderungen). Danach wurden die Proben mit dem *Bioanalyzer* vermessen. Die fragmentierte und aufgereinigte cDNA wurde entweder direkt im Anschluss für die Herstellung von Sequenzierbibliotheken verwendet oder ü/N bei -20 °C gelagert.

## 5.8 Synthese und Aufreinigung von cDNA-Sequenzierbibliotheken

Für die Herstellung von Sequenzierbibliotheken müssen Sequenzierplattform-spezifische Adapter an die cDNA ligiert werden. Wenn mehrere Proben gleichzeitig sequenziert werden sollen, müssen außerdem spezifische Erkennungssequenzen, sogenannte *Barcodes*, angehängt werden um die einzelnen *Reads* anschließend den entsprechenden Proben zuordnen zu können.

Die Synthese der Sequenzierbibliotheken wurde in dieser Arbeit mit dem *Ovation Ultralow DR Multiplex System 1-8* durchgeführt. Ausgehend von 1 ng - 100 ng doppelsträngiger fragmentierter cDNA wird zunächst ein *End Repair* durchgeführt, in dem die in der Fragmentierung neu entstandenen cDNA-Enden geglättet werden. In dieser Arbeit wurden 50 ng fragmentierter cDNA für die Herstellung der Sequenzierbibliotheken verwendet. Anschließend werden die Illumina-Sequenzieradapter inklusive der *Barcode*-Sequenzen an die cDNA ligiert und die cDNA wird unter Verwendung der *Agencourt RNAClean XP Beads* aufgereinigt. Schließlich erfolgt eine Amplifikation der korrekt ligierten Fragmente mithilfe einer PCR in 15 Zyklen. Nach einer finalen Aufreinigung mit den magnetischen *Beads* wurde die amplifizierte cDNA-Bibliothek mithilfe des *Bioanalyzer* vermessen.

## 5.9 Sequenzierung

Die cDNA-Sequenzierbibliotheken wurden an der *Sequencing Core Facility* des Centrum für Biotechnologie der Universität Bielefeld sequenziert. Die Sequenzierung erfolgte mithilfe eines

*Illumina-MiSeq*-Gerätes unter Verwendung des *MiSeq Reagent Kit v2* nach Herstellerangaben. Dazu wurden pro Sequenzierung 6 – 8 Proben mit jeweils unterschiedlichen *Barcodes* in einem *Paired-End*-Lauf mit 500 Zyklen sequenziert. Somit betrug die Leseweite eines einzelnen *Reads* 250 Basen.

## 6 Bioinformatik

Für die taxonomische und funktionale Zuordnung der Sequenzier-Daten wurden zwei unterschiedliche Bioinformatik-*Pipelines* verwendet. Der Datensatz aus der Ohr-Probe wurde zum einen mit METRANS (Christina Ander, 2014) und zum anderen mit MG-RAST (Meyer et al., 2008) analysiert. Die Analyse der Datensätze aus den Achselhöhlenproben wurde nur mit MG-RAST durchgeführt. Da es sich bei letzteren um *Paired-End*-Datensätze handelte, mussten diese vor der Analyse mit dem SEQPREP-*Tool* (jstjohn, 2015) prozessiert werden.

### 6.1 Bearbeiten der Sequenzierdaten mit SEQPREP

Das SEQPREP-*Tool* wurde verwendet, um die die zusammengehörigen *Reads* der *Paired-End*-Sequenzierung der Sequenzierbibliotheken aus den Achselhohlen-Proben zu längeren *Reads* zu verbinden (jstjohn, 2015). Dazu müssen die beiden einzelnen *Reads* überlappen und/oder die Sequenzen müssen eine Adaptersequenz enthalten. *Read*-Paare, die nicht zusammengefügt werden konnten, wurden verworfen. Weiterhin wurden während der Prozessierung die Adaptersequenzen inklusive der *Barcode*-Sequenzen entfernt. Dazu wurden für jeden Datensatz die folgenden zwei Erkennungssequenzen aus den *Illumina-TruSeq*-Adaptersequenzen angegeben:

*Forward*-Adapter: AATGATACGGCGACCACCGA

*Reverse*-Adapter: GATCGGAAGAGCACACGTCT

Die Eingabe des Prozessierungsbefehls erfolgte über die Kommandozeile. Zusätzlich zu den Adaptersequenzen wurden folgende Eingaben gemacht: Dateiname für die *Forward*- und *Reverse-Read*-Datei (jeweils im *fastq*-Format), Dateinamen des *Forward*- und *Reverse-Output* (jeweils im *fastq*-Format), Dateiname für den Output der zusammengeführten *Reads* (im *fastq*-Format). Die *fastq*-Datei mit den zusammengeführten *Reads* diente als Grundlage für alle weiteren Analyseschritte.

### 6.2 Qualitätsanalyse der Sequenzier-*Reads* mit dem FASTX-*Toolkit*

Zur Berechnung der Qualität der Sequenzierdaten wurde das FASTX-*Toolkit* verwendet (Hannon Lab, 2009). Das *Tool FASTQ Information* berechnet über den gesamten Datensatz für jede

Position im Read die Anzahl der einzelnen Basen sowie den arithmetischen Mittelwert, den Median, die Quartilen und die Extremwerte der Qualität als *Phred Quality Score*. Der *Phred Quality Score* ist ein logarithmisches Maß für die Fehlerwahrscheinlichkeit der Zuordnung einer Base und wird für die Beschreibung der Qualität von Sequenzierdaten verwendet (Ewing et al., 1998). Unter Verwendung der arithmetischen Mittelwerte wurden mithilfe von *Microsoft Excel* Qualitätsplots der einzelnen Datensätze erstellt um die mittlere Qualität der einzelnen Sequenzier-Datensätze darzustellen.

### 6.3 Metatranskriptom-Analysen mit METRANS

Die Analyse der Sequenzier-Daten aus der Ohr-Probe wurde mithilfe der Analysepipeline METRANS durchgeführt, die in Zusammenarbeit mit Dr. Christina Ander und Prof. Dr. Jens Stoye vom Lehrstuhl für Genominformatik an der Technischen Fakultät der Universität Bielefeld entwickelt und zur Verfügung gestellt wurde. Dabei handelt es sich um ein Programm bestehend aus fünf Analyseschritten, das auf dem Hochleistungsrechencluster des CeBiTec betrieben werden kann. Die einzelnen Analyseschritte werden im Folgenden beschrieben.

1. *Sequenzvorbereitung / Quality Trimming*: Die Sequenzen werden auf ihre Qualität hin überprüft und entsprechend prozessiert. Dazu wird ein *Sliding Window Approach* angewandt, bei dem die mittlere Qualität innerhalb eines definierten Längenintervalls betrachtet wird. Das Intervall wird schrittweise um eine Base in Richtung des 3'-Endes verschoben und die Qualität wird in jedem Schritt ermittelt. Wenn die mittlere Qualität unter einen vom Benutzer definierten Schwellenwert fällt, wird die Sequenz so gekürzt, dass lediglich der Abschnitt vor dem Intervall erhalten bleibt. Da in den Sequenzier-Daten der Ohr-Probe eine Basenposition einen sehr niedrigen mittleren Qualitätswert aufweist, wird ein Intervall von 100 Basen gewählt, um ein Kürzen der Sequenzen an dieser Basenposition zu vermeiden. Als Qualitäts-Schwellenwert für das Verwerfen der Sequenzen wurde ein Wert von 35 (angegeben als *Phred Score*) gewählt. Somit wurden alle Reads mit einer mittleren Qualität < 35 aus dem Datensatz entfernt.
2. *Binning*: Um die Sequenzdatenmenge und die für die Analyse benötigte Rechenleistung und –zeit zu reduzieren, wird eine Klasseneinteilung der Sequenzen nach Ähnlichkeit vorgenommen. Dazu wird das implementierte Bioinformatik-Tool DNACLUSt verwendet (Ghodsí et al., 2011). Dabei werden Sequenzen mit einer Sequenzidentität von mindestens 98 % in eine gemeinsame Klasse (*bin*) eingeteilt und es wird eine repräsentative Sequenz gebildet, mit der alle weiteren Analyseschritte durchgeführt werden.

3. Herausfiltern von rRNA-Sequenzen: Der Großteil der RNA in bakteriellen Gemeinschaften besteht aus ribosomaler RNA. Da diese Sequenzen keine Enzym- oder Protein-Funktionen kodieren, werden sie entfernt und separat analysiert. Dies erfolgt durch den Vergleich der repräsentativen Sequenzen aus dem *Binning* mit den *SILVA*-Datenbanken der ribosomalen rRNA der kleinen Untereinheit (SSU), der ribosomalen rRNA der großen Untereinheit (LSU) und RFAM unter Verwendung des BLAST-Algorithmus (Altschul et al., 1990; Griffiths-Jones et al., 2003; Olsen et al., 1991). Dabei werden alle *Reads*, die mindestens 80 % Sequenzidentität zu einem Datenbankeintrag bei mindestens 80 % Sequenzüberlapp aufweisen, aus dem Datensatz herausgefiltert.
4. Klassifizierung der SSU-rRNA-Sequenzen mit dem RDP-CLASSIFIER: Die Sequenzen, die einen Treffer in der SSU-rRNA-Datenbank liefern, werden zur taxonomischen Beschreibung des untersuchten Datensatzes mithilfe des implementierten RDP-CLASSIFIER verwendet (Wang et al., 2007). In dieser Arbeit wurden die rRNA-basierten taxonomischen Darstellungen aus METRANS in *MS Excel* übertragen und taxonomische Profile wurden erstellt.
5. Funktionale und taxonomische Klassifizierung: Die Sequenzen, die nicht herausgefiltert werden, werden verwendet, um die von der RNA kodierten Funktionen und ihre taxonomische Herkunft zu beschreiben. Dazu wurde ein Vergleich der repräsentativen Sequenzen mit den funktionalen Datenbanken *KEGG*, *SwissProt* und *COG/EGGNOG* unter Verwendung des BLASTX-Algorithmus durchgeführt (Boeckmann et al., 2005; Kanehisa, 1997; Powell et al., 2012).

Für die taxonomische Klassifizierung der Funktionen werden die Daten auf die Taxonomie des *National Center for Biotechnology Information (NCBI)* gemappt (Sayers et al., 2010). Für die Ermittlung einer eindeutigen Klassifizierung wird der *Lowest Common Ancestor Approach* angewendet. Das bedeutet, dass für unterschiedliche taxonomische Zuordnungen einer Sequenz, beispielsweise im Falle einer stark konservierten Proteinfunktion, der kleinste gemeinsame Vorfahre für die taxonomische Zuordnung der Sequenz verwendet wird. Für die in dieser Arbeit durchgeführte Beschreibung der funktionalen Taxonomie wurden die entsprechenden Darstellungen aus METRANS in *MS Excel* übertragen und Profile wurden erstellt.

Die Identifizierung von Funktionen, die im Zusammenhang mit der Pathogenität von *Staphylococcus aureus* stehen, wurde auf Grundlage der Trefferliste gegen die *SwissProt*-Datenbank durchgeführt. Die Liste wurde von Dr. Christina Ander extrahiert und zur Verfügung gestellt. Als quantitatives Maß zur Beschreibung der Funktionen wurden die entsprechenden RPKM-Werte verwendet. Der RPKM-Wert errechnet sich nach der Formel

$$RPKM = \frac{\# Reads_{Funkt.} \times 1000}{Genlänge \times 1000000 \times \# Reads_{MT}}$$

mit:

$\# Reads_{Funkt.}$ : Anzahl der *Reads*, die einer bestimmten Funktion zugeordnet werden

$\# Reads_{MT}$ : Anzahl der *Reads* im entsprechenden Metatranskriptom-Datensatz

*Genlänge*: Länge der Referenzsequenz in Basen

und bietet eine Möglichkeit, die *Read*-Abundanz, die einer bestimmten Funktion zugeordnet wurde, sowohl auf die Länge der Referenzsequenz als auch auf die Größe des gesamten Metatranskriptom-Datensatzes zu normalisieren. Auf diese Weise können einzelne Annotationen innerhalb eines Datensatzes sowie aus mehreren Datensätzen in vergleichenden Beschreibungen zu einander in Bezug gesetzt werden.

## 6.4 Metatranskriptom-Analysen mit MG-Rast

Der Datensatz aus der Ohr-Probe und die Datensätze aus den Achselhöhlenproben wurden unter Verwendung der MG-RAST-Pipeline analysiert (Meyer et al., 2008). Die webbasierte Pipeline ermöglicht die funktionale und taxonomische Analyse von 16S-Amplikon-, Metagenom- und Metatranskriptom-Datensätzen sowie den Vergleich verschiedener Datensätze. Die aktuelle Version (v3) basiert im Wesentlichen darauf, dass vor einer Datenbankabfrage zur Bestimmung von Sequenzhomologien eine Genvorhersage durchgeführt wird, die es erlaubt zwischen RNA- und proteinkodierenden Genen zu unterscheiden und die Datenbankabfragen für jede Sequenz entsprechend zu wählen. Zur Benutzung der *Pipeline* wird ein Benutzer-*Account* angelegt. Verschiedene Metagenom- oder –transkriptom-Datensätze können dann in ein Projekt geladen und die Analyse kann gestartet werden. Die einzelnen Analyse-Schritte werden im Folgenden beschrieben:

- **Preprocessing**: Auf Grundlage der entsprechenden *fastq*-Dateien werden die Sequenzen getrimmt, um Bereiche mit einer schlechten Qualität zu entfernen. Entsprechend der Standardeinstellungen in MG-RAST wird der minimale Qualitätswert auf einen *Phred Quality Score* von 15 gesetzt. Die maximale Anzahl der Basen in einer getrimmten Sequenz, die diesen Wert unterschreitet, wird ebenfalls standardmäßig auf 5 gesetzt. Nach diesem Analyseschritt werden unterschiedliche Analysen für Protein- und RNA-kodierende Sequenzen durchgeführt.

### **Protein-Pipeline:**

- **Dereplication**: Eine bekannte Fehlerquelle bei der Analyse von Sequenzier-Daten steht im Zusammenhang mit sogenannten *Artificial Duplicate Reads* (ADRs). Da diese Artefakte des



Sequenzier-Vorganges bilden und nicht im Zusammenhang mit der Probe stehen, werden sie an dieser Stelle der Analysepipeline entfernt, indem Sequenzen, die in ihren ersten 20 Basen identisch zu anderen Sequenzen sind, entfernt werden. Dieser Analyseschritt ist in der *Pipeline* voreingestellt. Da die Herstellung der in dieser Arbeit analysierten cDNA jedoch auf einer isothermalen Amplifikation basiert, wurde diese Option für die Analyse der verhältnismäßig kleinen Datensätze aus den Achselhöhlenproben aus der Analyse-Pipeline entfernt. Da der Sequenzier-Datensatz der Ohr-Probe jedoch relativ groß ist, wird die *Dereplication* in diesem Fall beibehalten um eine weitere Datenreduktion und somit eine bessere Handhabung der Daten zu erwirken.

- DRISEE**: *DRISEE* steht für *Duplicate Read Inferred Sequencing Error Estimation* und stellt eine Methode zur Beurteilung von Sequenzierfehlern unter Verwendung der ARDs dar. Dieser Analyseschritt wird aufgrund der eliminierten *Dereplication* nicht für die Datensätze aus den Achselhöhlenproben durchgeführt, sondern nur für den Datensatz aus der Ohr-Probe.

- Screening**: Die Sequenzensatzdaten werden mithilfe des BOWTIE-Algorithmus auf Homologien zu einem vom Benutzer ausgewählten Modellorganismus durchsucht (Langmead et al., 2009). Sequenzen, die Homologien aufweisen, werden entfernt und nicht weiter analysiert. Dieser Analyseschritt eignet sich besonders für Metadatenätze von mikrobiellen Lebensgemeinschaften, die mit einem Wirtsorganismus assoziiert sind. Für die in dieser Arbeit analysierten Datensätze wurde das humane Genom (*H. sapiens*, NCBI v36) gemäß den Voreinstellungen in der *Pipeline* für diesen Schritt gewählt.

- Gene calling**: Zur Verringerung des Rechenaufwandes im Vergleich zur Homologie-basierten Genvorhersage wird eine *de-novo*-Genvorhersage unter Verwendung des FRAGGENESCAN-Tools durchgeführt (Rho et al., 2010).

- AA clustering**: Zur weiteren Verringerung der benötigten Rechenleistung in Form einer Datenreduktion werden die Proteinsequenzen zu Clustern mit mindestens 90 % Sequenzidentität zusammengefügt. Dazu wird das im QIME-Toolkit implementierte Programm verwendet (Caporaso et al., 2010; Edgar, 2010). Die relativen Häufigkeiten der Sequenzen bleiben für weitere Betrachtungen in der Auswertung jedoch erhalten.

- Protein identification**: Die repräsentative Sequenz eines jeden Clusters wird für die Suche nach Sequenzhomologien mithilfe des sBLAT-Tools verwendet (Kent, 2002). Als Referenzdatenbank wird die M5NR-Datenbank verwendet, die einen nicht redundanten Zusammenschluss der Datenbanken *GenBank*, *SEED*, *IMG*, *UniProt*, *KEGG* und *eggNOGs* darstellt (Benson et al., 2013;

Kanehisa, 1997; Magrane und Consortium, 2011; Markowitz et al., 2012; Overbeek et al., 2005; Wilke et al., 2012; Jensen et al., 2008). Die berechneten Homologien werden in der Auswertung ebenfalls für die Rekonstruktion der taxonomischen Zusammensetzung der Datensätze verwendet.

### **RNA-Pipeline:**

•**rRNA-detection:** Zur schnellen Identifizierung von rRNA-Sequenzen innerhalb der Sequenzier-Datensätze wird eine Datenbankabfrage unter Verwendung des BLAT-Algorithmus durchgeführt. Als Referenz dient eine von der *SILVA*-Datenbank abgeleitete Datenbank, in der die Einträge der *SILVA*-Datenbank mit einem Identitätsschwellenwert von 90 % geclustert vorliegen (Kent, 2002).

•**rRNA clustering:** Für eine Datenreduktion werden die im vorherigen Analyseschritt identifizierten Sequenzen zu Clustern mit mindestens 97 % Sequenzidentität zusammengefasst und die repräsentative Sequenz eines jeden Clusters wird für die Identifizierung verwendet.

•**rRNA identification:** Die Identifizierung der rRNA-Sequenzen erfolgt über eine Datenbankabfrage unter Verwendung des BLAT-Algorithmus. Als Referenzdatenbank dient die *M5RNA*-Datenbank, die (analog zur *M5NR*) eine nicht redundante Integration der Datenbanken *SILVA*, *Greengenes* und *RDP* darstellt (Cole et al., 2003; DeSantis et al., 2006; Pruesse et al., 2007).

•**Annotation mapping:** Aufgrund der Verwendung der *M5NR*-Datenbank als Referenz für die Identifizierung von Proteinsequenzen können die Ergebnisse mithilfe unterschiedlicher funktionaler Hierarchien, wie *SEED Subsystems*, *IMG terms*, *COG* und *eggNOGs* dargestellt werden. Der Benutzer kann so zwischen verschiedenen Annotationssystemen wählen.

•**Abundance profiles:** Alle Sequenzähnlichkeiten auf funktionaler und auf taxonomischer Ebene werden mit den entsprechenden Häufigkeiten dargestellt und der Benutzer kann zwischen verschiedenen Ansichten mit unterschiedlichen Werten für die folgenden Parameter wählen: *E-value*, Identität und minimale *Alignment*-Länge. Die entsprechenden Daten können als Balkendiagramm, Baum oder Tabelle dargestellt werden. Vergleichende Darstellungen sind in Form einer *Heatmap* oder einer *Principal Coordinates Analyse* (PCoA) darstellbar. Funktionale Abundanzen sind zusätzlich noch mit dem *KEGG Mapper* darstellbar. Zusätzlich kann ein Datensatz mit der Funktion *Recruitment Plot* auf ein vom Benutzer definiertes Genom gemappt werden.

Für jeden analysierten Datensatz wird in MG-RAST eine Übersichtsseite, die *Metagenome Overview*, erstellt, auf der die Datenanalyse zusammenfassend dargestellt ist. Zusätzlich zur Darstellung von Metadaten und der Analysestatistik für jeden Datensatz, werden die Verteilungen der *Read*-Abundanzen in einem *Analysis Flowchart* dargestellt, so dass nachvollziehbar ist, welche Datenanteile rRNA oder bestimmten Funktionen zugeordnet werden konnten und welche Datenanteile im Rahmen der Qualitätsprüfung entfernt wurden. Weiterhin werden Fehlerabschätzungen, Nukleotid-Verteilungen und funktionale und taxonomische Verteilungen zusammenfassend dargestellt. Besonders hervorzuheben ist die *Source Hits Distribution*, in der die Datenanteile, die den jeweiligen Annotationssystemen zugeordnet werden konnten, farbkodiert nach den entsprechenden *E-Value*-Exponenten abgebildet werden. Diese Darstellung wird in dieser Arbeit verwendet, um einen geeigneten *E-Value*-Schwellenwert für die Datenanalyse auszuwählen.

Eine Besonderheit stellt die sogenannte *Workbench* innerhalb der Analyse-Optionen dar. Der Benutzer kann Teil-Datensätze, beispielsweise bestimmte funktionale oder taxonomische Annotationen, in den *Workbench*-Speicher laden und herunterladen oder weiterführend analysieren. Dies ermöglicht ebenfalls die Darstellung von Teildatensätzen in unterschiedlichen Annotationssystemen, die Darstellung einer funktionalen Taxonomie oder umgekehrt die funktionale Annotation bestimmter taxonomischer Gruppen.

Da die *MG-RAST-Pipeline* auf die Analyse von bakteriellen Metagenomen ausgelegt ist, sind nicht alle analytischen *Tools* mit allen Darstellungsoptionen für die Auswertung von Metatranskriptom-Datensätzen sinnvoll nutzbar. Im Folgenden werden die in dieser Arbeit durchgeführten Datenanalyse-Schritte in MG-RAST näher beschrieben.

#### **6.4.1 Taxonomische Analyse von Metatranskriptomik-Datensätzen in MG-RAST**

Die taxonomische Analyse von Sequenzier-Datensätzen kann in MG-RAST grundsätzlich auf drei unterschiedliche Weisen erfolgen. Der Benutzer kann zwischen den folgenden drei Klassifikationstypen für die taxonomische Zuordnung der Sequenzier-Daten wählen:

- *Lowest Common Ancestor*: Analog zum in METRANS verwendeten Ansatz werden hier Sequenzen mit unterschiedlichen taxonomischen Zuordnung dem kleinsten gemeinsamen Vorfahren zugeordnet.
- *Best Hit Classification*: Wenn ein *Read* unterschiedliche taxonomische Treffer liefert, wird der beste Treffer gewertet. Ist die Zuordnung nicht eindeutig, wird die Sequenz allen Treffern zugeordnet. Diese Methode resultiert vor allem im Falle von stark konservierten Sequenzen in einer scheinbaren Erhöhung der Datenmenge.

- *Representative Hit Classification*: Wenn ein *Read* Treffer unterschiedlicher taxonomischer Zuordnungen in gleicher Qualität liefert, wird nur eine Zuordnung ausgewählt. Die Auswahl erfolgt zufällig, beispielsweise nach der alphabetischen oder numerischen Reihenfolge des Auftretens in der Liste der Annotationen. Bei dieser Methode kann es zu falschen taxonomischen Zuordnungen konservierter Sequenzen kommen. Ein Vorteil liegt jedoch darin, dass die Anzahl der Annotationen gleich der Anzahl der *Reads* ist.

Zum einen kann der Benutzer eine rRNA-basierte Taxonomie auf Grundlage der *M5RNA*-Datenbank oder auf Grundlage der einzelnen rRNA-Datenbanken *SSU*, *LSU*, *RDP* und *Greengenes* erstellen. Zum anderen ist analog die Erstellung einer funktionalen Taxonomie auf Grundlage der *M5NR*-Datenbank oder der einzelnen Datenbanken *IMG*, *TrEMBL*, *PATRIC*, *SwissProt*, *GenBank*, *SEED*, *RefSeq*, *eggNOG* und *KEGG* möglich.

In dieser Arbeit wurde rRNA-basierte taxonomische Zuordnung auf Grundlage der *M5RNA*-Datenbank mit einem *E-value*-Schwellenwert von  $10^{-30}$ , einem Schwellenwert für die Sequenzidentität von 80 % und einem Schwellenwert für den Sequenzüberlapp von 80 Basen für den Datensatz aus dem Ohr und 110 Basen für die Datensätze aus den Achselhöhlenproben durchgeführt. Die funktionale Taxonomie wurde auf Basis der *M5NR*-Datenbank bei einem *E-value*-Schwellenwert von  $10^{-10}$ , einem Schwellenwert für die Sequenzidentität von 80 % und einem Schwellenwert für den Sequenzüberlapp von 26 Aminosäuren für den Datensatz aus dem Ohr und 36 Aminosäuren für die Datensätze aus den Achselhöhlenproben erstellt. Dazu wurde eine taxonomische Zuordnung der jeweiligen Datensätze in MG-RAST unter Verwendung der *Representative Hit Classification* durchgeführt und die Datenausgabe in Tabellenform gewählt. Die Tabelle wurde anschließend nach den jeweiligen taxonomischen Ebenen gruppiert und über die Download-Funktion heruntergeladen. Die Erstellung der taxonomischen Profile und die Illustrierung wurde mit MS EXCEL durchgeführt.

Die Berechnung der Diversität und Äquität erfolgte auf Grundlage der *Representative Hit Classification* in der *M5RNA*-Datenbank auf Spezies-Ebene mithilfe des Shannon Diversitätsindex nach den folgenden Formeln:

$$H' = - \sum p_i * \ln(p_i)$$

und

$$E_{H'} = \frac{H'}{\ln(S)}$$

$H'$ : Shannon Diversitätsindex

$p_i$ : relative Abundanz der i-ten Spezies

$E_{H'}$ : Shannon Äquitätsindex

$S$ : Anzahl der Spezies-Annotationen

Die Korrelation einzelner Spezies mit dem *Malodor Score* wurde wie folgt durchgeführt:

- Die Datensätze der *Representative Hit Classification* auf Grundlage der *M5RNA*-Datenbank auf Spezies-Ebene wurden so gestaltet, dass jede Spezies-Annotation eines jeden Datensatzes genau einmal in eine Liste geschrieben und die entsprechenden *Read*-Abundanzen der einzelnen Datensätze zugeordnet wurden.
- Anschließend erfolgte eine Normalisierung der *Read*-Abundanzen nach folgender Formel:

$$R_{S_{ij},n} = \frac{(R_{S_{ij}} * 10^{15})}{(R_j * K_j)}$$

$R_{S_{ij},n}$ : Normalisierte *Read*-Abundanz der i-ten Spezies des Metatranskriptoms j

$R_{S_{ij}}$ : *Read*-Abundanz der i-ten Spezies des Metatranskriptoms j

$R_j$ : Gesamtzahl der *Reads* des Metatranskriptoms j

$K_j$ : Keimzahl der Probe j [*KBE/ml*]

- Die normalisierten *Read*-Abundanzen wurden anschließend durch Logarithmierung zur Basis 2 transformiert.
- Die normalisierten und transformierten *Read*-Abundanzen wurden schließlich mit dem *Malodor Score* korreliert.

#### 6.4.2 Funktionale Analyse von Metatranskriptomik-Datensätzen in MG-RAST

Für die funktionale Analyse der Metatranskriptomik-Datensätze wurden in Abhängigkeit der Datensätze und der zu beantwortenden biologischen Fragestellungen jeweils unterschiedliche Ansätze gewählt, die im Folgenden näher beschrieben werden.

Erstellung von COG-Profilen

Zur funktionalen Analyse eines gesamten Metatranskriptoms wurde für den jeweiligen Datensatz ein funktionales Profil auf Grundlage der COG-Klasseneinteilung erstellt. Die dazu verwendeten Schwellenwerte der Analyseparameter sind in Tabelle 8 dargestellt. Mit der Analyse-Funktion von MG-Rast wurde zunächst die Einstellung *Hierarchical Classification* gewählt, die entsprechenden Analyseparameter eingestellt und eine Tabelle erstellt. Die Tabelle wurde anschließend nach *level 2* gruppiert und über die *Download*-Funktion heruntergeladen. Die COG-Profile wurden anschließend in MS EXCEL erstellt und illustriert.

**Tabelle 8: Schwellenwerte der Analyseparameter für die Erstellung von COG-Profilen in MG-RAST.**

| Metatranskriptom | <i>E-value</i> -Exponent | Sequenzidentität [%] | Sequenzüberlapp [aa] |
|------------------|--------------------------|----------------------|----------------------|
| Ohr              | -5                       | 60                   | 15                   |
| Achselhöhle      | -30                      | 80                   | 36                   |

Funktionale Analyse von *Staphylococcus aureus* im Metatranskriptom-Datensatz der Ohr-Probe

Für die funktionale Analyse der Sequenzen, die dem Pathogen *Staphylococcus aureus* zugeordnet wurden, wurde zunächst der Metatranskriptomik-Datensatz aus der Ohr-Probe taxonomisch auf Grundlage der *M5NR*-Datenbank unter Verwendung der *Representative Hit Classification* und der Standardeinstellungen für die Analyseparameter klassifiziert. Die Ausgabe wurde in Tabellenform gewählt und die Tabelle nach der Spezies-Ebene gruppiert. Danach wurde der *Staphylococcus aureus* zugeordnete Teil-Datensatz in die *Workbench* geladen. Anschließend wurde mit dem Teil-Datensatz eine funktionale Analyse auf Grundlage der *SwissProt*-Annotationen bei den Standardeinstellungen für die Analyseparameter durchgeführt. Für die Suche nach den Annotationen, die im Zusammenhang mit der Pathogenität von *S. aureus* stehen, wurde die Filterfunktion der Tabelle genutzt. Als quantitatives Maß zur Beschreibung der jeweiligen Funktion wurde die *Read*-Abundanz gewählt.

Identifizierung von Funktionen im Zusammenhang mit Geruchsentstehung aus *Corynebacterium striatum* in den Metatranskriptomik-Datensätzen der Achselhöhlen-Proben

Für die funktionale Beschreibung N<sup>α</sup>-Acylglutamin-Aminoacylase, Cystathionin-β-Lyase und *Thiol Precursor Dipeptidase* aus *Corynebacterium striatum* wurde zunächst eine taxonomische Klassifizierung der Metatranskriptomik-Datensätze nach der *Representative Hit Classification* auf Grundlage der *M5NR*-Datenbank unter Verwendung der Standardeinstellungen für die Analyseparameter durchgeführt. Die Datenausgabe wurde in Tabellenform gewählt und die

Tabelle wurde nach der Spezies-Ebene gruppiert und nach *C. striatum* gefiltert. Die jeweiligen Teil-Datensätze der einzelnen Metatranskriptomte wurden anschließend in die *Workbench* geladen und funktional auf Grundlage des *GenBank*-Annotationssystems unter Verwendung der Standardeinstellungen für die Analyseparameter analysiert. Die Datenausgabe wurde in Tabellenform gewählt und die Filterfunktion der Tabelle wurde genutzt um nach den funktionalen Annotationen N<sup>α</sup>-Acylglutamin-Aminoacylase, Cystathionin-β-Lyase und *Thiol Precursor Dipeptidase* zu suchen. Als quantitatives Maß zur Beschreibung der jeweiligen Funktion wurde die *Read*-Abundanz gewählt. Als qualitatives Maß diente die Sequenzidentität.

#### Analyse von Aminoacylase-Funktionen in den Metatranskriptomik-Datensätzen der Achselhöhlen-Proben

Für die Beschreibung der Aminoacylase-Sequenzen in den Metatranskriptomik-Datensätzen der Achselhöhlen-Proben in MG-RAST wurden zwei unterschiedliche analytische Ansätze gewählt, die im Folgenden beschrieben werden.

1. Analyse der COG1473-Annotationen: Zunächst wurde eine funktionale Klassifizierung der jeweiligen Metatranskriptomte auf Grundlage des COG-Annotationssystems unter Verwendung der Standardeinstellungen für die Analyseparameter erstellt. Die Datenausgabe wurde in Tabellenform gewählt und die Tabelle wurde mithilfe der Filterfunktion nach der Klasse COG1473 durchsucht. Die entsprechenden Teil-Datensätze wurden anschließend in die *Workbench* geladen und mithilfe der *Best Hit Classification* unter Verwendung der Standardparameter die jeweiligen *Read*-Abundanzen klassifiziert und auf Genus-Ebene dargestellt. Die entsprechenden Tabellen wurden heruntergeladen und mittels MS EXCEL visualisiert. Die Korrelation zwischen den Anzahlen der Annotationen zur Funktion COG1473 und dem *Malodor Score* wurde ebenfalls in MS EXCEL durchgeführt. Die Normalisierung und Transformation der Daten erfolgte analog zur in Kapitel 6.4.1 beschriebenen Methode.

2. Analyse der *GenBank*-Annotationen: Zunächst wurde eine funktionale Klassifizierung der Metatranskriptomte auf Grundlage des *GenBank*-Annotationssystems unter Verwendung der Standardeinstellungen für die Analyseparameter erstellt. Die Datenausgabe wurde in Tabellenform gewählt und die Tabelle wurde mithilfe der Filterfunktion auf Annotationen zur Aminoacylase (Suchbegriff „aminoacylase“) durchsucht. Im Falle mehrerer unterschiedlicher Annotationen der gesuchten Funktion wurde jede Annotation separat betrachtet. Die jeweiligen Teildatensätze wurden ausgewählt und in die *Workbench* geladen. Anschließend wurden zwei separate Analyseschritte durchgeführt:

- die Datensätze aus der *Workbench* wurden mittels *Best Hit Classification* auf Grundlage der *M5NR*-Datenbank unter Verwendung der Standardeinstellungen klassifiziert und auf Spezies-Ebene analysiert. Die Visualisierung erfolgte mithilfe von MS EXCEL.

- die Teildatensätze zur Annotation N<sup>α</sup>-Acylglutamin-Aminoacylase wurden im *fasta*-Format heruntergeladen und händisch mithilfe des BLASTX-Algorithmus in der *Online*-Version der *GenBank*-Datenbank (Stand 08.05.2016) abgefragt. Für den jeweils ersten Treffer wurden die folgenden Parameter notiert: *GenBank-ID*, Organismus und funktionale Annotation der Referenzsequenz, Sequenzidentität, *Reading Frame*, Proteinsequenz der abgefragten Sequenz und der Referenzsequenz. Darüber hinaus wurde die Sequenzidentität der abgefragten Sequenz zu N-AGA aus *C. striatum* notiert. Weiterhin wurde die Zuordnung der abgefragten Sequenz in der *CDD*-Datenbank notiert. Die übersetzten Aminosäure-Sequenzen sowie die Referenzsequenzen wurden mithilfe des CLUSTALOMEGA-Tools in einem *Multiple Sequence Alignment* dargestellt und mit der GENEDOC-Software visualisiert. Darüber hinaus wurde mit dem selben Tool ein phylogenetischer Baum erstellt. Chimäre Sequenzen wurden nicht in die Analyse aufgenommen.

#### Analyse von Cystathionin-β-Lyase-Funktionen in den Metatranskriptomik-Datensätzen der Achselhöhlen-Proben

Für die Beschreibung der Cystathionin-β-Lyase-Sequenzen in den Metatranskriptomik-Datensätzen der Achselhöhlen-Proben in MG-RAST wurden zwei unterschiedliche analytische Ansätze gewählt, die im Folgenden beschrieben werden.

1. Analyse der *GenBank*-Annotationen: Zunächst wurde eine funktionale Klassifizierung der jeweiligen Metatranskriptome auf Grundlage des *GenBank*-Annotationssystems unter Verwendung der Standardeinstellungen für die Analyseparameter erstellt. Die Datenausgabe wurde in Tabellenform gewählt und die Tabelle wurde mithilfe der Filterfunktion auf Annotationen der Cystathionin-β-Lyase und Cystein-S-Konjugat-β-Lyase durchsucht. Im Falle mehrerer unterschiedlicher Schreibweisen der gesuchten Annotation wurden alle Annotationen ausgewählt und in die *Workbench* geladen. Dann wurden die Subdatensätze in der *Workbench* mittels *Best Hit Classification* klassifiziert. Die Darstellung der taxonomischen Profile erfolgte in MS EXCEL. Darüber hinaus wurden in der *Workbench* die annotierten Treffer geladen und der Verlinkung auf die *GenBank-ID* gefolgt. Ausgehend von dem entsprechenden Eintrag in der *Online*-Version der *GenBank*-Datenbank wurde eine Suche in der *GenBank*-Datenbank unter Verwendung des BLASTP-Algorithmus bei Standardeinstellungen durchgeführt. Auf der Ergebnisseite wurde der Verlinkung zum Treffer in der *CDD*-Datenbank gefolgt und die nicht-spezifischen Treffer *Maly*, *MetC*, *CysK* und *Met\_gamma\_lyase* registriert. Als quantitatives Maß



zur Beschreibung der jeweiligen funktionalen Zuordnung wurde die *Read*-Abundanz in MG-RAST gewählt. Als qualitatives Maß diente die Sequenzidentität. Weiterhin wurde die taxonomische Zuordnung des jeweiligen BLASTP-Treffers registriert.

2. Analyse der COG1168-Annotationen: Zunächst wurde eine funktionale Klassifizierung der jeweiligen Metatranskriptome auf Grundlage des *COG*-Annotationssystems unter Verwendung der Standardeinstellungen für die Analyseparameter erstellt. Die Datenausgabe wurde in Tabellenform gewählt und die Tabelle wurde mithilfe der Filterfunktion nach der Klasse COG1168 durchsucht. Die entsprechenden Teil-Datensätze wurden anschließend in die *Workbench* geladen und über die *Download*-Funktion im *fasta*-Dateiformat heruntergeladen. Anschließend wurden mehrfach vorhandene Sequenzen aus den Listen entfernt und die einzelnen Sequenzen in der *GenBank*-Datenbank (lokal auf dem CeBiTec Rechencluster, Version vom März 2014) unter Verwendung des BLASTX-Algorithmus gesucht. Die entsprechenden PERL-Skripte wurden von Daniel Wibberg am CeBiTec zur Verfügung gestellt und ausgeführt. Die Datenausgabe erfolgte in Form einer *Alignment*-Datei. Für jede Sequenz wurden die folgenden Daten registriert: übersetzte Aminosäure-Sequenz aus dem *Alignment*, *GenBank*-ID, taxonomische Zuordnung, funktionale Zuordnung, *E-value*-Exponent und Position des *Alignments* auf der Treffer-Sequenz. Die Erstellung eines *Multiple-Sequence-Alignments* der jeweiligen Referenzsequenzen aus der *GenBank*-Datenbank wurde mit dem *ClustalW2*-Tool durchgeführt und die Visualisierung erfolgte unter Verwendung der *GeneDoc*-Software.

#### Analyse von Dipeptidase-Funktionen in den Metatranskriptomik-Datensätzen aus den Achselhöhlenproben

Die Analyse der Dipeptidase-Sequenzen erfolgte weitestgehend analog zur Analyse der Cystathionin- $\beta$ -Lyase (s. oben). Für die Analyse auf Grundlage der *GenBank*-Annotationen wurde die Annotation „thiol precursor dipeptidase“ gewählt. Für die Analyse der *COG*-Annotationen wurde die *COG*-Klasse COG2195 ausgewählt.

#### Analyse anderer im Zusammenhang mit Körpergeruchsentstehung beschriebener Funktionen in den Metatranskriptomik-Datensätzen der Achselhöhle

Die anderen im Zusammenhang mit Körpergeruchsentstehung in der Fachliteratur beschriebenen Funktionen wurden auf Basis der EC-Nomenklatur untersucht. Zunächst wurde eine funktionale Klassifizierung der jeweiligen Metatranskriptome auf Grundlage des *Subsystems*-Annotationssystems unter Verwendung der Standardeinstellungen für die Analyseparameter erstellt. Die Datenausgabe wurde in Tabellenform gewählt und die Tabelle wurde mithilfe der Filterfunktion auf die entsprechenden Annotationen und EC-Nummern durchsucht. Im Falle einer mehrfachen Zuordnung einer Annotation zu unterschiedlichen

übergeordneten Funktionskategorien wurden nur Annotationen aus einer Zuordnung ausgewählt und in die *Workbench* geladen. Anschließend wurden die Teil-Datensätze mithilfe der *Best Hit Classification* unter Verwendung der Standardeinstellungen für die Analyseparameter klassifiziert und in Tabellenform ausgegeben. Die Tabelle wurde nach der Genus-Ebene gruppiert, über die Download-Funktion heruntergeladen und mithilfe von MS EXCEL wurden Profile der funktionalen Taxonomie erstellt und illustriert.

## 6.5 Taxonomische Analyse der Metatranskriptomik-Datensätze auf Spezies-Ebene mit der *GenBank*-Datenbank

Zusätzlich zur taxonomischen Analyse der Metatranskriptomik-Datensätze aus den Achselhöhlen-Proben mit MG-RAST wurde eine 16S-rRNA-basierte Analyse der Daten auf Spezies-Ebene auf Grundlage der *GenBank*-Datenbank durchgeführt. Dazu wurde zunächst in MG-RAST eine taxonomische Klassifizierung der jeweiligen Datensätze auf Grundlage der SSU-Datenbank unter Verwendung der *Representative Hit Classification* und den Standardeinstellungen für die Analyseparameter durchgeführt. Die Datenausgabe wurde in Tabellenform gewählt und die Tabelle wurde nach der Domänen-Ebene gruppiert. Die Teil-Datensätze, die den *Bacteria* zugeordnet wurden, wurden in die *Workbench* geladen und über die *Download*-Funktion heruntergeladen. Anschließend wurden die Datensätze mit einer Sequenzidentität von mindestens 97 % geclustert und die repräsentative Sequenz eines jeden Clusters wurde in der *GenBank*-Datenbank (lokal auf dem CeBiTec Rechencluster, Version vom März 2014) unter Verwendung des BLASTN-Algorithmus abgefragt. Für jede Sequenz wurde die *GenBank*-ID sowie die taxonomische Zuordnung des ersten Treffers erfasst und mit der entsprechenden Anzahl an Sequenzier-*Reads* pro Cluster in einer Tab-separierten Text-Datei ausgegeben. Die entsprechenden PERL-Skripte wurden von Daniel Wibberg am CeBiTec zur Verfügung gestellt und ausgeführt. Die Erstellung der taxonomischen Profile und die Illustrierung erfolgten mithilfe von MS EXCEL.

## 6.6 Durchführung von Hauptkomponentenanalysen (PCAs)

Eine Hauptkomponentenanalyse (*Principal Component Analysis*, PCA) ist eine statistische Methode und wird häufig verwendet, um komplexe Datensätze vergleichend darzustellen. Die Datensätze müssen dabei so aufgebaut sein, dass jeweils einer Messgröße (Eigenschaft, *Identifizier*) ein Messwert (Abundanz, Intensität) zugeordnet wird. Eine PCA erfüllt im Wesentlichen zwei Funktionen:

1. Datentransformation: Datensätze werden mathematisch so transformiert, dass die größtmögliche Varianz zwischen ihnen entsteht. Anschließend werden Hauptkomponenten

(PCs) berechnet. Eine Hauptkomponente setzt sich aus unterschiedlichen Messgrößen in unterschiedlichen Proportionen (je nach Datentransformation) zusammen. Die PC1 beschreibt dabei die größte Varianz über alle Datensätze, die PC2 die nächstgrößere, usw. Die Anzahl der benötigten Hauptkomponenten ist abhängig vom Datensatz und wird individuell bestimmt.

2. Vergleichende Darstellung: Jeder Datensatz wird anschließend mithilfe von Eigenvektoren als Punkt in einem Koordinatensystem dargestellt. Die Koordinaten werden dabei von den jeweiligen PCs gebildet. In dieser Darstellung beschreibt der Abstand zwischen den Datensätzen die jeweilige Varianz, d. h. Datensätze, die nahe beieinander liegen, sind sich ähnlicher als Datensätze, die weiter voneinander entfernt sind. Zur Beschreibung der jeweiligen PCs dient der sogenannte *Loading Plot*. Dieser stellt die Gewichtung einzelner Messgrößen innerhalb des Koordinatensystems aus PCs dar und ermöglicht somit ebenfalls eine inhaltliche Bewertung der Varianzen in dem *PCA-Plot*.

In dieser Arbeit wurde eine PCA zur vergleichenden Beschreibung der Metatranskriptomik-Datensätze aus den Achselhöhlen-Proben auf taxonomischer und auf funktionaler Ebene durchgeführt. Dazu wurden die im Rahmen der rRNA-basierten taxonomischen Beschreibung heruntergeladenen Teil-Datensätze verwendet (s. Abschnitte 6.4.1). Für die Erstellung der taxonomischen PCA wurden alle Genera aller Datensätze betrachtet, die mindestens 0,3 % der gesamten *Reads* eines Datensatzes ausmachen. Die PCA wurde anschließend mithilfe von ORIGINPRO 9.1G unter Verwendung der voreingestellten Parameter für eine Kovarianzmatrix erstellt. Zur vergleichenden Beschreibung der Metatranskriptomik-Datensätze der Achselhöhlen-Proben auf funktionaler Ebene wurde ebenfalls eine PCA auf Basis der COG-Klasseneinteilung durchgeführt. Dazu wurden die entsprechenden Teil-Datensätze wie in Abschnitt 6.4.2 beschrieben, jedoch unter Verwendung der Standardeinstellungen für die Analyseparameter, heruntergeladen. Als Messgrößen für die Eingabe in ORIGINPRO 9.1G wurden die 13 COG-Klassen mit den höchsten *Read*-Abundanzen gewählt. Die Erstellung der PCA erfolgte unter Verwendung der voreingestellten Parameter für eine Kovarianzmatrix.

## 6.7 Analyse von Funktionen mit Korrelation zum Körpergeruch

Zur Identifizierung von Funktionen, die mit dem *Malodor Score* korrelieren, wurde in MG-RAST die KO-Ontologie und die *Heatmap*-Funktion verwendet. Da die Datensätze der drei Proben 13n, 18n und 19n nach Deo-Anwendung keinen sinnvollen Rückschluss auf Korrelationen zum Körpergeruch erwarten lassen, wurden diese drei Datensätze im Rahmen dieser Untersuchung nicht weiterführend analysiert. Zunächst wurde unter dem Menüpunkt *Functional Abundance* der Modus *Hierarchical Classification* ausgewählt. Anschließend wurde unter Auswahl aller 14

Metatranskriptomik-Datensätze und der Standardeinstellungen für die Analyseparameter das Annotationssystem *KO* ausgewählt und eine *Heatmap* erstellt. Die Gruppierung der *Heatmap* erfolgte nach der Funktion unter Verwendung der Rohdaten. Für alle übrigen Parameter wurden die Voreinstellungen in MG-RAST übernommen. Anschließend wurden die Daten über die *Download*-Funktion heruntergeladen und in MS EXCEL wie folgt weiter bearbeitet:

- Für jede funktionale Annotation wurde die Summe über die *Read*-Abundanzen über alle 14 Datensätze gebildet. Alle Annotationen, die eine Summe von  $\leq 14$  *Read*-Abundanzen aufwiesen, wurden nicht weiterführend analysiert.
- Für alle übrigen Annotationen wurden die Werte der *Read*-Abundanzen nach der folgenden Formel normalisiert:

$$R_{i,j,n} = \frac{R_{i,j}}{R_j} \times 10^7$$

mit:

$R_{i,j,n}$ : normalisierte *Read*-Abundanz der funktionalen Annotation *i* des Datensatzes *j*

$R_{i,j}$ : *Read*-Abundanz der funktionalen Annotation *i* des Datensatzes *j*

$R_j$ : Gesamtzahl der *Reads* des Datensatzes *j*

- Alle normalisierten *Read*-Abundanzen wurden anschließend logarithmisch zur Basis 2 transformiert. Da einige Einzelwerte aufgrund einer sehr geringen *Read*-Abundanz einen normalisierten Wert nahe 0 lieferten, wurden diese gleich Null gesetzt.
- Für alle funktionalen Annotationen wurden weiterhin das Minimum und das Maximum über alle 11 Datensätze bestimmt. Anschließend wurde der Faktor von Maximum zu Minimum bestimmt. Alle funktionalen Annotationen mit einem Faktor  $\leq 2$  wurden nicht weiterführend analysiert.
- Schließlich wurde für jede funktionale Annotation eine Korrelation zwischen dem *Malodor Score* und dem transformierten Wert der normalisierten *Read*-Abundanzen berechnet. Korrelationen von  $\geq 0,5$  wurden als positiv, Korrelationen von  $\leq -0,5$  als negativ und alle übrigen Korrelationen als neutral bewertet.
- Auf Grundlage der entsprechenden *KO*-Nummern wurden unter Verwendung der *Search&Color-Pathway*-Funktion des KEGG-MAPPER die entsprechenden Enzymfunktionen dargestellt. Dabei wurden positiv korrelierte Funktionen grün, negativ korrelierte Funktionen rot und neutrale Funktionen gelb oder schwarz eingefärbt. Es wurde sowohl die gesamte metabolische Karte, als auch Karten einzelner im Zusammenhang mit Körpergeruchsentstehung relevanter Stoffwechselwege dargestellt und analysiert.

- Schließlich wurden ausgewählte funktionale Annotationen in MG-RAST klassifiziert. Dazu wurden die Teil-Datensätze mithilfe der *Best Hit Classification* unter Verwendung der Standardeinstellungen für die Analyseparameter klassifiziert und in Tabellenform ausgegeben. Die Tabelle wurde nach der Genus-Ebene gruppiert, über die *Download*-Funktion heruntergeladen und mithilfe von MS EXCEL wurden Profile der funktionalen Taxonomie erstellt und illustriert.

## IV. Ergebnisse

Die Metatranskriptomik ermöglicht die kultivierungsunabhängige ganzheitliche Beschreibung und Analyse des Transkriptoms einer mikrobiellen Gemeinschaft an einem bestimmten Ort zu einer bestimmten Zeit. Im Rahmen dieser Arbeit wurden Metatranskriptom-Analysen von Proben des menschlichen Ohrs und der menschlichen Achselhöhle durchgeführt. Für die technische Durchführung wurde jeweils ein Laborverfahren etabliert, das den Umgang mit sehr kleinen Probenmengen und die Herstellung von cDNA-Sequenzierbibliotheken auf Grundlage von cDNA-Synthese und -Amplifikation ermöglichte.

Die anschließende Datenanalyse erfolgte unter Verwendung der beiden Bioinformatik-*Pipelines* MG-RAST und im Falle der Ohrprobe, METRANS. Unabhängig von der untersuchten Probe und der verwendeten *Analyse-Pipeline* waren die folgenden Aspekte für die Beschreibung des Metatranskriptoms zentral:

- taxonomische Klassifizierung des Metatranskriptoms auf Grundlage der Sequenzen mit Homologie zu rRNA-Sequenzen (rRNA-basierte Taxonomie) und auf Grundlage der Sequenzen mit Homologie zu proteinkodierenden RNA-Sequenzen (funktionale Taxonomie)
- Funktionale Charakterisierung von Sequenzen mit Homologie zu proteinkodierenden Sequenzen
- Beschreibung von Funktionen und Organismen, die im Zusammenhang mit dem für die jeweilige Probe relevanten biologischen Phänomen (Gehörgangsentzündung für die Ohrprobe und Körpergeruchsentstehung für die Achselhöhlen-Proben) stehen.

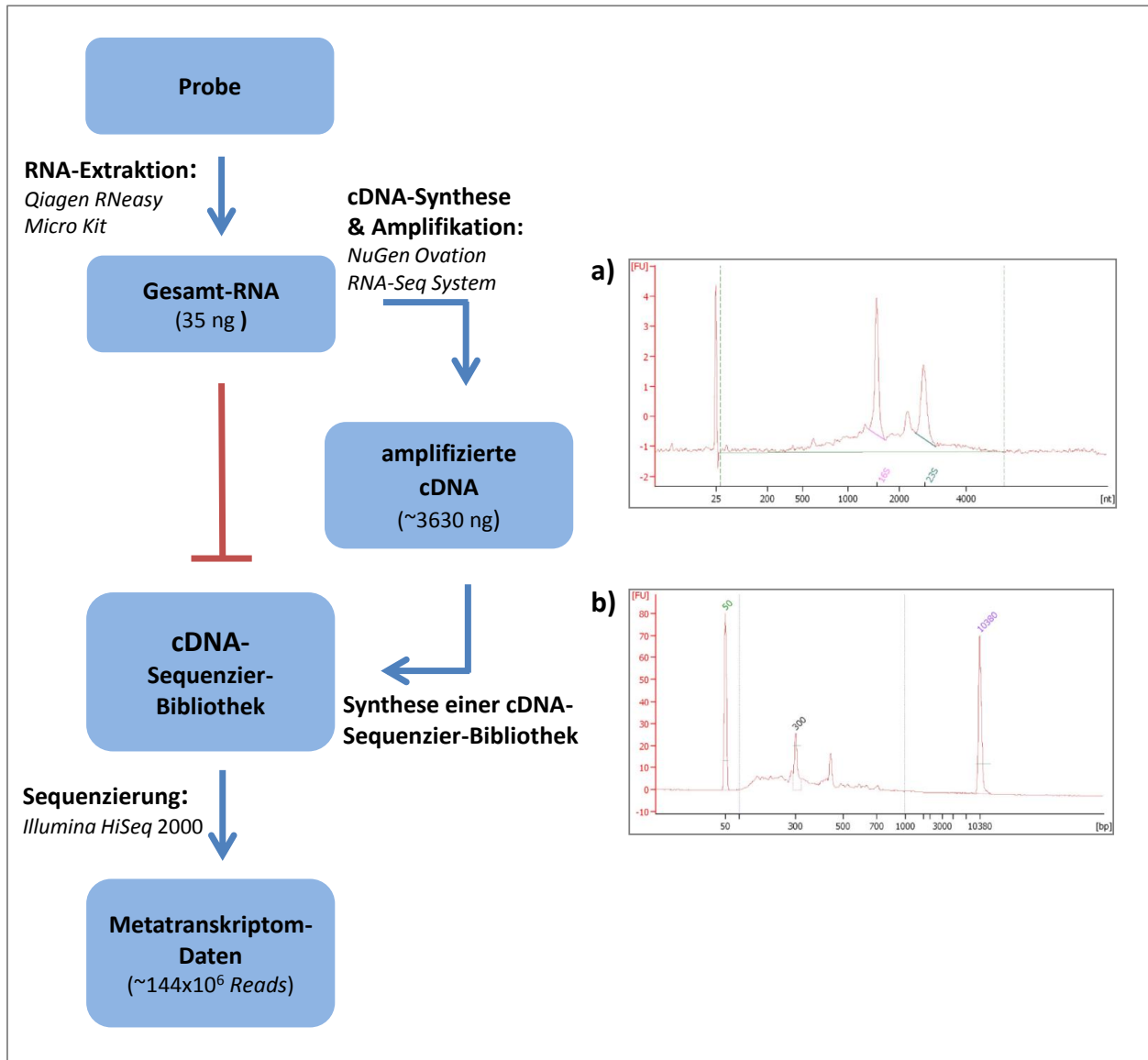
# 1 Metatranskriptomik des menschlichen Ohrs im Zusammenhang mit Außenohrentzündung

Im Rahmen der Metatranskriptom-Analyse des menschlichen Außenohrs wurde mithilfe einer Probe eines Otitis-externa-Patienten ein Laborverfahren zur Herstellung amplifizierter cDNA etabliert. Für die Analyse des aus dieser cDNA gewonnenen Sequenzierdatensatzes wurden zwei Ansätze gewählt. Zum einen wurde die online verfügbare MG-RAST-*Pipeline* verwendet. Zum anderen wurde die im Rahmen einer Zusammenarbeit mit Dr. Christina Ander und Prof. Dr. Jens Stoye am CeBiTec entwickelte Bioinformatik-*Pipeline* METRANS für die Auswertung benutzt. Im Folgenden werden zunächst die Ergebnisse der Laborarbeiten beschrieben und anschließend die Ergebnisse der Datenanalysen mit MG-RAST und METRANS dargestellt. Neben den oben beschriebenen zentralen Aspekten der Metatranskriptom-Analyse lag der Fokus der funktionalen Analyse dieses Datensatzes auf der Identifizierung und Beschreibung von Funktionen, die im Zusammenhang mit Pathogenität und Gehörgangsentzündung (Otitis externa) stehen.

## 1.1 Laborverfahren der Metatranskriptomik des menschlichen Ohrs

Das angewandte Laborverfahren für die Metatranskriptomik-Analyse des Ohrs basiert im Wesentlichen auf der Synthese und Amplifikation von cDNA aus isolierter Gesamt-RNA. Nach einem mechanischen Zellaufschluss wurde eine säulenbasierte RNA-Isolierung durchgeführt, die insgesamt 35 ng RNA mit einer guten Qualität (RIN = 7,7) lieferte (Abbildung 2 a). Für die Synthese von cDNA mithilfe des *Ovation RNA-Seq System V2* wurden 7 ng dieser RNA verwendet. Nach der cDNA-Synthese und Amplifikation wurden insgesamt 3630 ng cDNA mit einer Längenverteilung von 100 bp bis 1000 bp aufgereinigt (Abbildung 2 b). Die Herstellung der cDNA-Sequenzierbibliothek sowie die Sequenzierung wurde von der Firma *vertis Biotechnologie AG* durchgeführt. Die Sequenzierung erfolgte auf einer *Lane* eines *Illumina HiSeq-2000*-Gerätes mit einer Leseweite von 100 bp und lieferte insgesamt 143.923.092 *Reads*. Die Sequenzen wurden in Form einer *fastq*-Datei zur Verfügung gestellt.

In Abbildung 3 ist für jede Basenposition im *Read* der mittlere Qualitätswert als *Phred-Score* (logarithmisches Maß für die Fehlerwahrscheinlichkeit einer Basendetektion) dargestellt. Die Kurve besitzt einen für Illumina-Sequenzierungen charakteristischen Verlauf, bei dem die Qualität im vorderen Abschnitt stark ansteigt und im hinteren Sequenz-Abschnitt sukzessiv abfällt. Im Falle des vorliegenden Datensatzes ist der sprunghafte Abfall der mittleren Qualität auf einen Wert von 32 bei der Basenposition 40 besonders auffällig und verdient bei der



**Abbildung 2:** Schematische Darstellung der Labormethode zur Metatranskriptomik-Analyse der Ohrprobe. Ausgehend von der Probe eines Otitis-externa-Patienten wurde eine RNA-Isolierung durchgeführt. Aufgrund der geringen RNA-Mengen wurde cDNA synthetisiert und anschließend amplifiziert um die Herstellung von cDNA-Sequenzierbibliotheken und die High-Throughput-Sequenzierung zu ermöglichen. Dargestellt sind zudem Elektropherogramme der Gesamt-RNA (a) und der amplifizierten cDNA (b). Die Teilschritte der Synthese der cDNA Sequenzier-Bibliothek und der Sequenzierung wurden von der *vertis Biotechnologie AG* durchgeführt.

bioinformatischen Auswertung des Datensatzes besondere Beachtung. Mit Ausnahme der 100. Base besitzen alle Basen einen mittleren *Phred Score* größer als 25, was einer Fehlerwahrscheinlichkeit von weniger als 1 % entspricht. Die Sequenzen dienen als *Input* für die taxonomische und funktionale Metatranskriptom-Analyse mithilfe von MG-RAST und METRANS.



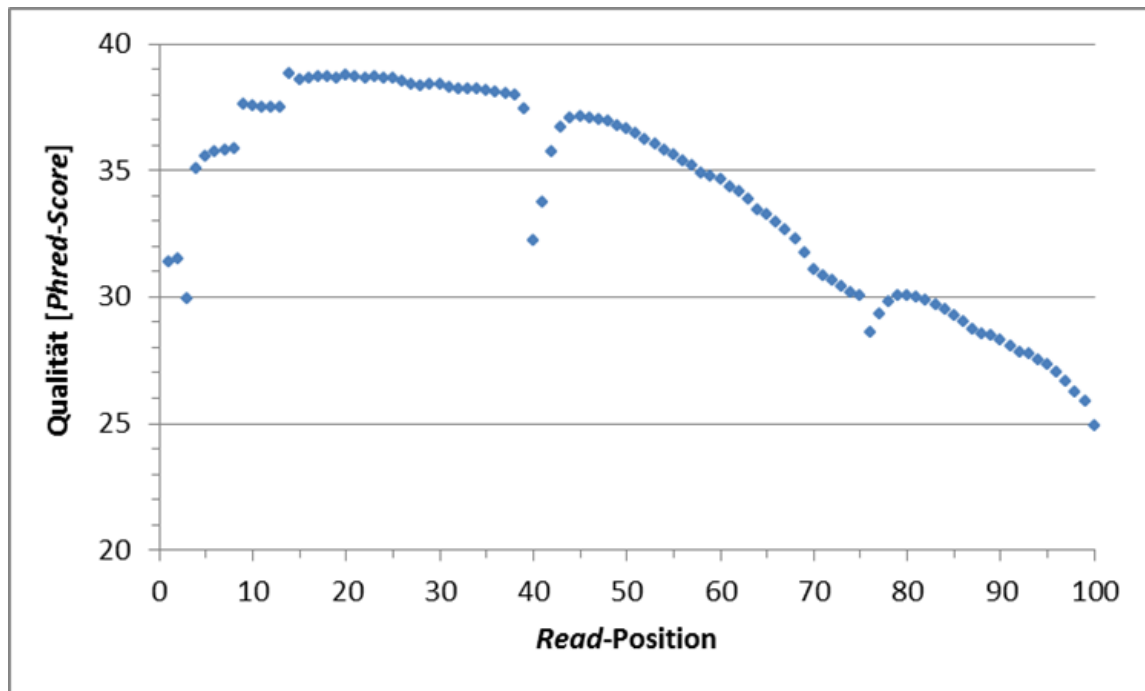


Abbildung 3: *Quality-Plot* des Metatranskriptomik-Datensatzes aus dem menschlichen Ohr. Dargestellt ist die mittlere Qualität als *Phred-Score* in Abhängigkeit von der Basen-Position im *Read*. Die Gesamtzahl der Reads beträgt 143.923.092.

## 1.2 Metatranskriptomik-Analyse der Ohr-Probe mithilfe von MG-RAST

### 1.2.1 Bioinformatik-basierte Beschreibung der Datenanalyse mittels MG-RAST

Die *MG-RAST-Pipeline* besteht aus mehreren bioinformatischen Analyse-Schritten, die im Wesentlichen in einer *Protein-Pipeline* und in einer *RNA-Pipeline* angeordnet sind. Als *Input* für die Datenanalyse in MG-RAST dienen die Illumina-Sequenzierdaten bestehend aus etwa 144 Millionen *Reads*. Die Analyse wurde unter Verwendung der in der Eingabemaske voreingestellten Parameter durchgeführt. Der Datensatz ist unter der Projektbezeichnung *Metatranscriptome ear* (Projekt-ID 5282) abgelegt. In Abbildung 4 ist eine Übersicht der Datenanalyse auf der Grundlage von *Read*- bzw. Cluster-Anzahlen dargestellt. Das *Preprocessing* diente als Qualitätskontrolle, während der etwa 3,5 Millionen *Reads* entfernt wurden. Dazu wurden die Sequenzen so getrimmt, dass sie anschließend maximal 5 Basen mit einem *Phred-Quality-Score* kleiner 15 aufweisen. Eine besonders starke Datenreduktion um einen Faktor von etwa 14 wird während des darauffolgenden *Dereplication*-Schrittes erzielt, bei dem Sequenz-Duplikate entfernt werden. Der Grund für diese starke Datenreduktion liegt maßgeblich in der Amplifikation während der cDNA-Synthese und während der Herstellung der

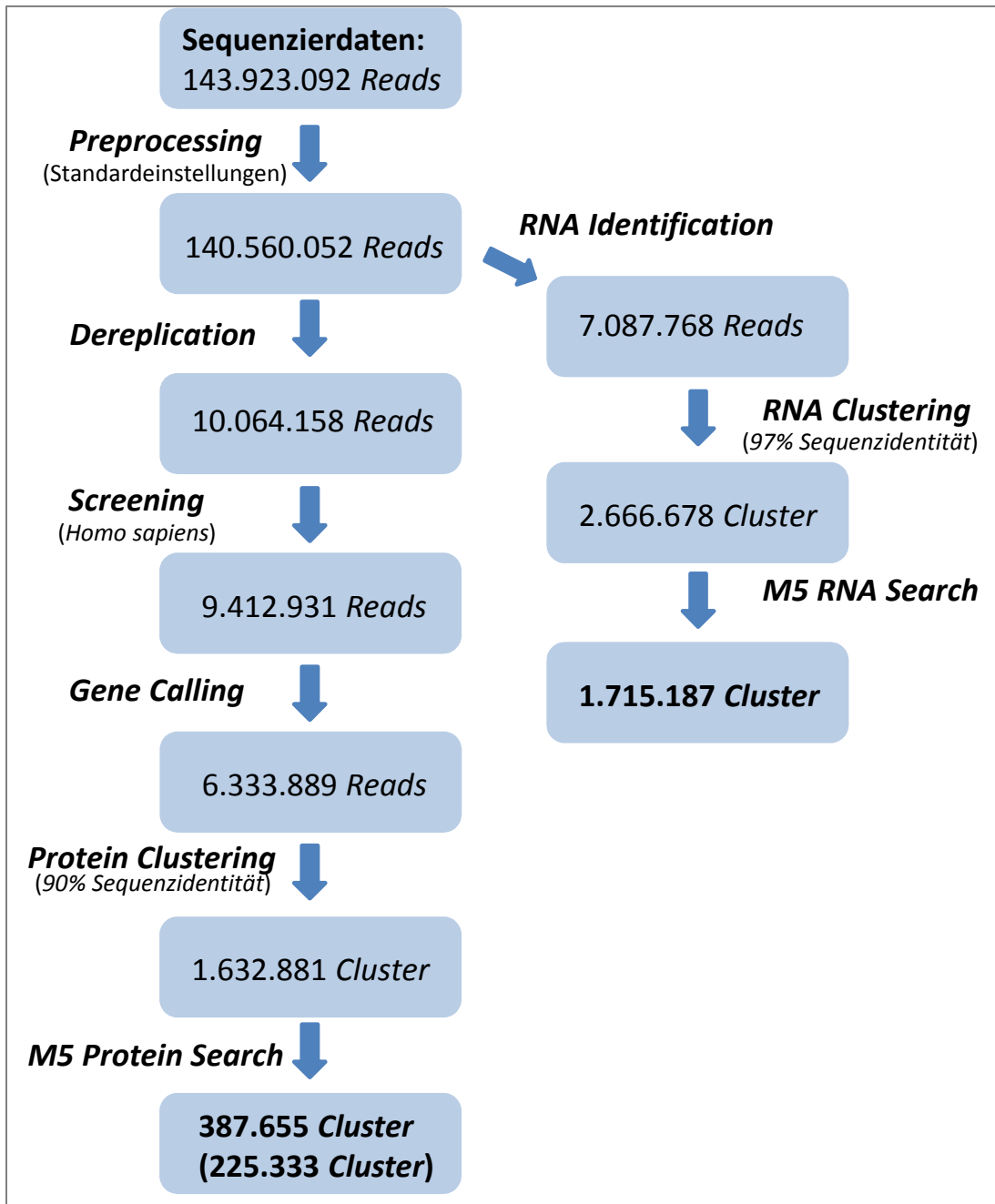


Abbildung 4: Übersicht der Sequenzdatenanalyse des Metatranskriptomik-Datensatzes aus dem Ohr in MG-RAST. Dargestellt ist jeweils die absolute Anzahl der *Reads* oder *Cluster*, die nach den jeweiligen Prozessierungsschritten für die weitere Analyse zur Verfügung stehen. Neben den Pfeilen werden die jeweiligen Analyseschritte beschrieben. Unten in Klammern ist die Anzahl an *Clustern* angegeben, die einer funktionalen Ontologie in der *M5NR*-Datenbank zugeordnet werden konnten.

Sequenzierbibliothek. Jedoch können auch Transkripte, die in mehreren Kopien im Metatranskriptom vorliegen, ebenfalls *Read*-Duplikate erzeugen. An dieser Stelle innerhalb der Analyse-Pipeline ist der Ursprung von Sequenz-Duplikaten nicht mehr nachvollziehbar. Während

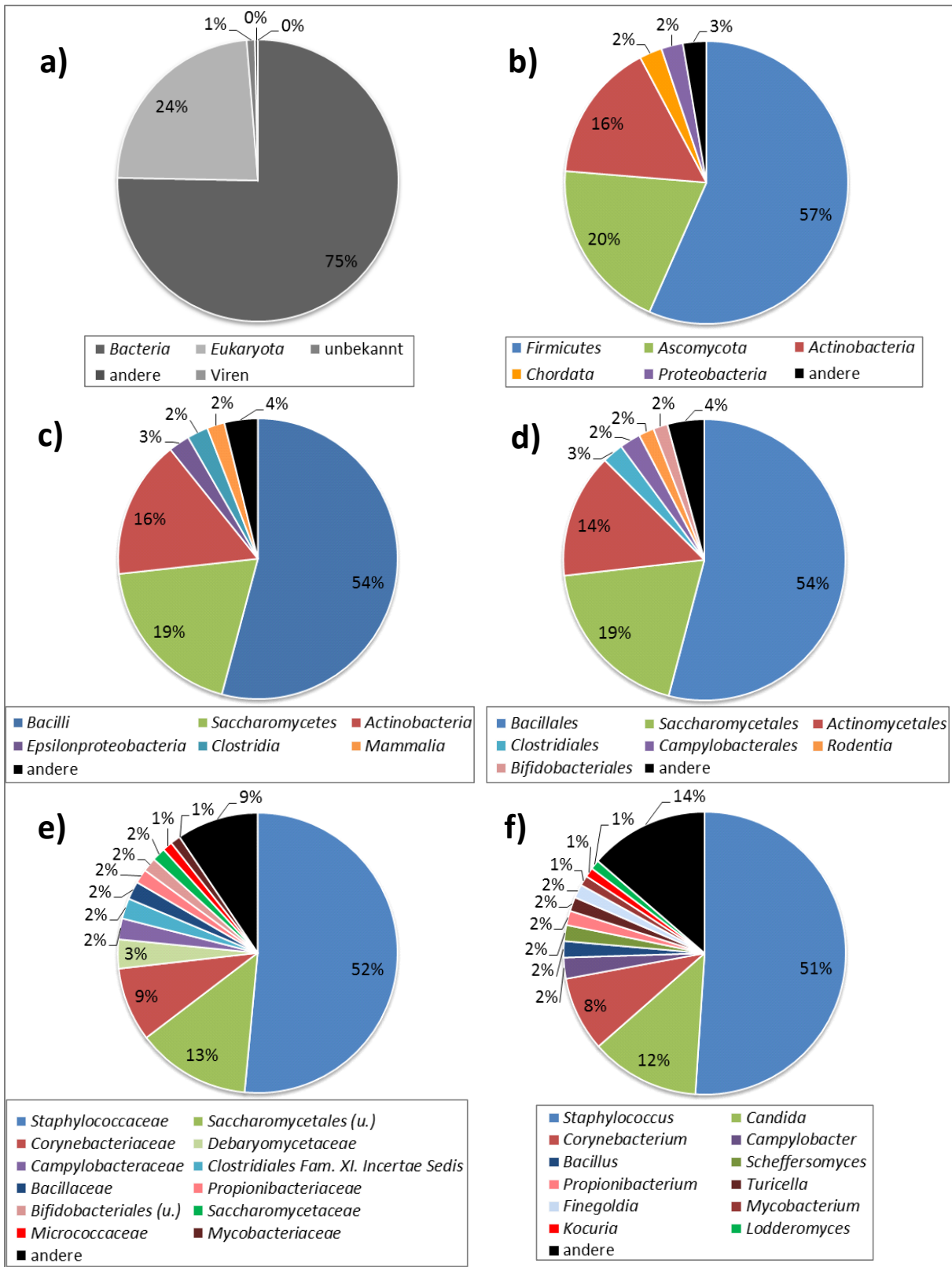
des *Screening* werden lediglich etwa 651.000 *Reads* mit Homologie zum menschlichen Genom entfernt, sodass noch 9.412.931 *Reads* in der *Analyse-Pipeline* verbleiben. Nach dem folgenden Analyse-Schritt *Gene Calling* verbleiben noch 6.333.889 *Reads*, die anschließend zu 1.632.881 Clustern mit mindestens 90 % Sequenzidentität zusammengefasst werden. Die Abfrage in der *M5NR*-Datenbank liefert 387.655 Sequenzcluster, die eine Sequenzhomologie bei den Standardeinstellungen für die Analyseparameter (Schwellenwerte: 60 % Sequenzidentität, *E-value*  $1 \times 10^5$ , 15 Aminosäuren Sequenzüberlappung) aufweisen. Davon können 255.333 Cluster den in der *M5NR*-Datenbank repräsentierten funktionalen Ontologien zugeordnet werden.

In der *RNA-Pipeline* wurden etwa 7 Millionen *Reads* als potenzielle rRNA-Sequenzen identifiziert und zu 2.666.678 Clustern mit mindestens 97 % Sequenzidentität zusammengefasst. Die Abfrage in der *M5RNA*-Datenbank lieferte 1.715.187 Treffer bei den Standardeinstellungen für die Analyseparameter. Insgesamt konnte im Rahmen der Datenanalyse mit MG-RAST eine Datenreduktion um einen Faktor von 68 erzielt werden und 2.102.842 Sequenzcluster stehen schließlich für die taxonomische und funktionale Analyse zur Verfügung.

### 1.2.2 Taxonomische Analyse mit MG-RAST

Im Rahmen der taxonomischen Analyse des Metatranskriptomik-Datensatzes aus dem Ohr wird zum einen die rRNA-basierte und zum anderen die funktionale Taxonomie beschrieben. Mithilfe der rRNA-basierten Taxonomie wird die Zusammensetzung der aktiven Mikrobiota durch taxonomische Zuordnung der Sequenzen mit Homologie zu rRNA-Sequenzen in der *M5RNA*-Datenbank beschrieben. Bei der *M5RNA* handelt es sich um eine nicht redundante Datenbank, in der die Referenzsequenzen der Datenbanken *SSU*, *LSU*, *RDP*, *ITS* und *Greengenes* zusammengefasst sind und die sich besonders gut für die umfassende taxonomische Klassifizierung eines Metatranskriptomik-Datensatzes eignet, da sowohl eukaryotische als auch prokaryotische rRNA der großen und der kleinen ribosomalen Untereinheiten dargestellt werden kann.

In Abbildung 5 ist die rRNA-basierte Zusammensetzung der Lebensgemeinschaft auf unterschiedlichen taxonomischen Ebenen auf Grundlage der den einzelnen Taxa zugeordneten *Read*-Abundanzen dargestellt. Auf der Domänen-Ebene wird die Mikrobiota von Bakterien mit einem Anteil von 75 % und die der Eukaryoten mit einem Anteil von 24 % dominiert. Sequenzhomologien zu Viren, unbekanntem Sequenzen und anderen Sequenzen liegen im Bereich von etwa 1 % der Abundanz auf der Domänen-Ebene. Auf der Ebene des Stammes sind



**Abbildung 5: rRNA-basierte Taxonomie des Metatranskriptomik-Datensatzes aus dem Ohr. Dargestellt ist die taxonomische Verteilung der *Read*-Abundanzen auf der Ebene der Domäne (a), des Stammes (b), der Klasse (c), der Ordnung (d), der Familie (e) und des Genus (f) bei einem Schwellenwert von 1 %. (u): unklassifiziert.**

die *Firmicutes* (57 %), *Ascomycota* (20 %) und *Actinobacteria* (16 %) dominant. *Chordata* (2 %), *Proteobacteria* (2 %) sowie andere Stämme (3 %) machen nur einen geringen Anteil aus. Im Einklang mit der Verteilung auf der Phylum-Ebene stellen auf der Ebene der Klasse *Bacilli* (54 %), *Saccharomycetes* (19 %) und *Actinobacteria* (16 %) die überwiegende Mehrheit dar. Lediglich ein geringer Anteil wird durch die *Epsilonproteobacteria* (3 %), *Clostridia* (2 %), *Mammalia* (2 %) sowie andere Klassen beschrieben. Auf der taxonomischen Ebene der Ordnung dominieren entsprechend die *Bacillales* (54 %), *Saccharomycetales* (19 %) und *Actinomycetales* (14 %). Außerdem werden die Ordnungen *Clostridiales* (3 %), *Campylobacterales* (2 %), *Rodentia* (2 %), *Bifidobacteriales* (2 %) sowie andere Ordnungen beschrieben. Die dominanten Familien stellen die *Staphylococcaceae* (52 %) und unklassifizierte *Saccharomycetales* (13 %) dar. Auf Genus-Ebene ergibt sich eine stark von *Staphylococcus* (51 %), *Candida* (12 %) und *Corynebacterium* (8 %) dominierte Lebensgemeinschaft. Die übrigen 29 % der aktiven Mikrobiota werden durch die Bakteriengattungen *Campylobacter*, *Bacillus*, *Propionibacterium*, *Turicella*, *Fingoldia*, *Mycobacterium*, *Kocuria*, die Pilzgattungen *Scheffersomyces* und *Lodderomyces* sowie andere Gattungen beschrieben.

Die funktionale Taxonomie wird für diesen Datensatz auf Grundlage der Sequenzhomologien zu proteinkodierenden Sequenzen in der *M5NR*-Datenbank beschrieben und ist in Abbildung 6 dargestellt. Dieses taxonomische Profil weist Ähnlichkeiten zu dem rRNA-basierten Profil auf. Die Dominanz der Sequenzhomologien zu *Staphylococcus* (38 %) auf der Genus-Ebene und entsprechend auf den höheren taxonomischen Ebenen (38 % *Staphylococcaceae* auf der Familien-Ebene, 39 % *Bacillales* auf der Ebene der Ordnung, 39 % *Bacilli* auf der Ebene der Klasse und 41 % der *Firmicutes* auf der Ebene des Stammes) ist vergleichbar mit dem rRNA-basierten taxonomischen Profil (51 % - 57 %), jedoch ist die Dominanz hier mit 38 % - 41 % schwächer ausgeprägt. Weiterhin sind die Anteile der Gattung *Candida* (7 %) und der entsprechenden Zuordnung auf höheren taxonomischen Ebenen (7 % unklassifizierte *Saccharomycetales* und 3 % *Saccharomycetaceae* auf der Ebene der Familie, 16 % *Saccharomycetales* auf der Ebene der Ordnung, 16 % *Saccharomycetes* auf der Ebene der Klasse und 19 % *Ascomycota* auf der Ebene des Phylums) in etwa vergleichbar mit denen im rRNA-Profil. Neben Zuordnungen zu den Pilzgattungen *Scheffersomyces* (1 %) und *Lodderomyces* (2 %), die ebenfalls in dem rRNA-basierten Profil zu finden sind, lassen sich in der funktionalen Taxonomie ebenfalls die Pilzgattungen *Meyerozyma* (1 %) und *Debaryomyces* (1 %) beschreiben.

Ein deutlicher Unterschied des Profils der funktionalen Taxonomie im Vergleich zum rRNA-basierten taxonomischen Profil liegt in dem Anteil der Sequenzen mit Homologie zu

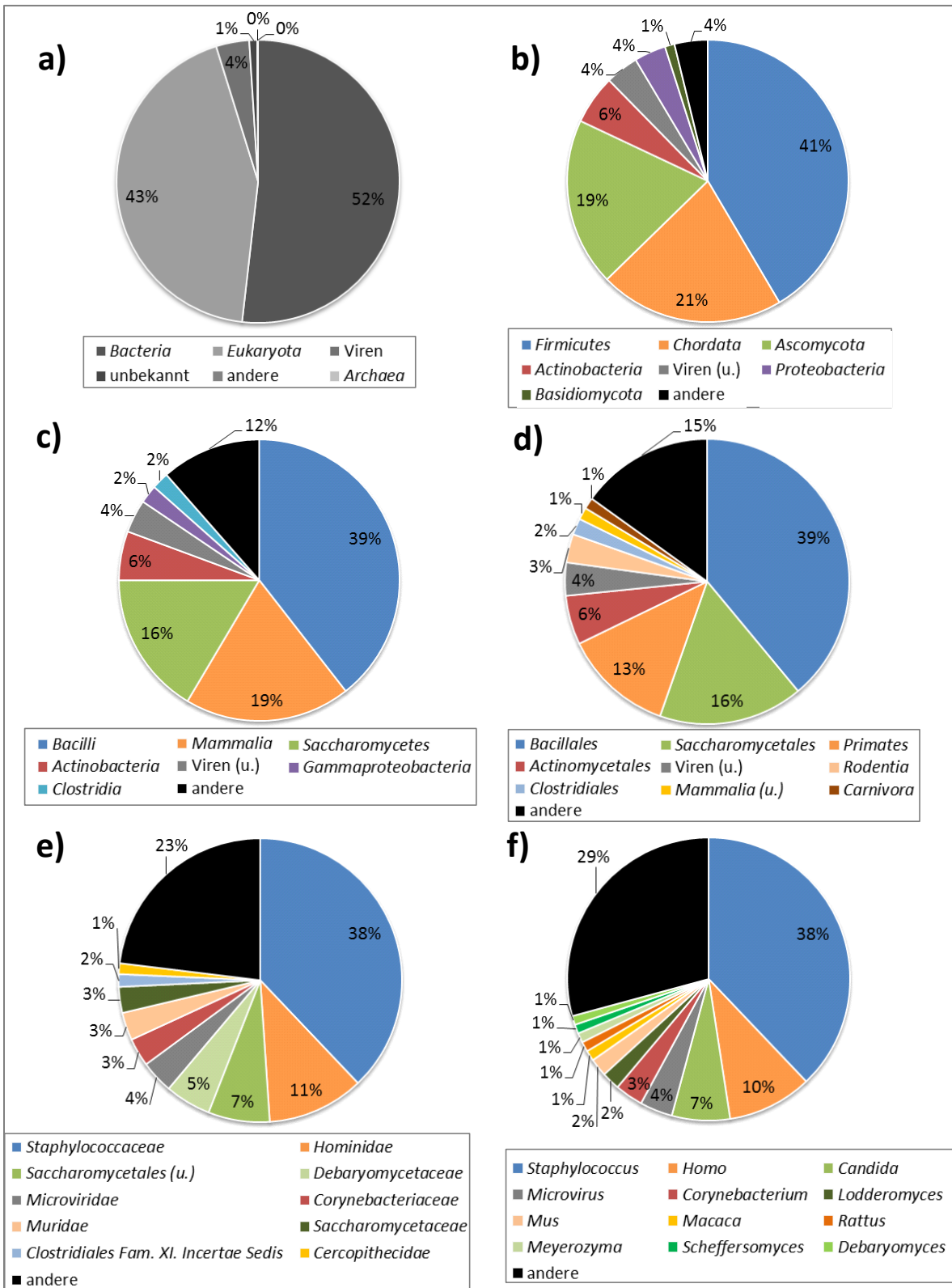
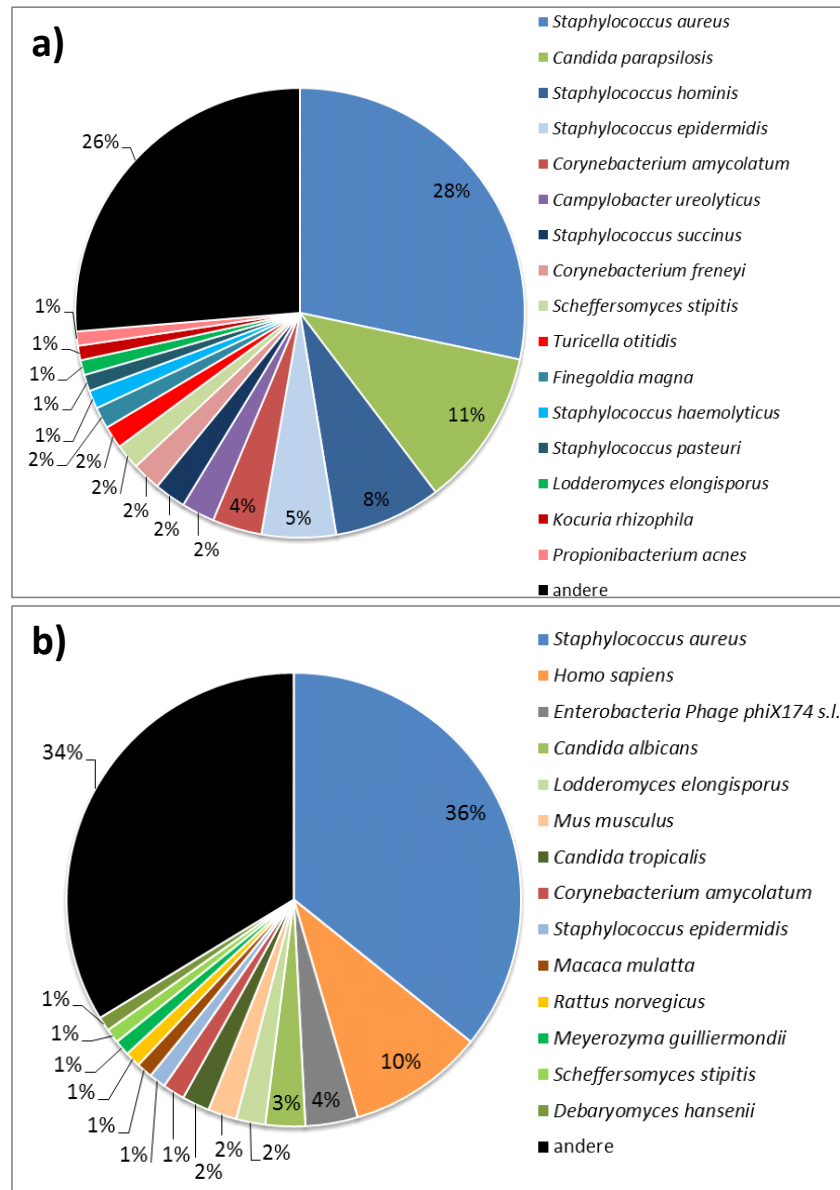


Abbildung 6: Funktionale Taxonomie des Metatranskriptomik-Datensatzes aus dem Ohr. Dargestellt ist die taxonomische Verteilung der *Read*-Abundanzen auf der Ebene der Domäne (a), des Stammes (b), der Klasse (c), der Ordnung (d), der Familie (e) und des Genus (f) bei einem Schwellenwert von 1 %. (u): unklassifiziert.

proteinkodierenden Sequenzen der *Chordata* (21 %) (Abbildung 6, b). Dies äußert sich bereits auf der Domänen-Ebene durch den deutlich höheren Anteil der Eukaryoten (43 % im Vergleich zu 24 %). Auf der Ebene der Klasse werden 19 % der Lebensgemeinschaft durch die *Mammalia* beschrieben. Auf der Ebene der Ordnung stellen die *Primates* (13 %), *Rodentia* (3 %) und unklassifizierte *Mammalia* (1 %) die nicht pilzlichen Eukaryoten dar. Auf der Ebene der Familie werden 11 % der Lebensgemeinschaft den *Hominidae* und auf Gattungsebene 10 % der Gattung *Homo* zugeordnet. Bei den entsprechenden Sequenzen handelt es sich wahrscheinlich um RNA-Sequenzen des Patienten, die beim *Screening* fälschlicherweise nicht entfernt wurden. Die Dominanz dieser Sequenzen in der funktionalen Taxonomie könnte durch alternative Spleißvarianten proteinkodierender mRNAs und ihre Unbekanntheit in den funktionalen Datenbanken erklärbar sein. Der Anteil der Sequenzen, der den Gattungen *Mus* (2 %), *Macacca* (1 %) und *Rattus* (1 %) zugeordnet wird, ist somit vermutlich auf das gleiche Phänomen zurückzuführen. Eine weitere Besonderheit der funktionalen Taxonomie im Vergleich zur RNA-basierten stellt die Anwesenheit der Viren dar (4 % Viren auf der Domänen-Ebene, jeweils 4 % unklassifizierte Viren auf Ebene des Phylums, der Klasse und der Ordnung, 4 % *Microviridae* auf Ebene der Familie und 4 % *Microvirus* auf der Gattungsebene). Da Viren und Phagen keine rRNA besitzen, können Sie nicht in der rRNA-basierten Taxonomie dargestellt werden und treten lediglich in der funktionalen Taxonomie in Erscheinung. Die Rolle der Viren in der aktiven Lebensgemeinschaft der untersuchten Probe ist jedoch gering. Dennoch deutet dieses Phänomen auf die Komplementarität der funktionalen und rRNA-basierten Taxonomie bei der Beschreibung der aktiven Mikrobiota hin. Insgesamt wird das Profil der funktionalen Taxonomie von den Gattungen *Staphylococcus* (38 %), *Homo* (10 %) und *Candida* (7 %) dominiert und deutet somit auf mögliche Pathogen-Wirt-Interaktionen innerhalb einer polymikrobiellen Lebensgemeinschaft hin. Im Vergleich zum rRNA-basierten taxonomischen Profil erscheint die funktionale Taxonomie diverser. Dies ist insbesondere auf die Zuordnung von Sequenzen zu den *Chordata* und den Viren auf der Ebene des Phylums und entsprechend auf allen anderen Ebenen zurückzuführen.

Die Verteilung der Sequenzen auf Spezies-Ebene ist in Abbildung 7 dargestellt. Das rRNA-basierte Profil der Lebensgemeinschaft (Abbildung 7 a) wird eindeutig von den Gattungen *Staphylococcus aureus* (28 %) und *Candida parapsilosis* (11 %) dominiert. Weiterhin können Sequenzen mit Homologie zu *Staphylococcus hominis* (8 %), *Staphylococcus epidermidis* (5 %), *Staphylococcus succinus* (2 %), *Staphylococcus haemolyticus* (1 %) und *Staphylococcus pasteurii* (1 %) beschrieben werden. Homologien zu weiteren Bakterienspezies, die im Zusammenhang mit der Haut-Mikrobiota bekannt sind, können hier ebenfalls durch *Corynebacterium*



**Abbildung 7: Taxonomische Darstellung des Metatranskriptomik-Datensatzes aus dem Ohr auf Spezies-Ebene. Dargestellt ist die rRNA-basierte taxonomische Verteilung der *Read*-Abundanzen auf Grundlage der *M5RNA*-Datenbank (a) und auf Grundlage der *M5NR*-Datenbank (b) bei einem Schwellenwert von 1 %. (s.l.): *sensu lato*.**

*amycolatum* (4%), *Corynebacterium freneyi* (2 %), *Finegoldia magna* (1 %) und *Propionibacterium acnes* (1 %) beschrieben werden. Als einziger Vertreter der *Proteobacteria* wird die pathogene Bakterienspezies *Campylobacter ureolyticus* (2 %) beschrieben (O'Donovan et al., 2014). Ebenfalls weisen 2 % der Sequenzen Homologien zu *Turicella otitidis*, einem im Zusammenhang mit Mittel- und Außenohrentzündungen beschriebenen Pathogen auf (von Graevenitz und Funke, 2014). Etwa 1 % der Sequenzen weisen Homologie zu Sequenzen des Bodenbakteriums *Kocuria rhizophila* auf. Weitere Pilzspezies, die durch Sequenzhomologien auf



der rRNA-Ebene beschrieben werden, sind *Scheffersomyces stipitis* (2 %) und *Lodderomyces elongisporus* (1 %). Sequenzhomologien zu anderen Spezies mit Anteilen < 1 % werden durch etwa 26 % der Sequenzen beschrieben.

Abbildung 7 b) zeigt die taxonomische Zuordnung proteinkodierender Sequenzen auf Grundlage der *M5NR*-Datenbank auf Spezies-Ebene. Dieses Profil wird ebenfalls von Sequenzen mit Homologie zu *S. aureus* dominiert (36 %). Weiterhin haben 1 % der Sequenzen Homologie zu *S. epidermidis*. Weitere 10 % der Sequenzen weisen Homologie zu *Homo sapiens* auf und 4 % der Sequenzen werden den *Enterobacteria Phagen phiX174 sensu lato* zugeordnet. Weiterhin werden geringere Anteile an Sequenzen den Pilzspezies *Candida albicans* (4 %), *Lodderomyces elongisporus* (2 %), *Meyerozyma guilliermondii* (1 %), *Scheffersomyces stipitis* (1 %) und *Debaryomyces hansenii* (1 %) zugeordnet. Der Anteil an Sequenzen, die anderen Spezies zugeordnet werden, liegt bei 34 % und ist somit deutlich höher als die entsprechende Zuordnung auf rRNA-Ebene (26 %, vgl. Abbildung 7 a).

Im Rahmen der rRNA-basierten Taxonomie wurden Annotationen zu 1180 Spezies mit einem Shannon-Diversitätsindex von 3,422 und einem Äquitätsindex von 0,484 beschrieben. Die funktionale Taxonomie lieferte Annotationen zu 1032 verschiedenen Spezies sowie einen etwas höheren Shannon-Diversitätsindex von 3,580 und einen Äquitätsindex von 0,516.

Das im Rahmen der Datenanalyse mit MG-RAST erstellte taxonomische Profil entspricht teilweise dem mikrobiologischen Befund und beschreibt das Bakterium *Staphylococcus aureus*, welches auch als Verursacher der Otitis externa bekannt ist, als die dominanteste Spezies innerhalb der Lebensgemeinschaft. Abweichend von dem mikrobiologischen Befund werden mithilfe der Metatranskriptomik zusätzlich *Candida*-Spezies und andere Pilzspezies wie *Lodderomyces elongisporus*, *Scheffersomyces stipitis*, *Meyerozyma guilliermondii* und *Debaryomyces hansenii* beschrieben.

### **1.2.3 Funktionale Analyse des Metatranskriptomik-Datensatzes des menschlichen Ohrs mit MG-RAST**

Zur Beschreibung des funktionalen Profils des Metatranskriptomik-Datensatzes aus dem Ohr wurde die *COG*-Klasseneinteilung verwendet (s. Abbildung 8). Dabei werden die Treffer gegen die *M5NR*-Datenbank, die sich der *COG*-Ontologie zuordnen lassen, als relative *Read*-Abundanzen betrachtet. Die vier häufigsten biologischen Funktionen werden durch die *COG*-Klassen „Translation, ribosomale Struktur und Biogenese“ (14,7 %), „Posttranslationale Modifikation, Protein-Turnover, Chaperone“ (14,2 %), „Energiegewinnung und -umwandlung“

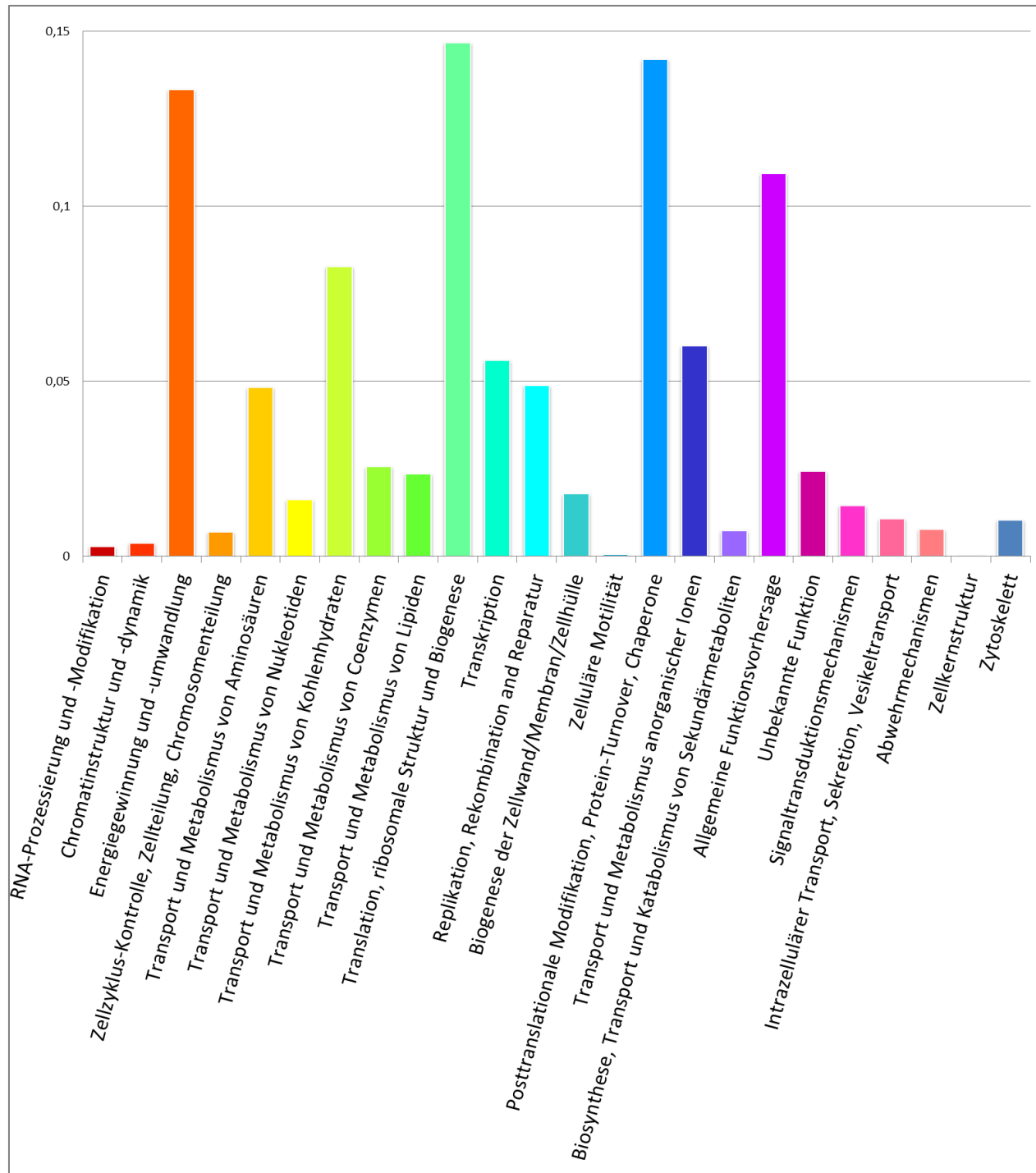


Abbildung 8: Funktionale Klassifizierung des Metatranskriptomik-Datensatzes aus der Ohr-Probe auf Grundlage der COG-Klasseneinteilung. Dargestellt ist der den COG-Klassen zugeordnete Teil-Datensatz bei den folgenden Einstellungen für die Schwellenwerte der Analyseparameter: *E-value*-Exponent: -5, Sequenzidentität: 80 % und Sequenzüberlappung 15 Aminosäuren. Als quantitatives Maß dient der Anteil der den jeweiligen Klassen zugeordneten *Read*-Abundanzen.

(13,3 %) und „Allgemeine Funktionsvorhersage“ (10,9 %) beschrieben und entsprechen etwa 53,1 % der funktionalen Zuordnungen für den gesamten Datensatz. Weiterhin beschreiben die COG-Klassen „Transport und Metabolismus von Kohlenhydraten“ (8,3 %), Transport und Metabolismus anorganischer Ionen“ (6,0 %), „Transkription“ (5,6 %), „Replikation, Rekombination und Reparatur“ (4,9 %) und „Transport und Metabolismus von Aminosäuren“ (4,8 %) weitere 29,7 % der funktionalen Annotationen im COG-System. Die übrigen 15 COG-Klassen beschreiben somit weniger als 20 % des gesamten Datensatzes.

Aufgrund der Dominanz von Sequenzen mit Homologie zu *Staphylococcus aureus* in dem Datensatz des Ohres wurden im nächsten Schritt Virulenzfaktoren des bekannten Pathogens analysiert. Dazu wurde der gesamte Datensatz zunächst taxonomisch auf Grundlage des M5NR-Annotationssystems mithilfe der *Representative Hit Classification* in MG-RAST klassifiziert. Anschließend wurden die Annotationen zu *S. aureus* auf Grundlage des *SwissProt*-Annotationssystems funktional klassifiziert. Die Ergebnisliste umfasste insgesamt etwa 1400 Einträge und wurde schließlich nach den relevanten Einträgen durchsucht und die jeweiligen *Read*-Abundanzen wurden erfasst. In Tabelle 9 sind die auf diese Weise identifizierten Virulenzfaktoren von *S. aureus* im Metatranskriptomik-Datensatz des Ohres zusammengefasst.

Pathogenitätsfaktoren von *S. aureus* sind in der Literatur sehr gut beschrieben und umfassen verschiedene Proteine aus den funktionalen Gruppen der Toxine, Adhäsionsfaktoren, Enzyme und Proteine, sowie Faktoren, die es *S. aureus* ermöglichen, die Immunabwehr des Wirtsorganismus zu umgehen (Zecconi und Scali, 2013). Insgesamt wurden in dem vorliegenden Datensatz Sequenzen mit Homologie zu 29 Pathogenitätsfaktoren aus acht unterschiedlichen Gruppen identifiziert (Tabelle 9). Dabei wurden besonders viele Sequenzen mit Homologie zu den Proteinen *Clumping factor A* (ClfA, 3169 *Reads*) und *Clumping factor B* (ClfB, 953 *Reads*) identifiziert. ClfA und ClfB gehören zu den Oberflächen-Adhäsinen, die auch als *MSCRAMMs* (*microbial surface recognizing adhesive matrix molecules*) eine wichtige Rolle bei der Besiedlung des Wirtsorganismus und bei der Umgehung der Immunantwort spielen (Chavakis et al., 2007; Speziale et al., 2009). Die beiden Oberflächen-Proteine binden Fibrinogen und Fibrin und es kommt zur Ausbildung von Aggregaten zwischen *S. aureus* und den Thrombozyten des Wirtsorganismus, wodurch die Immunerkennung des Wirtes umgangen wird. Weiterhin bindet ClfB die Proteine Cytokeratin 8 und Cytokeratin 10 und ist somit an der Besiedlung der Epithelzellen beteiligt (Clarke und Foster, 2006; Haim et al., 2010). Darüber hinaus spielt ClfB eine Rolle bei der Ausbildung von Biofilmen (Speziale et al., 2014). Insgesamt 752 *Reads* weisen Homologie zum *Elastin-binding protein* (Ebbs) auf, welches ebenfalls ein Adhäsionsfaktor ist und

Tabelle 9: Beschreibung der Virulenzfaktoren von *Staphylococcus aureus* aus dem Metatranskriptomik-Datensatz des menschlichen Ohres in MG-RAST auf Grundlage des SwissProt-Annotationssystems. Als quantitatives Maß für die funktionalen Annotationen dienen die jeweiligen *Read*-Abundanzen.

| Virulenzfaktoren-Gruppe                   | Virulenzfaktor   | # Reads |
|---|--|---------|
| Enterotoxine                              | <i>Enterotoxin type A</i>  | 3       |
|   | <i>Enterotoxin type E</i>  | 4       |
|   | <i>Enterotoxin type G</i>  | 5       |
| Adhäsionsfaktoren                         | <i>Fibronectin-binding protein A</i>                             | 131     |
|   | <i>Immunglobulin G-binding protein A</i>                         | 4       |
|   | <i>Serin-aspartate repeat protein C</i>                          | 224     |
|   | <i>Serin-aspartate repeat protein D</i>                          | 73      |
|   | <i>Serin-aspartate repeat protein E</i>                          | 9       |
|   | <i>Elastin-binding protein Ebps</i>                              | 752     |
|   | <i>Iron-regulated surface determinant A</i>                      | 15      |
|   | <i>Iron-regulated surface determinant B</i>                      | 3       |
|   | <i>Iron-regulated surface determinant H</i>                      | 5       |
|   | <i>Surface protein</i>   | 2       |
|   | <i>Surface protein G</i>   | 32      |
|   | <i>Serine-rich adhesin for platelets</i>                         | 65      |
|   | <i>Extracellular matrix protein-binding protein Emp</i>          | 1       |
| <i>Immunoglobulin-binding protein Sbi</i> | 10   |         |
| Enzyme und andere Proteine                | <i>O-acetyltransferase OatA</i>                                  | 8       |
| Resistenz gegen oxidativen Burst          | <i>Catalase CatA</i>   | 414     |
| Komplement-Inaktivierung                  | <i>Capsular polysaccharide type 5 biosynthesis protein Cap5A</i> | 69      |
|   | <i>Staphylokinase</i>  | 3       |
|   | <i>Staphylococcal complement inhibitor</i>                       | 19      |
|   | <i>Clumping factor A</i>   | 3169    |
| Umgehung der angeborenen Immunantwort     | <i>Clumping factor B</i>   | 953     |
|   | <i>Zinc metalloproteinase aureolysin</i>                         | 1       |
| Antibiotika-Resistenzen                   | <i>Transposase for insertion sequence-like element IS431mec</i>  | 4       |
|   | <i>Tetracycline resistance protein</i>                           | 1       |

darüber hinaus eine Rolle bei der Biofilm-Ausbildung spielt (Nakakido et al., 2014). Das Protein *Fibronectin-binding protein A* ist ebenfalls ein Adhäsion, welches an der Besiedlung des Wirtes durch die Bindung von Fibronectin und Fibrin beteiligt ist und ebenfalls bei der Ausbildung von Biofilmen eine Rolle spielt (Chavakis et al., 2007; Clarke und Foster, 2006; Speziale et al., 2009, 2014). Weiterhin wurden Sequenzhomologien zu den folgenden an der Adhäsion und an der Biofilm-Bildung beteiligten Proteinen in *S. aureus* identifiziert: *S. aureus surface protein G* (32 Reads), *serin-aspartate repeat protein C* (224 Reads) (Speziale et al., 2014; Zecconi und Scali, 2013). Insgesamt 414 Reads weisen Sequenzhomologie zu der Sequenz des Proteins Catalase A von *S. aureus* auf (s. Tabelle 9). Catalase A inaktiviert Wasserstoffperoxid, welches an der Entstehung reaktiver Sauerstoffspezies beteiligt ist, und spielt somit eine entscheidende Rolle bei der Besiedlung der Nasenschleimhäute (Chavakis et al., 2007). Sequenzhomologien zu weiteren Virulenzfaktoren von *S. aureus* werden ebenfalls identifiziert (s. Tabelle 9). Die entsprechenden Annotationen weisen jedoch im Vergleich zu den oben beschriebenen Adhäsionsfaktoren relativ geringe Read-Abundanzen auf. Somit stellt die Adhäsion die abundanteste virulente Funktion von *S. aureus* im Metatranskriptomik-Datensatz der hier untersuchten Probe dar. Zu den Virulenzmechanismen von *Candida spp.* ist in der Literatur relativ wenig bekannt. Im Allgemeinen wird jedoch angenommen, dass die Virulenz im Wesentlichen durch Protease- und Phospholipase-Aktivitäten beschrieben wird (Arsović et al., 2009; Oksuz et al., 2007). In Analogie zu der für Virulenzfaktoren von *S. aureus* durchgeführten Analyse wurde der Metatranskriptomik-Datensatz des Ohrs mithilfe der *Representative Hit Classification* auf Grundlage des *M5NR*-Annotationssystems klassifiziert. Alle Sequenzen, die dem Genus *Candida* zugeordnet wurden, wurden weiterführend mithilfe des *KO*-Annotationssystems klassifiziert. Die Filtereinstellung nach der EC-Nummer 3.4. (Peptidase) lieferte insgesamt 37 funktionale Annotationen zu 8.332 Reads, die unterschiedlichen physiologischen Funktionen zugeordnet wurden. Eine weitere Filtereinstellung nach der EC-Nummer 3.1. (Phospholipase) lieferte ebenfalls 37 funktionale Annotationen zu 7.916 Reads. Somit können Peptidase- und Phospholipase-Funktionen von *Candida*-Spezies im Metatranskriptomik-Datensatz der hier untersuchten Probe beschrieben werden.

Die Koisolation von *Candida spp.* und *S. aureus* im Zusammenhang mit diversen Krankheiten oder biomedizinischen Materialien wie Prothesen oder Implantaten kommt häufig vor, da beide Organismen menschliche Schleimhäute besiedeln und als opportunistische Pathogene Infektionen verursachen können (Nair et al., 2014; Peters und Noverr, 2013; Shirtliff et al., 2009). Da *C. albicans* den bekanntesten Vertreter der Gattung *Candida* darstellt, ist diese Spezies besonders gut untersucht. Kooperative Interaktionen zwischen *S. aureus* und *C. albicans* sind insbesondere im Zusammenhang mit der Biofilm-Bildung bekannt (Kojic und Darouiche,

2004). Dabei bildet *C. albicans* die Basis des Biofilms und erleichtert somit die Bildung eines gemischten Biofilms mit *S. aureus*. Die Organisation in einem gemischten Biofilm erhöht die Vancomycin-Resistenz von *S. aureus* (Harriott und Noverr, 2009). Das Hyphen-Protein Als3p (*agglutinin-like sequence*) vermittelt dabei die Bindung zwischen *C. albicans* und *S. aureus*. In dem vorliegenden Metatranskriptomik-Datensatz wurden keine Sequenzen mit Homologie zu Als3p identifiziert. Auch die Sekretion von Proteasen durch *S. aureus* und *C. albicans* stellen einen kooperativen molekularen Mechanismus dar, der sowohl die Adhäsion von *C. albicans* an Schleimhäute als auch die Degradation der F<sub>c</sub>-Region von Immunglobulin G ermöglicht (Nair et al., 2014).

Im Zusammenhang mit der angeborenen Immunabwehr des Menschen bei *S. aureus*-Infektionen ist insbesondere die Aktivierung spezifischer Rezeptoren zur Erkennung von *PAMPs* (*pathogen-associated molecular patterns*) in der Literatur beschrieben (Fournier und Philpott, 2005). Dabei handelt es sich im Wesentlichen um die *Toll-like*-Rezeptoren, die *NOD-like* Rezeptoren, *C-type-lectin*-Rezeptoren sowie die *RIG-I-like*-Rezeptoren (Kawai und Akira, 2011). Während der Analyse des Metatranskriptomik-Datensatzes des Ohrs wurden im Rahmen des *Screening* Sequenzen mit Homologie zu genomischen Sequenzen des Menschen entfernt und stehen somit nicht für die Analyse zur Verfügung (s. Abschnitt 1.2.1). Jedoch sind im Rahmen der funktionalen Taxonomie Sequenzen mit Homologie zu menschlichen Sequenzen identifiziert worden (s. Abbildung 7 b). Diese Sequenzen wurden vermutlich aufgrund alternativer Spleißvarianten der mRNAs nicht im Rahmen des *Screening* entfernt und stehen somit für die weitere Analyse zur Verfügung. In Analogie zur Analyse der Virulenzfaktoren von *S. aureus* wurde der Metatranskriptomik-Datensatz der Ohr-Probe mithilfe der *Representative Hit Classification* auf Grundlage des *M5NR*-Annotationssystems klassifiziert. Alle Sequenzen, die auf taxonomischer Ebene der Klasse der *Mammalia* zugeordnet wurden, wurden weiterführend mithilfe des *SwissProt*-Annotationssystems funktional klassifiziert und die Ergebnisliste wurde schließlich mit der Filter-Funktion nach Annotationen zu den oben genannten Rezeptoren durchsucht. Dabei wurden insgesamt 32 Sequenzen mit Homologie zu Sequenzen der *Toll-like* Rezeptoren 1, 2, 4, 6 und 8 identifiziert. Insgesamt 155 *Reads* wurden zu Genprodukten der *C-type domain member family* annotiert, welche die Lectin-Rezeptoren des Typen C kodieren. Sequenzen mit Homologie zu den anderen beiden Rezeptortypen wurden nicht identifiziert.

Zusammenfassend lässt sich festhalten, dass die Analyse des Metatranskriptomik-Datensatzes des Ohres die im Zusammenhang mit Gehörgangsentzündung aus der Literatur bekannte Mikrobiota darstellt. Die funktionale Analyse mithilfe der *COG*-Ontologie liefert ein Profil, das von Funktionen des Primärmetabolismus dominiert wird. Sequenzhomologien zu

Virulenzfaktoren von *S. aureus*, die im Zusammenhang mit Adhäsion und Biofilm-Bildung stehen, wurden im Rahmen gezielter funktionaler Analysen identifiziert. Weiterhin wurden Sequenzhomologien zu Peptidasen und Phospholipasen von *Candida*-Spezies beschrieben. Bezüglich der Immunabwehr des Menschen können Sequenzhomologien zu *Toll-like*-Rezeptoren sowie zu Genprodukten der *C-type domain member family* zugeordnet werden.

## 1.3 Metatranskriptomik-Analyse des menschlichen Ohrs mithilfe von METRANS

Als Alternative zur Metatranskriptomik-Analyse mittels MG-RAST wurde METRANS verwendet. Dabei handelt es sich um eine Analyse-Pipeline für Metatranskriptomik-Datensätze, die von Dr. Christina Ander und Prof. Dr. Jens Stoye am CeBiTec entwickelt und zur Verfügung gestellt wurde. Im Folgenden wird die Analyse des Metatranskriptomik-Datensatzes aus dem Ohr mithilfe von METRANS beschrieben.

### 1.3.1 Bioinformatik-basierte Beschreibung der Datenanalyse mittels METRANS

Die METRANS-Pipeline setzt sich zusammen aus drei bioinformatischen Analyseschritten und einer abschließenden taxonomischen Charakterisierung der Sequenzen mit Homologie zu rRNA sowie einer funktionalen und taxonomischen Beschreibung von Sequenzen mit Homologie zu proteinkodierenden Sequenzen. Eine Übersicht der Kennzahlen der einzelnen Analyseschritte ist in Abbildung 9 dargestellt. Als *Input* diente die *fastq*-Datei mit den Sequenzierdaten bestehend aus 143.923.092 *Reads*. Den ersten Schritt in der Analysepipeline stellt das *Quality Trimming* mit einem Qualitäts-Schwellenwert für einen *Phred-Score* von 35 dar. Dabei wurde ein *Sliding Window* von 100 Basen gewählt, da der deutliche Abfall der mittleren Qualität der Sequenzierdaten an der Basenposition 40 bei einem kürzeren Betrachtungsintervall die Sequenzen verkürzt hätte. Nach dem *Quality Trimming* verblieben noch 64.804.207 *Reads* in der Pipeline, was etwa 45 % des ursprünglichen Sequenzierdatensatzes entspricht. Um den Rechenaufwand zu minimieren und eine Datenreduktion zu erreichen wurden die verbliebenen Sequenzen unter Verwendung des Bioinformatik-Tools *dnaclust* in Cluster (*Bins*) mit mindestens 98 % Sequenzidentität eingeteilt. Auf diese Weise wurden 5.000.434 *Bins* generiert, was einer Datenreduktion um einen Faktor von etwa 29 entspricht. Für jedes *Bin* wurde eine repräsentative Sequenz generiert, die für alle folgenden Analyseschritte verwendet wurde. Für die quantitativen Darstellungen der Taxonomie und der Funktionen bleiben die Sequenzhäufigkeiten der einzelnen *Bins* jedoch erhalten und werden bei der Normalisierung berücksichtigt.

Den nächsten Analyseschritt stellt das *Filtering* dar. Dabei werden Sequenzen mit Homologie zu rRNA-Sequenzen mit den *SILVA*-Datenbanken *SSU* und *LSU* sowie mit der *Rfam*-Datenbank abgeglichen und aus dem Datensatz entfernt. Die Abfrage gegen die *SSU*-Datenbank lieferte 923.058 Treffer, was 16.616.409 *Reads* entspricht. Die Abfrage gegen die *LSU*-Datenbank ergab 2.781.474 Treffer (53.723.195 *Reads*) und die Abfrage gegen die *Rfam*-Datenbank lieferte 532.778 Treffer (7.494.571 *Reads*). Da die kumulative *Read*-Anzahl mit Homologien in den drei Datenbanken größer ist als die Anzahl der nach dem *Quality Trimming* verbliebenen *Reads*, kann



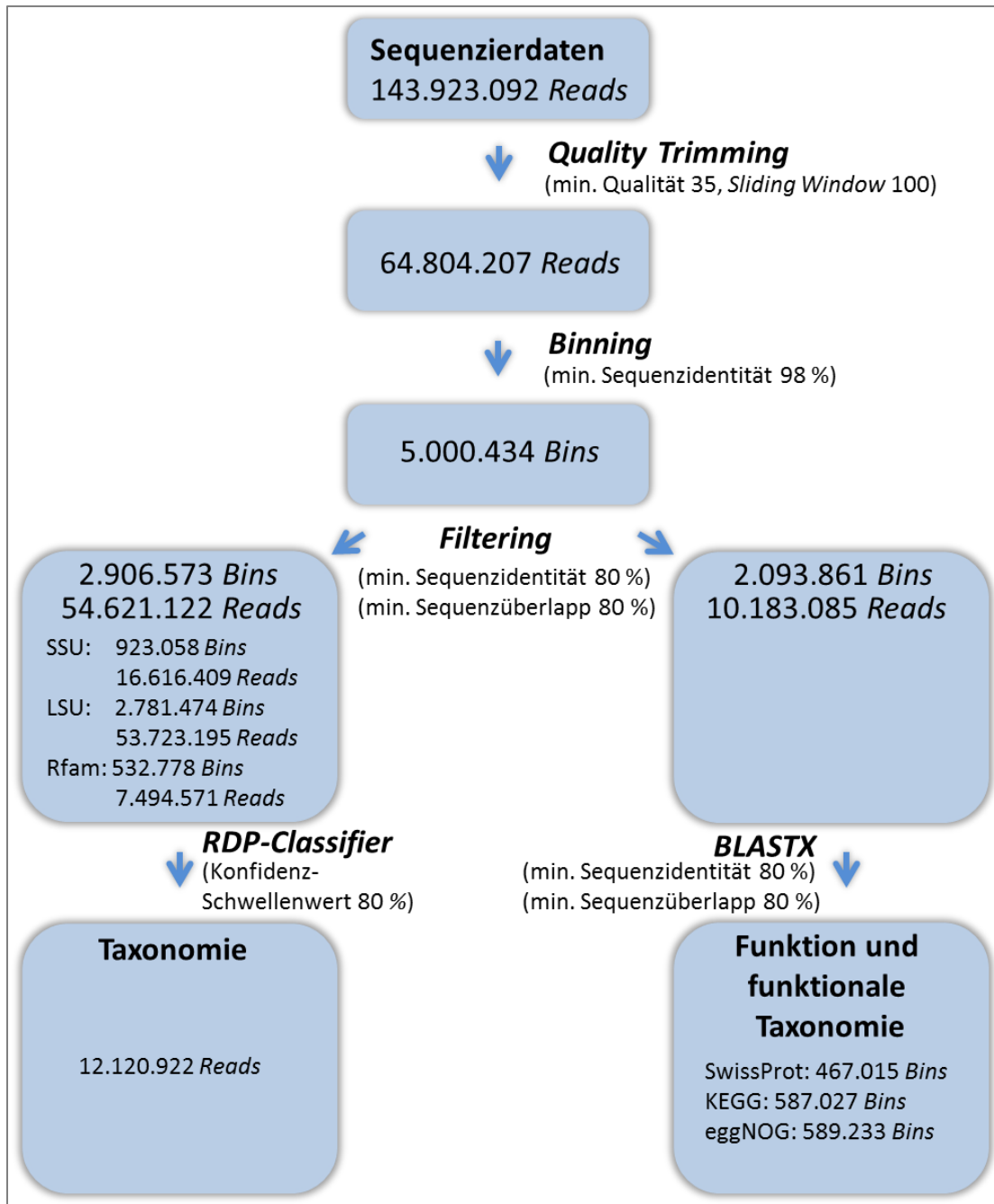


Abbildung 9: Übersicht der Sequenzdatenanalyse des Metatranskriptomik-Datensatzes aus dem Ohr mittels METTRANS. Dargestellt ist jeweils die absolute Anzahl der Reads oder Bins, die nach den jeweiligen Prozessierungsschritten für die weitere Analyse zur Verfügung stehen. Neben den Pfeilen werden die jeweiligen Analyseschritte und Tools beschrieben.

angenommen werden, dass bestimmte Sequenzen Homologien in mehr als einer Datenbank aufweisen. Alle Sequenzen mit Homologien in mindestens einer der drei Datenbanken werden in dem Datensatz entsprechend markiert und nicht mehr für die funktionale Analyse verwendet. Alle Sequenzen mit einem Treffer gegen die SSU-Datenbank werden anschließend in der

taxonomischen Analyse mit dem RDP-CLASSIFIER verwendet. Bei dem gewählten Konfidenz-Schwellenwert von 80 % wurden insgesamt 12.120.922 *Reads* klassifiziert.

Die Sequenzen, die im Rahmen des *Filtering*-Schrittes keine Treffer gegen die rRNA-Datenbanken liefern, werden in 2.093.861 Bins zusammengefasst und für die funktionale Klassifizierung und für die Darstellung der funktionalen Taxonomie verwendet. Für die Zuordnung funktionaler Annotationen werden die repräsentativen Sequenzen der einzelnen *Bins* mithilfe des BLASTX-Algorithmus bei einem Schwellenwert für die Sequenzidentität und für die Sequenzüberlappung von  $\geq 80$  % in den Datenbanken *SwissProt*, *KEGG* und *eggNOG* abgefragt. Mithilfe der *SwissProt*-Datenbank konnten insgesamt 467.015 *Bins* zugeordnet werden. Die Abfrage in der *KEGG*-Datenbank lieferte Treffer für 587.027 *Bins* und mithilfe der *eggNOG*-Datenbank konnten 589.233 *Bins* funktional klassifiziert werden.

### 1.3.2 16S-rRNA-basierte Taxonomie mittels METRANS

Für die taxonomische Analyse der aktiven Mikrobiota der Ohr-Probe wird die 16S-rRNA-basierte Klassifizierung mithilfe des RDP-CLASSIFIER verwendet. Die normalisierten Häufigkeitsverteilungen der *Reads* auf allen taxonomischen Ebenen bis einschließlich der Gattungsebene sind in Abbildung 10 dargestellt. Auf der Ebene des Reiches dominieren Bakterien (ca. 94,5 %), während etwa 6 % der Sequenzen nicht klassifiziert werden konnten. Weiterhin konnten insgesamt 36 Sequenzen mit Homologie zu Archaeen-Sequenzen beschrieben werden. Diese konnten jedoch nicht mit hinreichender Konfidenz auf den folgenden taxonomischen Ebenen klassifiziert und somit nicht genauer beschrieben werden. Daher sind auf der Ebene des Stammes ausschließlich Sequenzen mit Homologie zu rRNA-Sequenzen aus bakteriellen Stämmen vertreten. Die Verteilung wird dabei von den *Firmicutes* (ca. 94,7 %) dominiert. Ca. 5,1 % der klassifizierten Sequenzen können den Aktinobakterien und etwa 0,2 % den Proteobakterien zugeordnet werden. Homologien zu anderen Bakterienstämmen sind nur in sehr geringen Anteilen (ca. 0,1 %) vorhanden. Auf der Ebene der Klasse dominieren die *Bacilli* mit einem Anteil von 94 %. Weitere 6 % der *Reads* werden durch die *Actionobacteria* beschrieben. Auf der Ebene der Ordnung weisen 92 % der klassifizierten Sequenzen Homologie zu den *Bacillales* und 7 % zu den *Actinomycetales* auf. Auf der taxonomischen Ebene der Familie weisen 97 % der Sequenzen Homologie zu den *Staphylococcaceae* und 2 % der Sequenzen Homologie zu den *Brevibacteriaceae* auf. Auf der Genus-Ebene stellt *Staphylococcus* (96 %) den weitaus größten Anteil der Sequenzen dar. Weitere 3 % der Sequenzen werden der Gattung *Brevibacterium* zugeordnet und 1 % der Sequenzen wird durch Zuordnung zu anderen Gattungen beschrieben.

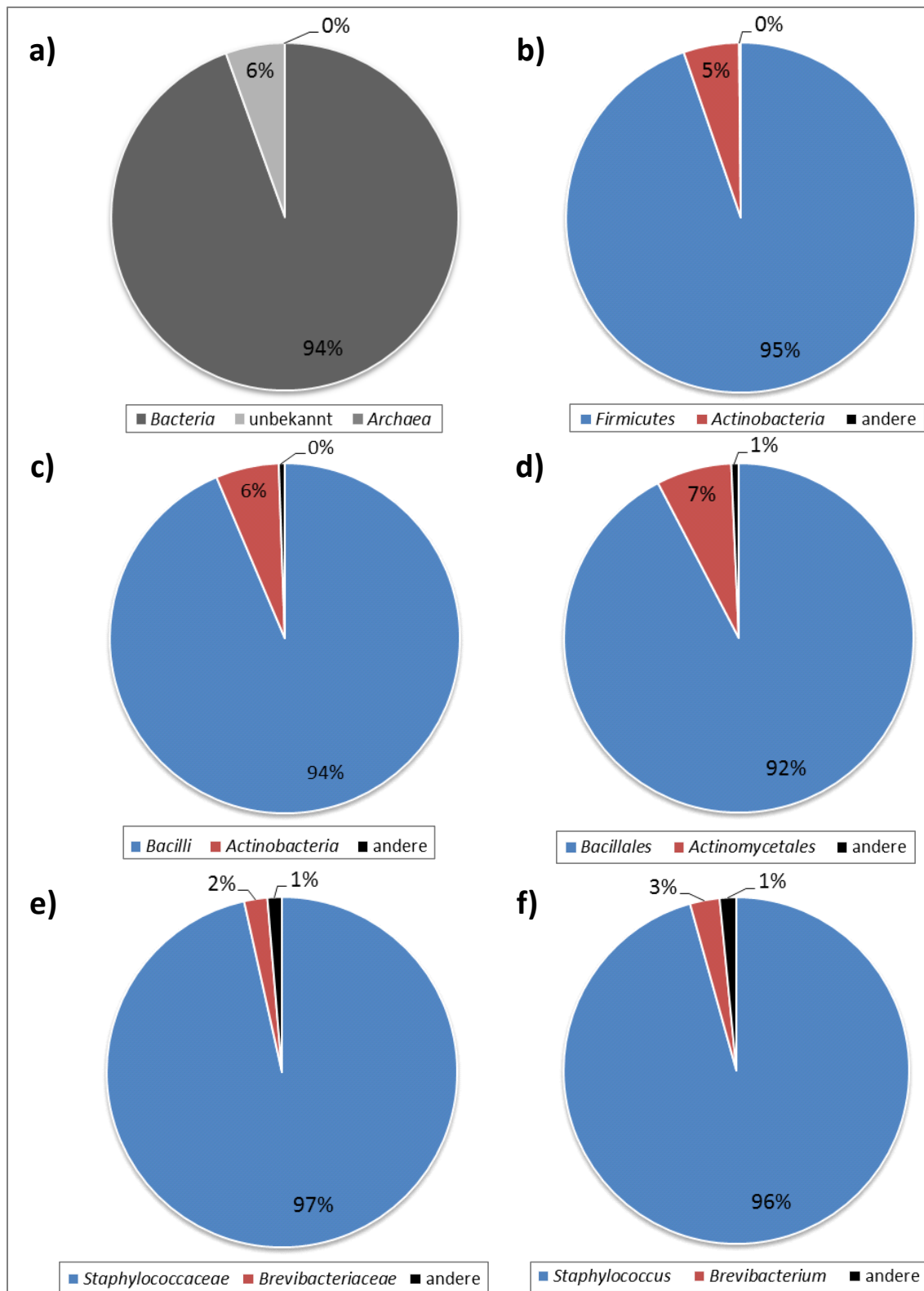


Abbildung 10: 16S-rRNA-basierte taxonomische Zuordnung des Metatranskriptomik-Datensatzes der Ohr-Probe mittels METTRANS unter Verwendung des RDP-CLASSIFIER bei einem Konfidenz-Schwellenwert von 0,8. Dargestellt ist die taxonomische Verteilung der *Read*-Abundanzen auf der Ebene der Domäne (a), des Stammes (b), der Klasse (c), der Ordnung (d), der Familie (e) und des Genus (f).

Eine 16S-rRNA-basierte taxonomische Klassifizierung des Datensatzes auf der Spezies-Ebene ist auf Grundlage des RDP-CLASSIFIER nicht möglich. Die sehr ausgeprägte Dominanz von Sequenzen mit Homologie zu Staphylokokken im Vergleich zu dem rRNA-basierten taxonomischen Profil aus MG-RAST ist auf die unterschiedlichen Klassifizierungsmethoden zurückzuführen (vgl. Abbildung 5). Da die Klassifizierung mithilfe des RDP-CLASSIFIER ausschließlich auf Grundlage von 16S-rRNA-Sequenzen erfolgt, werden ausschließlich taxonomische Zuordnungen zu Bakterien und Archaeabakterien gemacht. Sequenzen aus Eukaryoten und Viren können auf diese Weise nicht klassifiziert werden. Da die Klassifizierung lediglich bis zur Gattungsebene hinreichend zuverlässig möglich ist, ist nicht feststellbar, ob es sich bei den Staphylokokken um *S. aureus* handelt und ob die Metatranskriptomik-Analyse mit dem mikrobiologischen Befund im Einklang steht.

### **1.3.3 Funktionale Taxonomie des Metatranskriptomik-Datensatzes aus dem Ohr mittels METRANS**

Die Beschreibung der funktionalen Taxonomie erfolgt in METRANS unter Verwendung der Sequenzen, die keine Treffer gegen die SILVA-Datenbanken *SSU*, *LSU* und *Rfam* liefern. Diese Sequenzen werden unter Verwendung des BLASTX- Algorithmus in den Datenbanken *Swissprot*, *KEGG* und *eggNOG* abgefragt. Die taxonomische Zuordnung der jeweiligen Treffer wird mithilfe des *Lowest-Common-Anccestor*-Ansatzes ausgewertet und auf die *NCBI*-Taxonomie gemappt. Bei der Darstellung der funktionalen Taxonomie in METRANS werden die Datensätze auf Grundlage des RPKM-Wertes normalisiert, indem unter anderem die Länge der Referenzsequenzen einbezogen wird.

In Abbildung 11 sind die abundantesten taxonomischen Zuordnungen der Treffer gegen die drei funktionalen Datenbanken kumulativ für alle taxonomischen Ebenen dargestellt. Bei Betrachtung der Anzahl der zugeordneten *Reads* aus den einzelnen Datenbanken zeigen sich deutliche Unterschiede sowohl zwischen den einzelnen taxonomischen Zuordnungen als auch zwischen den einzelnen taxonomischen Ebenen. Im Einklang zu der funktionalen Taxonomie in MG-RAST dominieren auf der Ebene des Reiches Eukaryoten und Bakterien (Abbildung 11 a; vgl. Abbildung 6). Die scheinbare Dominanz der Sequenzen mit Homologie zu eukaryotischen Sequenzen ist vermutlich teilweise auf die Normalisierung zurückzuführen, bei der die Länge der Referenzsequenz ebenfalls berücksichtigt wird. Auf der taxonomischen Ebene des Stammes dominieren die *Chordata*. Weitere Stämme werden durch die *Firmicutes*, *Ascomycota*, *Actinobacteria*, *Proteobacteria* sowie unklassifizierte Sequenzen beschrieben. Auf der Ebene der Klasse dominieren die *Mammalia* und *Bacilli*. Weitere Sequenzhomologien werden den

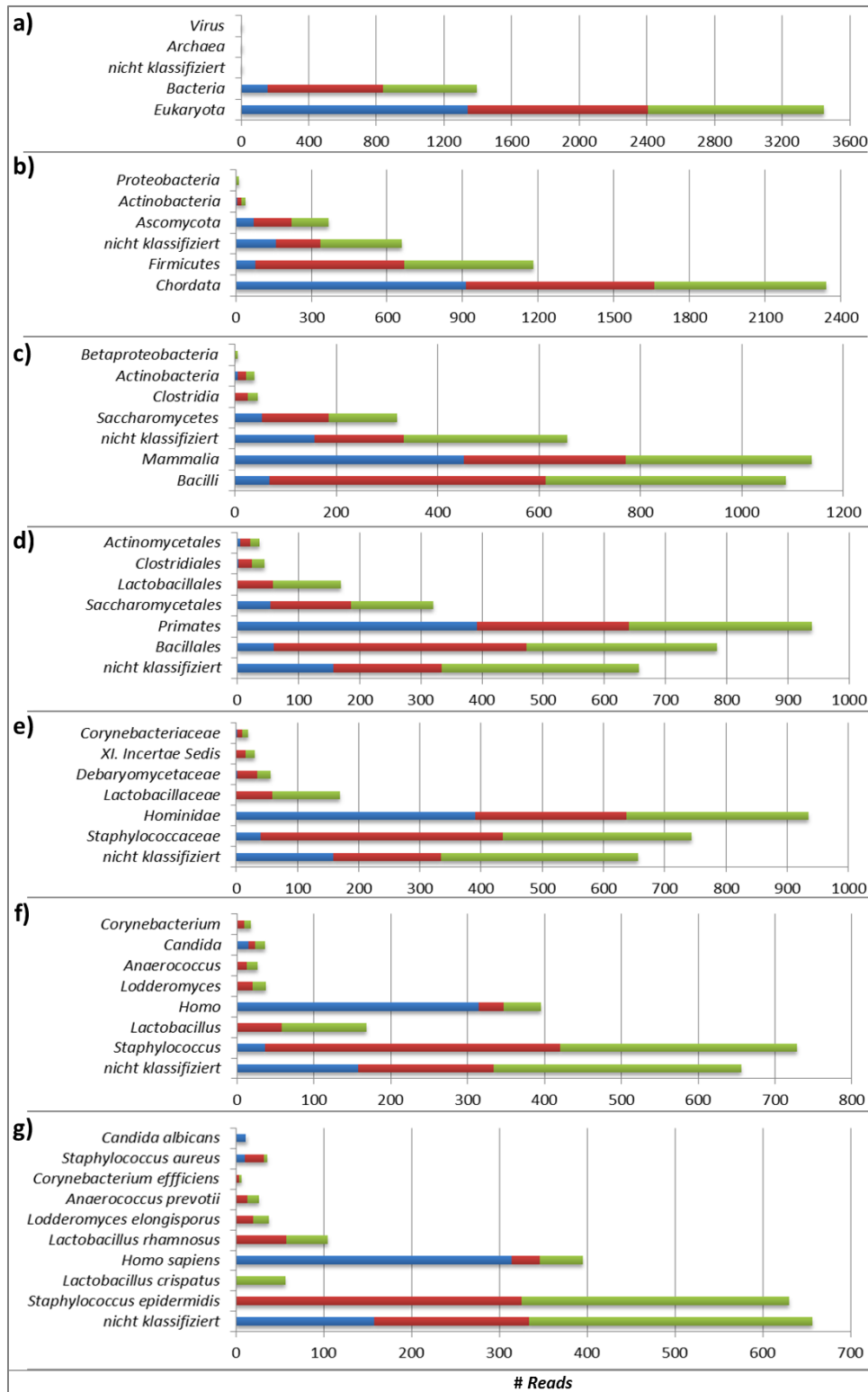


Abbildung 11: Funktionale Taxonomie der normalisierten Sequenzabundanzen mit Homologie zu proteinkodierenden Genen ermittelt durch Abfrage der Datenbanken *SwissProt* (blau), *eggNOG* (rot) und *KEGG* (grün) unter Verwendung des BLASTX-Algorithmus. Die normalisierten Sequenzhäufigkeiten der abundantesten Taxa sind kumulativ auf den folgenden taxonomischen Ebenen dargestellt: Reich (a), Stamm (b), Klasse (c), Ordnung (d), Familie (e), Gattung (f) und Art (g).

*Saccharomycetes*, *Clostridia*, *Actinobacteria* sowie nicht klassifizierten Sequenzen zugeordnet. Sequenzhomologien zu den *Primates*, *Bacillales* und zu nicht klassifizierten Sequenzen dominieren auf der taxonomischen Ebene der Ordnung. Entsprechend sind auf der Ebene der Familie Sequenzen mit Homologie zu den *Hominidae* und *Staphylococcaceae* sowie zu unbekanntem Sequenzen dominant. Auf der Genus-Ebene dominieren Sequenzen mit Homologie zu *Staphylococcus*, *Homo* und nicht klassifizierten Sequenzen. Auf der Spezies-Ebene dominieren Sequenzhomologien zu nicht klassifizierten Sequenzen, *Staphylococcus epidermidis* und *Homo sapiens*. Weiterhin werden die Spezies *Lactobacillus crispatus*, *Lactobacillus rhamnosus*, *Lodderomyces elongisporus*, *Anaerococcus prevotii*, *Corynebacterium efficiens*, *Staphylococcus aureus* und *Candida albicans* beschrieben. Dabei ist auffallend, dass die taxonomischen Zuordnungen zu *S. epidermidis* sowie zu *L. crispatus*, *L. rhamnosus*, *L. elongisporus*, *A. prevotii* und *C. efficiens* keine Treffer aus der *SwissProt*-Datenbank aufweisen.

Die Dominanz von Sequenzhomologien zu *S. epidermidis* stimmt weder mit den in der Fachliteratur beschriebenen Erkenntnissen noch mit dem mikrobiologischen Befund überein. Auch *Lactobacillus*-Spezies sind im Zusammenhang mit Gehörgangserkrankungen bislang nicht beschrieben worden.

#### **1.3.4 Funktionale Beschreibung des Metatranskriptoms des menschlichen Ohrs in METRANS**

Zur funktionalen Beschreibung der aktiven Mikrobiota werden in METRANS Ergebnislisten der Abfragen in der *SwissProt*-Datenbank ausgegeben. Als quantitatives Maß für die Häufigkeit der funktionalen Zuordnungen wird der RPKM-Wert verwendet.

In Tabelle 10 sind die aus *S. aureus* bekannten Virulenzfaktoren dargestellt, die Treffer bei der funktionalen Zuordnung basierend auf der *SwissProt*-Datenbank liefern. Insgesamt wurden Sequenzhomologien zu 31 Virulenzfaktoren aus sieben Gruppen identifiziert. Ähnlich wie in der entsprechenden Darstellung in MG-RAST werden auch in METRANS überwiegend Virulenzfaktoren im Zusammenhang mit Adhäsion und Biofilm-Bildung beschrieben (vgl. Tabelle 9). Dabei werden den Treffern zu den Proteinen *Clumping factor A*, *Clumping factor B* und *Elastin binding protein* die höchsten RPKM-Werte zugeordnet. Der RPKM-Wert von ClfA ist um einen Faktor von 10,5 größer als der entsprechende Wert für ClfB. Der Unterschied zwischen den entsprechend in MG-RAST identifizierten *Read*-Abundanzen liegt etwa bei einem Faktor von 3,3. Die RPKM-Werte aller übrigen in METRANS identifizierten Virulenzfaktoren liegen unter einem Wert von 1. Im Gegensatz zur Analyse in MG-RAST konnten in METRANS keine Sequenzhomologien zu den Funktionen *Immunglobulin G-binding protein* und *Immunglobulin G-binding protein Sbi*

Tabelle 10: Beschreibung der Virulenzfaktoren von *Staphylococcus aureus* aus dem Metatranskriptomik-Datensatz des menschlichen Ohres in METRANS auf Grundlage der Treffer gegen die SwissProt-Datenbank. Als quantitatives Maß für die funktionalen Annotationen dient der RPKM-Wert.

| Virulenzfaktoren-Gruppe               | Virulenzfaktor   | RPKM    |
|---------------------------------------|--|---------|
| Enterotoxine                          | <i>Enterotoxin type A</i>  | 0,2400  |
|                                       | <i>Enterotoxin type E</i>  | 0,1160  |
|                                       | <i>Enterotoxin type G</i>  | 0,0598  |
| Adhäsionsfaktoren                     | <i>Fibronectin-binding protein A</i>   | 0,1836  |
|                                       | <i>Serin-aspartate repeat protein C</i>                                      | 0,0809  |
|                                       | <i>Serin-aspartate repeat protein D</i>                                      | 0,0341  |
|                                       | <i>Serin-aspartate repeat protein E</i>                                      | 0,0134  |
|                                       | <i>Elastin-binding protein Ebps</i>  | 2,2888  |
|                                       | <i>Iron-regulated surface determinant A</i>                                  | 0,0439  |
|                                       | <i>Iron-regulated surface determinant B</i>                                  | 0,0345  |
|                                       | <i>Iron-regulated surface determinant H</i>                                  | 0,0240  |
|                                       | <i>Surface protein</i>   | 0,0094  |
|                                       | <i>Surface protein G</i>   | 0,0190  |
| Enzyme und andere Proteine            | <i>Serine-rich adhesin for platelets</i>                                     | 0,0989  |
|                                       | <i>Extracellular matrix protein-binding protein emp</i>                      | 0,0454  |
| Resistenz gegen oxidativen Burst      | <i>Staphylocoagulase</i>   | 0,0450  |
|                                       | <i>O-acetyltransferase OatA</i>  | 0,2559  |
| Komplement-Inaktivierung              | <i>Catalase CatA</i>   | 0,0311  |
|                                       | <i>Capsular polysaccharide type 5 biosynthesis protein Cap5A</i>             | 0,4170  |
|                                       | <i>Capsular polysaccharide biosynthesis protein CapA</i>                     | 0,0698  |
|                                       | <i>Staphylococcal complement inhibitor</i>                                   | 0,2660  |
|                                       | <i>Staphylokinase</i>  | 0,0946  |
|                                       | <i>Clumping factor A</i>   | 16,489  |
| Umgehung der angeborenen Immunantwort | <i>Clumping factor B</i>   | 1,5741  |
|                                       | <i>Protein DltD</i>  | 0,03937 |
| Antibiotika-Resistenzen               | <i>Zinc metalloproteinase aureolysin</i>                                     | 0,03031 |
|                                       | <i>Transposase for insertion sequence-like element IS431mec</i>              | 0,06889 |
|                                       | <i>Multidrug resistance ABC transporter ATP-binding and permease protein</i> | 0,0615  |
|                                       | <i>Tetracycline resistance protein</i>                                       | 0,02415 |
|                                       | <i>Beta-lactamase</i>  | 0,04930 |

---

identifiziert werden. Jedoch wurden in METRANS Sequenzhomologien zu den Funktionen *Staphylocoagulase*, *Capsular polysaccharide biosynthesis protein CapA*, *Protein DltD* aus den funktionalen Gruppen der Komplement-Inaktivierung und der Umgehung der angeborenen Immunantwort beschrieben. Sequenzhomologien zu Funktionen, die im Zusammenhang mit der Immunabwehr des Menschen stehen, konnten ebenfalls mithilfe der Abfrage der *SwissProt*-Datenbank dargestellt werden. Die RPKM-Werte der Zuordnungen zu den *Toll-like*-Rezeptoren 1, 2, 4, 6, 8 und 10 lagen jeweils bei Werten von unter 1. Weiterhin konnten Zuordnungen zu Proteinen der *C-type domain family* gefunden werden. Sequenzen mit Homologie zu Proteasen und Phospholipasen von *Candida* spp. konnten nicht analysiert werden, da die Zuordnung einer funktionalen Annotation zu einem ausgewählten Taxon in METRANS nicht durchführbar ist.



## 2 Metatranskriptomik der menschlichen Achselhöhlen-Mikrobiota

Im Rahmen dieser Arbeit wurde ein Laborprotokoll entwickelt, das die Metatranskriptom-Analyse von Proben aus der menschlichen Achselhöhle ermöglicht. Im Folgenden wird zunächst die in Kooperation mit dem industriellen Partner durchgeführte Studie sowie das eigens für diesen Zweck entwickelte Laborprotokoll dargestellt. Danach wird die Sequenzdaten-Prozessierung mit verschiedenen Bioinformatik-Tools sowie mittels MG-RAST dargestellt. Anschließend wird die inhaltliche Analyse der Daten vorgestellt, mit Fokus auf die Zusammensetzung der Metatranskriptomik-Datensätze auf taxonomischer Ebene, basierend auf rRNA- und auf funktionalen Annotationen. Schließlich werden funktionale Sequenzen in den Datensätzen beschrieben, die im Zusammenhang mit Körpergeruchsentstehung stehen.

### 2.1 Etablierung einer Labormethode für die Metatranskriptomik-Analyse von Probenmaterial aus der menschlichen Achselhöhle

#### 2.1.1 Ergebnisse der Deo-Studie

Im Rahmen der in Kapitel III.1.2 beschriebenen Deo-Studie wurden vom industriellen Kooperationspartner insgesamt 96 Proben von 24 männlichen Probanden zur Verfügung gestellt. Jeder Proband wurde insgesamt viermalig an zwei Tagen im Abstand von etwa 19 Tagen beprobt. Als Kontrollbehandlung zum Deodorant (Deo) diente Wasser ( $H_2O$ ). Die Probennahme erfolgte jeweils unmittelbar vor der Behandlung ( $t_0$ ) und sechs Stunden nach der Behandlung ( $t_6$ ). Jeweils 1 ml einer jeden Probe wurde für die Metatranskriptomik zur Verfügung gestellt. Das restliche Probenmaterial wurde unter anderem für die Bestimmung der koloniebildenden Einheiten verwendet.

Die Anzahl der koloniebildenden Einheiten aller Proben lag im Bereich von  $10^3 - 10^7$  KBE/ml mit einem Mittelwert von  $4,17 \times 10^5$  KBE/ml (s. Tabelle 11). Die mittlere Keimzahl erhöhte sich nach Behandlung mit Wasser nur geringfügig (von  $4,07 \times 10^5$  KBE/ml vor Behandlung auf  $4,36 \times 10^5$  KBE/ml nach Behandlung). Dahingegen kam es nach Deodorant-Behandlung zu einer Reduktion im Mittel um etwa eine Zehnerpotenz von  $7,12 \times 10^5$  KBE/ml auf  $8,36 \times 10^4$  KBE/ml. Dieser Effekt ist ebenfalls in den Mittelwerten der einzelnen Monate Mai und Juni zu beobachten (s. Tabelle 11). Auffallend ist, dass die mittleren  $t_0$ -Werte der Keimzahlen für die Probennahmen im Mai ( $7,15 \times 10^5$  KBE/ml) um den Faktor 1,77 höher sind als für die Probennahmen im Juni ( $4,04 \times 10^5$  KBE/ml). Generell wurden im Mittel jedoch nur geringfügig höhere Keimzahlen im Mai gemessen als im Juni ( $4,88 \times 10^5$  und  $3,47 \times 10^5$ ). Insgesamt variieren die gemessenen Keimzahlen sehr, so

dass die Standardabweichungen z. T. höhere Werte annehmen als die zugehörigen Mittelwerte. Dies deutet auf eine sehr große Variabilität der Achselhöhlen-Mikrobiota zwischen den einzelnen Probanden sowie zwischen den einzelnen Proben einzelner Probanden hin.

**Tabelle 11:** Übersicht der im Rahmen der Deo-Studie beobachteten Ergebnisse. Dargestellt sind jeweils Mittelwerte und Standardabweichungen für die beschriebenen Keimzahlen [KBE/ml] und *Malodor Scores* [1-4] für die jeweiligen Probennahme-Zeitpunkte  $t_0$  H<sub>2</sub>O,  $t_6$  H<sub>2</sub>O,  $t_0$  Deo,  $t_6$  Deo aufgeschlüsselt nach den jeweiligen Probennahmen im Mai und Juni sowie für alle Probennahmen (Mai und Juni).

| Probennahme<br>[Monat] | Proben<br>[Behandlung]   | Keimzahl [KBE/ml] |                 | Malodor Score [0-4] |              |
|------------------------|--------------------------|-------------------|-----------------|---------------------|--------------|
|                        |                          | Mittelwert        | Standardabw.    | Mittelwert          | Standardabw. |
| Mai                    | $t_0$ H <sub>2</sub> O   | 3,36E+05          | 5,62E+05        | 1,18                | 0,70         |
|                        | $t_6$ H <sub>2</sub> O   | 3,76E+05          | 5,22E+05        | 1,58                | 1,03         |
|                        | $t_0$ Deo                | 1,09E+06          | 2,88E+06        | 0,89                | 0,78         |
|                        | $t_6$ Deo                | 1,46E+05          | 1,94E+05        | 0,27                | 0,42         |
|                        | $t_0$ Mai                | 7,15E+05          | 2,07E+06        | 1,04                | 0,74         |
|                        | $t_6$ Mai                | 2,61E+05          | 4,02E+05        | 0,92                | 1,02         |
|                        | <b>Gesamt Mai</b>        | <b>4,88E+05</b>   | <b>1,51E+06</b> | <b>0,98</b>         | <b>0,88</b>  |
| Juni                   | $t_0$ H <sub>2</sub> O   | 4,78E+05          | 8,14E+05        | 0,84                | 0,70         |
|                        | $t_6$ H <sub>2</sub> O   | 4,96E+05          | 9,16E+05        | 1,21                | 0,56         |
|                        | $t_0$ Deo                | 3,30E+05          | 3,30E+05        | 0,60                | 0,36         |
|                        | $t_6$ Deo                | 8,36E+04          | 1,07E+05        | 0,35                | 0,28         |
|                        | $t_0$ Juni               | 4,04E+05          | 6,12E+05        | 0,72                | 0,56         |
|                        | $t_6$ Juni               | 2,90E+05          | 6,72E+05        | 0,78                | 0,62         |
|                        | <b>Gesamt Juni</b>       | <b>3,47E+05</b>   | <b>6,39E+05</b> | <b>0,75</b>         | <b>0,58</b>  |
| Mai und Juni           | $t_0$ H <sub>2</sub> O   | 4,07E+05          | 6,88E+05        | 1,01                | 0,70         |
|                        | $t_6$ H <sub>2</sub> O   | 4,36E+05          | 7,32E+05        | 1,40                | 0,84         |
|                        | $t_0$ Deo                | 7,12E+05          | 2,04E+06        | 0,60                | 0,35         |
|                        | $t_6$ Deo                | 8,36E+04          | 1,56E+05        | 0,31                | 0,35         |
|                        | $t_0$ Mai + Juni         | 5,60E+05          | 1,52E+06        | 0,88                | 0,67         |
|                        | $t_6$ Mai + Juni         | 2,75E+05          | 5,48E+05        | 0,85                | 0,84         |
|                        | <b>Gesamt Mai + Juni</b> | <b>4,17E+05</b>   | <b>1,14E+06</b> | <b>0,87</b>         | <b>0,75</b>  |

Vor jeder Probennahme erfolgte die semiquantitative Bestimmung des *Malodor Score* durch Riechen und Klasseneinteilung des Geruches von 0 bis 4 durch jeweils 4 *Sniffer*. Der mittlere *Malodor Score* aller Proben lag mit 0,87 deutlich im unteren Bereich der möglichen Klasseneinteilungen (s. Tabelle 11). Besonders auffällig ist jedoch der Höchstwert von 3,5 für den Probanden 7 nach Behandlung mit Wasser. Für beide Probennahme-Monate kann beobachtet werden, dass der mittlere *Malodor Score* sechs Stunden nach Behandlung mit Wasser um etwa 39 % erhöht wird (von 1,18 auf 1,58 im Mai und von 0,84 auf 1,21 im Juni). Die Deo-Behandlung führt hingegen zu beiden Zeitpunkten zu einer Herabsenkung des mittleren *Malodor Score* um einen Faktor von 2 bzw. 3 (von 0,89 auf 0,27 im Mai und von 0,60 auf 0,35 im Juni). Ein ähnlicher Effekt ist auch für die über alle Proben gemittelten Werte in Tabelle 11 zu beobachten. Analog zum entsprechenden Phänomen bei den Keimzahlen (s. oben) ist der mittlere *Malodor Score* der  $t_0$ -Werte im Mai (1,04) um den Faktor 1,44 höher als der entsprechende Wert im Juni (0,72). Die zu den jeweiligen Mittelwerten berechneten Standardabweichungen sind ebenfalls sehr hoch und übersteigen z. T. die Mittelwerte, so dass auch für den Körpergeruch eine hohe inter- und intraindividuelle Variabilität zu beobachten ist (s. Tabelle 11).

Insgesamt ist weder für die  $t_0$ -Werte im Mai noch für die im Juni ein Zusammenhang zwischen der Keimzahl und dem Körpergeruch erkennbar. Dies bestätigt die Darstellung in der Fachliteratur und deutet darauf hin, dass spezielle Organismen mit der Entstehung von Körpergeruch im Zusammenhang stehen. Dieser Zusammenhang soll mithilfe des im Rahmen dieser Arbeit entwickelten Metatranskriptomik-Ansatzes im Folgenden detailliert untersucht werden.

### **2.1.2 Laborprotokoll für die Metatranskriptomik-Analyse der Achselhöhlen-Mikrobiota**

Die Untersuchung der Proben aus der menschlichen Achselhöhle stellte aufgrund der sehr geringen Zelldichte sowie aufgrund des sehr geringen Probenvolumens eine große Herausforderung an die Labormethode dar. Daher handelte es sich im Hinblick auf die Machbarkeit der Metatranskriptom-Analyse bei der RNA-Isolierung aus den Proben um den kritischen Arbeitsschritt, nach dem entschieden wurde, ob die jeweiligen Proben weiterführend analysiert werden konnten. Im Folgenden werden das entwickelte Laborprotokoll sowie die Proben, auf die es angewendet werden konnte, beschrieben (s. Tabelle 12 und Abbildung 12).

Die RNA-Isolierung wurde säulenbasiert mithilfe des *RNeasy Micro Kit* von *Qiagen* durchgeführt. Die spektrophotometrische qualitative und quantitative Analyse der isolierten RNA mithilfe des *NanoDrop* lieferte aufgrund der sehr geringen RNA-Mengen keine auswertbaren Ergebnisse. Daher wurde die quantitative und qualitative Analyse elektrophoretisch mittels *Agilent*

Bioanalyzer auf einem RNA Pico Chip durchgeführt (s. Abbildung 12a). Diese geringen RNA-Mengen erforderten ähnlich wie bei der Ohr-Probe zunächst eine cDNA-Synthese und eine Amplifikation bevor die Herstellung von cDNA-Sequenzierbibliotheken erfolgen konnte. Für die Anwendung des Metatranskriptomik-Protokolles konnten nur die Proben verwendet werden, die genügend RNA für den darauffolgenden Arbeitsschritt der cDNA-Synthese und Amplifikation

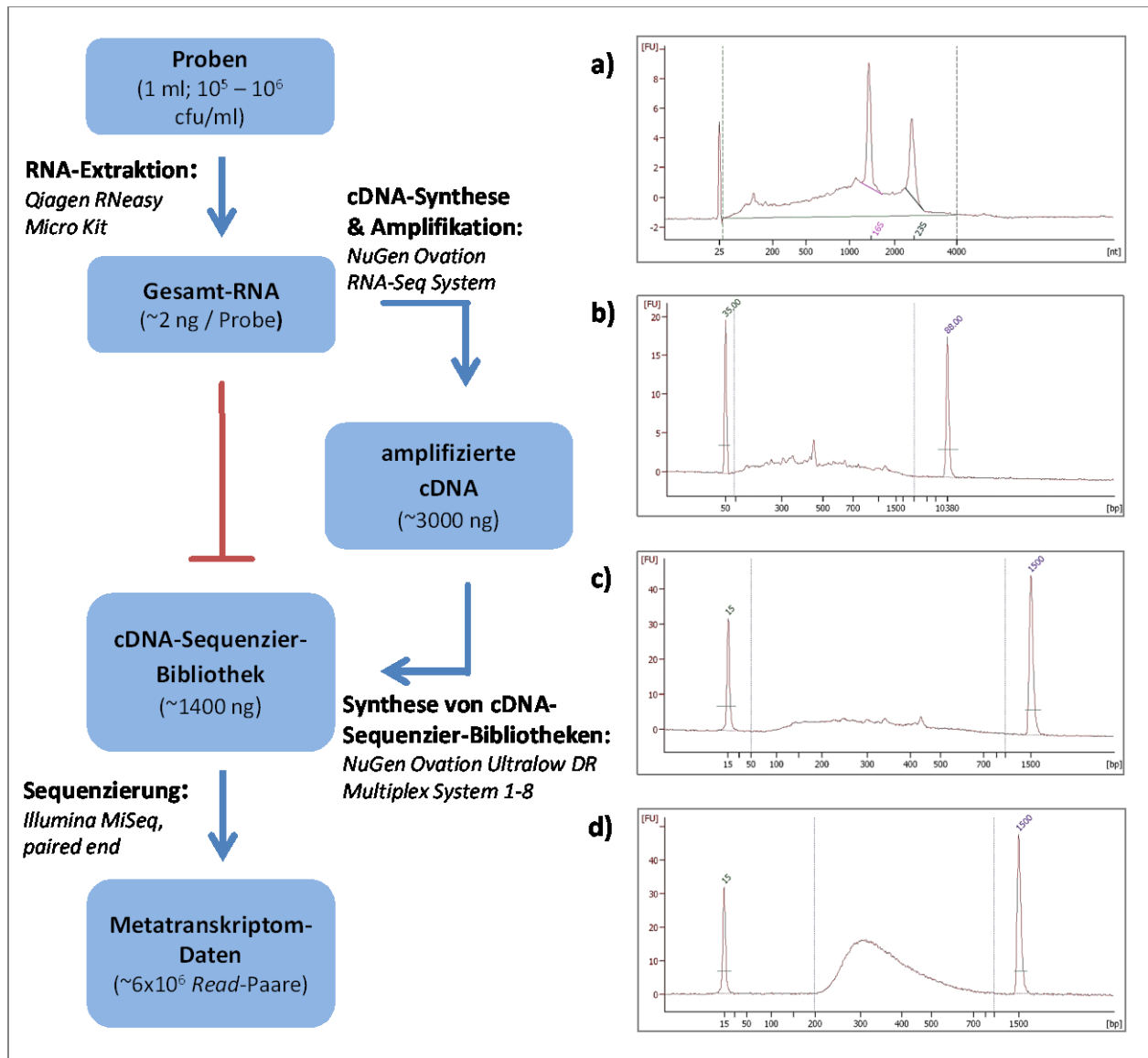


Abbildung 12: Fließdiagramm der Metatranskriptom-Analyse von Proben der Achselhöhlen-Mikrobiota. Ausgehend von den Proben aus der Deo-Studie wurde eine RNA-Isolierung durchgeführt. Aufgrund der geringen RNA-Mengen wurde cDNA synthetisiert und anschließend amplifiziert um die Herstellung von cDNA-Sequenzierbibliotheken und die *High-Throughput*-Sequenzierung zu ermöglichen. Dargestellt sind zudem Elektropherogramme der Analyse der Probe 48n. a) Gesamt-RNA, b) amplifizierte cDNA, c) fragmentierte cDNA, d) cDNA-Sequenzierbibliothek.

**Tabelle 12: Übersicht des für die Metatranskriptomik verwendeten Probenmaterials der Achselhöhle. Dargestellt sind 14 Proben von sechs Probanden mit den jeweiligen Zeitpunkten der Probennahme ( $t_0$  und  $t_6$ ), der jeweiligen Behandlung ( $H_2O$  und Deo), der *Malodor Score* sowie der Anzahl der koloniebildenden Einheiten.**

| Proband           | Probe | Probennahme | <i>Malodor Score</i> [0-4] | Keimzahl [KBE/ml]                    |
|-------------------|-------|-------------|----------------------------|--------------------------------------|
| 6                 | 6n    | $t_6 H_2O$  | 2,375                      | $2,81 \times 10^5$                   |
| 7                 | 7n    | $t_6 H_2O$  | 3,5                        | $9,76 \times 10^5$                   |
| 13                | 13n   | $t_6 Deo$   | 0                          | $1,91 \times 10^5$                   |
|                   | 37n   | $t_6 H_2O$  | 1,875                      | $3,19 \times 10^5$                   |
| 24                | 48v   | $t_0 H_2O$  | 0,75                       | $9,60 \times 10^5$                   |
|                   | 48n   | $t_6 H_2O$  | 1,25                       | $8,40 \times 10^5$                   |
| 18                | 18v   | $t_0 Deo$   | 1,125                      | $1,02 \times 10^7$                   |
|                   | 18n   | $t_6 Deo$   | 0                          | $6,71 \times 10^5$                   |
|                   | 42v   | $t_0 H_2O$  | 0,875                      | $3,03 \times 10^5$                   |
|                   | 42n   | $t_6 H_2O$  | 1,125                      | $1,53 \times 10^5$                   |
| 19                | 19v   | $t_0 Deo$   | 1,875                      | $6,11 \times 10^5$                   |
|                   | 19n   | $t_6 Deo$   | 0,25                       | $1,29 \times 10^4$                   |
|                   | 43v   | $t_0 H_2O$  | 1,875                      | $4,39 \times 10^5$                   |
|                   | 43n   | $t_6 H_2O$  | 1,5                        | $3,03 \times 10^5$                   |
| <b>Mittelwert</b> | -     | -           | <b>1,3125</b>              | <b><math>1,16 \times 10^6</math></b> |

lieferten (Tabelle 12). Zur Anwendung des entsprechenden Kit-Systems musste eine RNA-Menge von mindestens 500 pg RNA pro Probe vorliegen. Diese Anforderung wurde von insgesamt 14 Proben von sechs unterschiedlichen Probanden erfüllt (Tabelle 12 und Tabelle 13). Typische Werte für die Keimzahlen dieser Proben lagen im Bereich von  $10^5$  KBE/ml mit Ausnahme von Probe 18v ( $1,02 \times 10^7$  KBE/ml) und Probe 19n ( $1,29 \times 10^4$  KBE/ml). Für die Probanden 18 und 19 konnten jeweils alle vier Proben ( $t_0 H_2O$ ,  $t_6 H_2O$ ,  $t_0 Deo$  und  $t_6 Deo$ ) mit dem entwickelten Metatranskriptomik-Protokoll analysiert werden. Für beide Probensätze kann beobachtet werden, dass sowohl der *Malodor Score* als auch die Keimzahl nach Behandlung mit Deodorant deutlich abfällt. Die Keimzahl wird dabei im Falle des Probanden 18 um zwei Zehnerpotenzen und im Falle des Probanden 19 um eine Zehnerpotenz reduziert. Die *Malodor Scores* liegen nach Behandlung mit Deodorant bei 0 bzw. 0,25. Nach Behandlung mit Wasser finden in beiden Fällen geringere Veränderungen statt. Weiterhin wurden je zwei Proben der Probanden 24 ( $t_0 H_2O$  und  $t_6 H_2O$ ) und 13 ( $t_6 H_2O$  und  $t_6 Deo$ ), sowie je eine Probe der Probanden 6 ( $t_6 H_2O$ ) und 7

(t<sub>6</sub> H<sub>2</sub>O) für die Metatranskriptomik-Analysen ausgewählt. Besonders auffällig ist der sehr hohe *Malodor Score* von 3,5 für die Probe 7n des Probanden 7 (vgl. Kapitel 2.1.1). Aufgrund der starken Reduktion der Keimzahlen nach Behandlung mit Deodorant und der damit einhergehenden geringen RNA-Ausbeute konnten insgesamt nur drei Proben nach Deo-Behandlung (13n, 18n und 19n) im Rahmen der Metatranskriptom-Analysen betrachtet werden. Die erzielten Ausbeuten an Gesamt-RNA lagen im Mittel bei etwa 2 ng, mit einem Minimum von 420 pg für Probe 19n und einem Maximum von 12,15 ng für Probe 48n (Tabelle 13). Die Qualität der isolierten RNA wurde anhand der RIN (*RNA Integrity Number*) im Bioanalyser bestimmt und variierte stark in Abhängigkeit von der Probe (Tabelle 13). Die höchste im Rahmen dieser Studie erzielte RIN lag bei einem Wert von 8 (Probe 7n). Für die Probe 43n konnte keine RIN bestimmt werden. In Abbildung 12 a) ist beispielhaft das Elektropherogramm der Gesamt-RNA der Probe 48n mit einer RIN von 6,3 dargestellt. Dabei sind die Peaks für die 16S- und die 23S-rRNA gut aufgelöst und andere RNAs mit einer Längenverteilung im Bereich von 25 nt bis etwa 4000 nt sind erkennbar.

Für die Synthese und Amplifikation von cDNA aus der isolierten Gesamt-RNA der 14 Proben wurde das *NuGen Ovation RNA-Seq System* verwendet. Das *Priming* der Synthese des ersten und zweiten cDNA-Stranges erfolgte dabei mit *Random-Primern*. Die *Primer* lagerten sich dabei zufällig und unspezifisch an Bereiche der Transkripte an und dienten als Startpunkt der cDNA-Synthese, was in einer sehr breiten Längenverteilung der doppelsträngigen cDNA im Bereich von 50 bp bis 10.380 bp resultierte (Abbildung 12 b). Anschließend erfolgte eine isothermale Amplifikation der cDNA, wobei die 5'-Überhänge der *Random-Primer* als *Priming Sites* für die Amplifikation fungieren. Auf diese Weise konnten im Mittel etwa 3 µg cDNA synthetisiert werden (Abbildung 12, Tabelle 13).

Für die Herstellung der cDNA-Sequenzierbibliotheken wurde die cDNA zunächst mithilfe von Scherkräften im Ultraschall-Bad auf eine Länge von 100 bp bis 700 bp fragmentiert (s. Abbildung 12 c). Die Synthese der Sequenzierbibliotheken wurde anschließend unter Verwendung des *NuGen Ovation Ultralow DR Multiplex System 1-8* durchgeführt und lieferte im Mittel 1400 ng Material mit einer Längenverteilung von 200 bp bis 700 bp und einem Maximum bei 300 bp (s. Abbildung 12 d, Tabelle 13). Bei diesem Schritt wurden die für die *Illumina*-Sequenzierung benötigten *Tru-Seq*-Adaptersequenzen eingefügt. Weiterhin wurde jede Probe durch einen *Barcode*, eine individuelle Erkennungssequenz bestehend aus sechs Nukleotiden innerhalb des *Reverse-Primers*, gekennzeichnet (Tabelle 13). Dies ermöglichte die gleichzeitige Sequenzierung von Proben mit unterschiedlichem *Barcode* in einem Sequenzierlauf und die anschließende Zuordnung der Sequenzen zu den jeweiligen Proben anhand der *Barcode*-Sequenz.

**Tabelle 13: Übersicht der Daten zum Laborprotokoll der Metatranskriptomik der Achselhöhlen-Mikrobiota. Dargestellt ist für jede Probe die Menge isolierter Gesamt-RNA und die zugehörige RNA Integrity Number (RIN), die Menge an cDNA nach Amplifikation, die Menge der synthetisierten cDNA-Bibliothek und der Barcode für jede Probe.**

| Probe             | Gesamt-RNA<br>[ng] | RIN<br>[1-10] | Gesamt-cDNA<br>[ng] | cDNA-Sequenzierbibliothek<br>[ng] | Barcode |
|-------------------|--------------------|---------------|---------------------|-----------------------------------|---------|
| 6n                | 0,522              | 1,1           | 2.538               | 1.307                             | 8       |
| 7n                | 2,210              | 8             | 3.807               | 1.458                             | 2       |
| 13n               | 1,010              | 1             | 1.686               | 1.157                             | 5       |
| 37n               | 2,760              | 7,3           | 3.913               | 1.461                             | 1       |
| 48v               | 3,570              | 5,7           | 4.098               | 1.456                             | 3       |
| 48n               | 12,150             | 6,3           | 3.427               | 1,526                             | 4       |
| 18v               | 2,930              | 5,9           | 2.366               | 1.626                             | 3       |
| 18n               | 0,630              | 3,1           | 2.163               | 1.393                             | 5       |
| 42v               | 0,558              | 1             | 2.055               | 2.040                             | 7       |
| 42n               | 1,690              | 7             | 4.368               | 1.300                             | 6       |
| 19v               | 2,520              | 6,3           | 2.627               | 2.031                             | 4       |
| 19n               | 0,420              | 1,4           | 2.263               | 1.610                             | 6       |
| 43v               | 0,690              | 1,3           | 4.508               | 1.135                             | 1       |
| 43n               | 0,650              | n. d.         | 3.759               | 1.435                             | 2       |
| <b>Mittelwert</b> | <b>2,31</b>        | -             | <b>3.113</b>        | <b>1.495</b>                      | -       |

Besonders auffallend ist, dass trotz der sehr großen Unterschiede hinsichtlich der Qualität und Menge an Gesamt-RNA der Proben die Quantität und Qualität der cDNA-Sequenzierbibliotheken sehr homogen ist (Tabelle 13) Die Ergebnisse der Sequenzierung sowie die Bearbeitung der Sequenzierdaten mithilfe von Bioinformatik-Tools und der *Pipeline* MG-RAST wird im folgenden Kapitel beschrieben.

## 2.2 Bioinformatik-Analyse der Sequenzdaten mittels MG-RAST

Die Sequenzierung der cDNA-Bibliotheken der Achselhöhlen-Mikrobiota erfolgte auf einer *Illumina-MiSeq*-Plattform im *Paired-End*-Modus mit einer Leseweite von  $2 \times 250$  Basen (Abbildung 12). Dabei wurden jeweils die Proben 7n, 13n, 37n, 48v, 48n und 42n gemeinsam in

drei Durchgängen und die Proben 6n, 18v, 18n, 42v, 19v, 19n, 43v und 43n gemeinsam in vier Durchgängen sequenziert. Im Mittel wurden für jede Probe etwa 6 Millionen *Read*-Paare generiert (Abbildung 12, Tabelle 14 ). Um die zusammengehörigen *Read*-Paare der *Paired-End*-Sequenzierung zu verbinden (*Read Merging*) und die Adapter-Sequenzen an den *Read*-Enden zu entfernen (*Adaptor Trimming*) wurde das Bioinformatik-Tool SEQPREP verwendet. Insgesamt konnten im Mittel etwa 95 % der *Reads* zusammengefügt werden (Tabelle 14).

**Tabelle 14:** Übersicht der Sequenzdatenbearbeitung mit SEQPREP. Dargestellt ist für jede Probe die Anzahl der *Read*-Paare aus den Sequenzierdaten, die Anzahl der mithilfe von SEQPREP zusammengefügt *Reads* und die jeweiligen Mittelwerte über alle Proben.

| Proband           | Probe | Sequenzierdaten [# <i>Read</i> -Paare] | SeqPrep Output [# <i>Reads</i> ] |
|-------------------|-------|--|----------------------------------|
| 6                 | 6n    | 7.318.943                              | 6.921.702                        |
| 7                 | 7n    | 5.579.679                              | 5.565.871                        |
| 13                | 13n   | 4.542.026                              | 4.535.700                        |
|                   | 37n   | 4.883.820                              | 4.874.757                        |
| 24                | 48v   | 4.720.913                              | 4.708.049                        |
|                   | 48n   | 4.968.348                              | 4.959.168                        |
| 18                | 18v   | 7.375.815                              | 6.986.061                        |
|                   | 18n   | 6.496.806                              | 6.066.352                        |
|                   | 42v   | 8.098.784                              | 7.575.779                        |
|                   | 42n   | 5.482.504                              | 5.473.010                        |
| 19                | 19v   | 7.179.182                              | 6.662.466                        |
|                   | 19n   | 7.134.245                              | 6.502.162                        |
|                   | 43v   | 6.311.528                              | 5.753.660                        |
|                   | 43n   | 7.644.753                              | 7.132.499                        |
| <b>Mittelwert</b> |       | <b>6.266.953</b>                       | <b>5.979.803 (95 %)</b>          |

Nach dem *Read Merging* und dem Adapter *Trimming* in SEQPREP wurden Qualitätsstatistiken für jeden Datensatz mit dem FASTX-Toolkit erstellt und *Quality-Plots* generiert. In Abbildung 13 ist beispielhaft der *Quality-Plot* der Probe 48n dargestellt. Zu beobachten ist der für eine *Illumina*-Sequenzierung charakteristische leichte Anstieg des Qualitätswertes im vorderen (Basen 1-100)



und der leichte Abfall im hinteren Bereich der Sequenz (Basen 450-500). Besonders auffällig ist der abrupte Abfall des mittleren Qualitätswertes von einem *Phred-Score* von etwa 60 auf einen *Phred-Score* von etwa 35 bei der Basenposition 250. Der Grund hierfür liegt in der Leseweite eines *Reads* von 250 Basen. Sequenzen mit einer Länge von bis zu 250 Basen werden auf diese Weise im *Paired-End-Modus* doppelt sequenziert und anschließend zusammengefügt, was letztlich in einer sehr geringen Fehlerrate und einem entsprechend hohen *Quality-Score* resultiert. Alle Basenpositionen von der Position 250 an aufwärts werden lediglich einmalig sequenziert. Insgesamt befindet sich der mittlere Qualitätswert an jeder Basenposition oberhalb eines *Phred-Score* von 30, was auf eine mittlere Fehlerrate von maximal 0,001 schließen lässt.

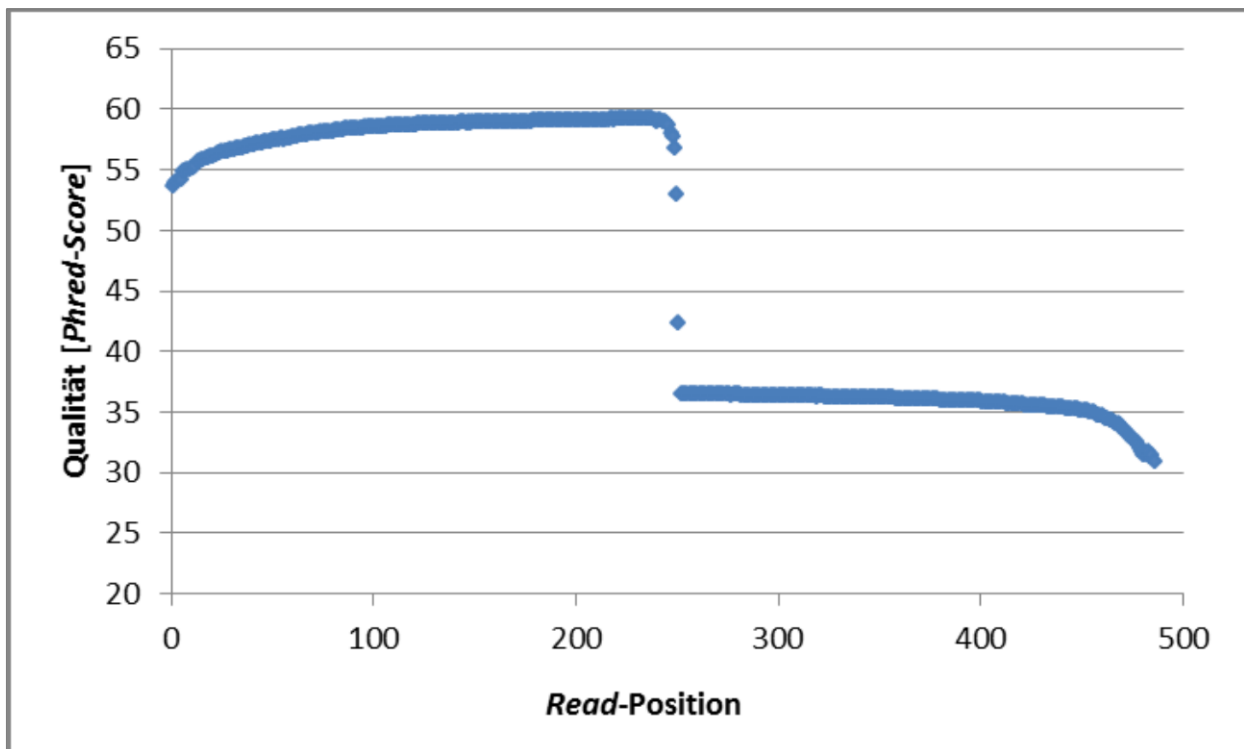


Abbildung 13: *Quality-Plot* des Metatranskriptomik-Datensatzes von Probe 48n. Dargestellt ist die mittlere Qualität als *Phred-Score* in Abhängigkeit von der Basen-Position im *Read* nach Bearbeitung der Sequenzen mit SEQPREP und Erstellung von Qualitätsstatistiken mithilfe des FASTX-Toolkit.

Die mit SEQPREP prozessierten Daten standen schließlich für den *Upload* in MG-RAST zur Verfügung. Die Datensätze wurden unter der Projektbezeichnung *Metatranscriptome axilla* (Projekt-ID 7812) hinterlegt. Die Kennzahlen der Sequenzdatenanalyse in der MG-RAST-*Pipeline* sind in Abbildung 14 dargestellt. Aufgrund der sehr guten Qualität der Sequenzdaten wurden im ersten Schritt der *Pipeline*, dem *Preprocessing*, im Mittel lediglich 5.979 Sequenzen (0,1 %) entfernt. Da der Prozessierungsschritt *Dereplication* in der *Pipeline* deaktiviert wurde, wurden an dieser Stelle keine weiteren Sequenzen entfernt. In dem darauffolgenden *Screening*-Schritt

wurden im Mittel lediglich 544 Sequenzen mit Ähnlichkeit zum menschlichen Genom entfernt. Diese relativ geringe Zahl (0,009 %) zeigt, dass das Probenmaterial überwiegend mit Organismen der Haut-Mikrobiota angereichert war und dass die Methode der Probennahme für Metatranskriptomik-Analysen der Achselhöhle sehr gut geeignet ist.

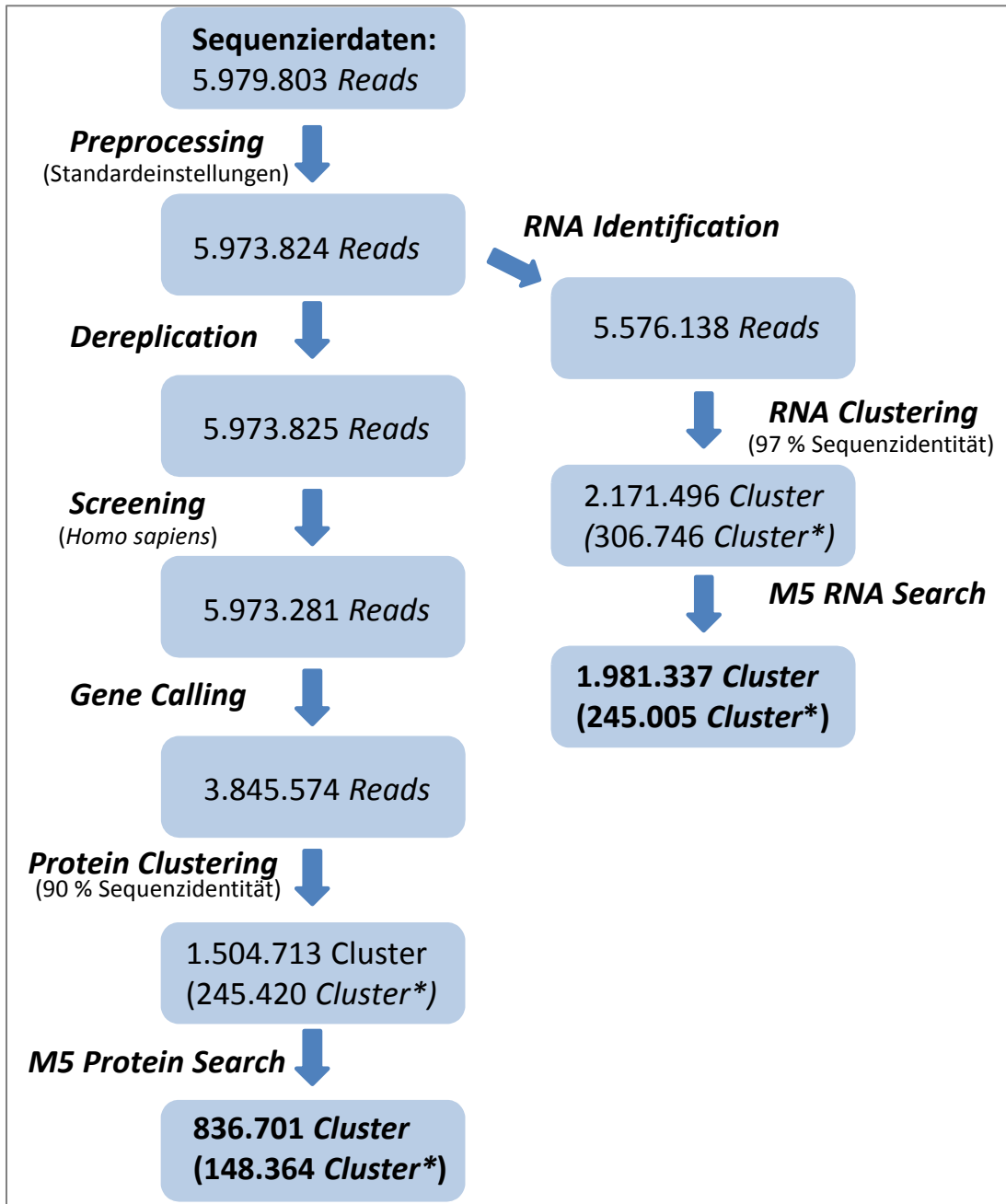


Abbildung 14: Übersicht der Sequenzdatenanalyse der Metatranskriptomik-Datensätze aus der Achselhöhle mittels MG-RAST. Dargestellt ist jeweils die über alle Proben gemittelte absolute Anzahl der Reads oder Cluster, die nach den jeweiligen Prozessierungsschritten für die weitere Analyse zur Verfügung stehen. Neben den Pfeilen werden die jeweiligen Analyseschritte beschrieben. Mit \* gekennzeichnete Zahlen beziehen sich nur auf die Datensätze der Proben 6n, 18v, 18n, 42v, 19v, 19n, 43v und 43n.

Im Rahmen des *Gene Calling* innerhalb der *Protein-Pipeline* wurden im Mittel 3.485.574 potentiell proteinkodierende *Reads* identifiziert. Nach dem *Protein Clustering* standen im Mittel 1.504.713 Cluster für die Datenbankabfrage zur Verfügung. In diesem Fall stellen jedoch die Datensätze der Proben 7n, 13n, 37, 48v, 48n und 42n eine Besonderheit dar, da das *Clustering* in diesen Fällen fälschlicherweise nicht durchgeführt wurde. Gemittelt über die Datensätze der Proben 6n, 18v, 18n, 42v, 19v, 19n, 43v und 43n, bei denen das *Clustering* korrekt durchgeführt wurde, lassen sich 245.420 Cluster errechnen, was einer Datenreduktion von einem Faktor von etwa 16 für den *Clustering*-Schritt entspricht. Bei einer Abfrage in der *M5NR*-Datenbank konnten im Mittel insgesamt 836.701 Cluster bestimmten Funktionen zugeordnet werden. Im Falle der korrekt geclusterten Datensätze konnten 148.364 Cluster zugeordnet werden. Das nicht durchgeführte *Clustering* für die oben genannten Datensätze ging mit einer deutlich erhöhten Rechendauer einher, besitzt jedoch für die weitere Auswertung der Datensätze keine Relevanz, da diese sich auf *Read*-Abundanzen und Annotationen stützt.

In der *RNA-Pipeline* wird nach dem *Preprocessing* der Analyseschritt *RNA-Identification* durchgeführt. Dabei werden im Mittel 5.576.138 potentiell rRNA-kodierende *Reads* vorhergesagt. Dies entspricht etwa 93 % der in MG-RAST hochgeladenen Datenmenge. Vor dem Hintergrund der im Mittel 3.845.574 *Reads*, die in der *Protein-Pipeline* als proteinkodierend vorhergesagt wurden, lässt sich schließen, dass ein großer Anteil der Sequenzdaten sowohl in der *Protein-* als auch in der *RNA-Pipeline* analysiert wird. Beim *Clustering* werden durchschnittlich 2.171.496 Cluster generiert. Analog zu dem für die *Protein-Pipeline* beschriebenen Phänomen fand auch hier kein *Clustering* der Datensätze der Proben 7n, 13n, 37, 48v, 48n und 42n statt. Gemittelt über die Datensätze der Proben 6n, 18v, 18n, 42v, 19v, 19n, 43v und 43n, bei denen das *Clustering* korrekt durchgeführt wurde, lassen sich 306.746 Cluster errechnen, was einer Datenreduktion von einem Faktor von etwa 18 für den *Clustering*-Schritt entspricht. Bei einer Abfrage in der *M5RNA*-Datenbank wurden im Mittel insgesamt 1.981.337 Cluster mit Homologie zu rRNA-Sequenzen identifiziert. Im Falle der korrekt geclusterten Datensätze konnten 245.005 Cluster zugeordnet werden. Bezogen auf die ursprünglich in MG-RAST hochgeladene Anzahl von 5.979.803 *Reads* kann insgesamt eine Datenreduktion um den Faktor 15 für die korrekt geclusterten Datensätze beobachtet werden. Gemittelt über alle Datensätze erfolgte eine Datenreduktion um einen Faktor 2. Der für einige Proben nicht durchgeführte *Clustering*-Schritt wirkt sich ebenfalls nicht auf die Analyse und Vergleichbarkeit der Datensätze bei der taxonomischen und funktionalen Auswertung aus, da diese auf Grundlage der den entsprechenden Annotationen zugeordneten *Read*-Abundanzen durchgeführt werden.

**Tabelle 15: Übersicht der Sequenz-Datensätze der Achselhöhlen-Mikrobiota nach Bearbeitung in MG-RAST. Dargestellt ist für jede Probe die Anzahl der prozessierten Reads, die MG-RAST Metagenom-ID, die mittlere Read-Länge sowie der GC-Gehalt.**

| Proband           | Probe | MG-RAST<br>prozessiert<br>[# Reads] | MG-RAST<br>Metagenom-ID | Read-Länge (MW ±<br>SD) [Basen] | % GC<br>[MW ± SD] |
|-------------------|-------|-------------------------------------|-------------------------|---------------------------------|-------------------|
| 6                 | 6n    | 6.914.185                           | 4562377.3               | 201 ± 65                        | 49 ± 6            |
| 7                 | 7n    | 5.560.185                           | 4552114.3               | 198 ± 64                        | 53 ± 4            |
| 13                | 13n   | 4.529.943                           | 4552109.3               | 193 ± 59                        | 48 ± 5            |
|                   | 37n   | 4.870.816                           | 4552110.3               | 202 ± 63                        | 48 ± 5            |
| 24                | 48v   | 4.703.679                           | 4552113.3               | 208 ± 67                        | 50 ± 5            |
|                   | 48n   | 4.955.845                           | 4552112.3               | 200 ± 64                        | 47 ± 6            |
| 18                | 18v   | 6.977.013                           | 4562371.3               | 199 ± 63                        | 48 ± 4            |
|                   | 18n   | 6.060.408                           | 4562370.3               | 200 ± 62                        | 48 ± 5            |
|                   | 42v   | 7.564.246                           | 4562374.3               | 206 ± 65                        | 48 ± 4            |
|                   | 42n   | 5.467.865                           | 4552111.3               | 199 ± 65                        | 48 ± 5            |
| 19                | 19v   | 6.652.149                           | 4562373.3               | 199 ± 62                        | 53 ± 4            |
|                   | 19n   | 6.494.801                           | 4562372.3               | 196 ± 60                        | 52 ± 4            |
|                   | 43v   | 5.749.430                           | 4562376.3               | 196 ± 59                        | 51 ± 5            |
|                   | 43n   | 7.125.375                           | 4562375.3               | 194 ± 63                        | 48 ± 6            |
| <b>Mittelwert</b> | -     | <b>5.973.281</b>                    | -                       | <b>199 ± 63</b>                 | <b>49 ± 5</b>     |

Nach der Sequenzdatenprozessierung in MG-RAST stehen für jeden Datensatz durchschnittlich 5.973.281 Reads für die taxonomische und funktionale Analyse zur Verfügung (Tabelle 15). Dies entspricht einem Anteil von etwa 95 % der Rohdaten nach Sequenzierung. Die mittlere Länge der Sequenzen beträgt 199 Basen ± 63 Basen und der mittlere GC-Gehalt beträgt 49 % ± 5 %. Bei der Betrachtung des mittleren GC-Gehaltes der Datensätze der Probanden 18 und 19 fällt auf, dass der mittlere GC-Gehalt für die Datensätze des Probanden 18 konstant bei 48 % knapp unterhalb des Mittelwertes liegt. Dahingegen liegt der mittlere GC-Gehalt für die Datensätze des Probanden 19 im Falle der Proben 19v, 19n und 43v über dem Mittelwert. Da Corynebakterien im Allgemeinen einen relativ hohen GC-Gehalt in ihrer genomischen DNA

besitzen, lässt dieses Phänomen bereits auf Unterschiede in der taxonomischen Zusammensetzung der Achselhöhlen-Mikrobiota dieser beiden Probanden schließen.

## 2.3 Taxonomische Analyse des Metatranskriptoms der Achselhöhlen-Mikrobiota

Mithilfe der *Analyse-Pipeline* MG-RAST kann die taxonomische Zusammensetzung der Achselhöhlen-Mikrobiota beschrieben und die taxonomischen Profile in unterschiedlichen Proben verglichen werden. Grundsätzlich bietet MG-RAST die Möglichkeit, das taxonomische Profil einer Probe auf Ebene der rRNA und auf funktionaler Ebene auf diversen taxonomischen Ebenen darzustellen. Im Folgenden werden die unter Verwendung von MG-RAST und anderen *Tools* durchgeführten taxonomischen Analysen der Achselhöhlen-Mikrobiota beschrieben.

### 2.3.1 Taxonomische Beschreibung der Achselhöhlen-Mikrobiota auf Genus-Ebene

Die Beschreibung der Zusammensetzung der aktiven Lebensgemeinschaften der Achselhöhle wurde in MG-RAST zum einen rRNA-basiert auf Grundlage der *M5RNA*-Datenbank und zum anderen funktional auf Grundlage der *M5NR*-Datenbank durchgeführt. In Abbildung 15 sind beide Profile für alle untersuchten Proben dargestellt. Im Einklang mit den Beschreibungen aus der Literatur handelt es sich bei den dominanten Gattungen in beiden Profilen um Staphylokokken und Corynebakterien. Die rRNA-basierten und funktionalen taxonomischen Profile einzelner Proben ähneln sich sehr stark, wobei das funktionale Profil jedoch stets diverser erscheint.

In der rRNA-Taxonomie stellen die SP-beta-ähnlichen Viren (hellblau) einen großen Anteil der von Staphylokokken dominierten Lebensgemeinschaften dar (Abbildung 15 a). Diese Annotation steht nicht im Einklang mit Beschreibungen aus der Literatur und erscheint vor dem Hintergrund einer rRNA-basierten Annotation nicht plausibel, da Viren nicht über rRNA verfügen. Die ursprüngliche Quelle dieser Annotation in der *M5RNA*-Datenbank stellt die *Greengenes*-Datenbank dar. Eine händisch durchgeführte Abfrage von wenigen zufällig ausgewählten Sequenzen mit dieser Annotation in der *GenBank*-Datenbank lieferte Treffer gegen rRNA-Sequenzen aus unterschiedlichen *Staphylococcus*-Spezies. Somit kann angenommen werden, dass es sich bei diesen Sequenzen um konservierte Bereiche aus der rRNA von Staphylokokken handelt und dass es sich bei der Annotation SP-beta-ähnliche Viren um eine Fehlannotation auf Grundlage eines fehlerhaften Datenbank-Eintrages handelt. Ein Vergleich der relativen Abundanzen dieser Annotation mit den relativen Abundanzen der Annotationen zu *Staphylococcus* in unterschiedlichen Profilen zeigt einen Zusammenhang und legt somit nahe,

dass die zugrundeliegenden Sequenzen aus Staphylokokken stammen. Weiterhin werden auf rRNA-Ebene die Gram-positiven anaeroben Bakterien der Genera *Anaerococcus*, *Propionibacterium*, *Fingoldia* und *Peptoniphilus* beschrieben. Andere und nicht klassifizierte Bakteriengattungen beschreiben lediglich 1 % bis 7 % der *Reads*. Die Achselhöhlen-Mikrobiota ist in dieser Darstellung auf der Genus-Ebene somit durch die genannten sechs Bakteriengattungen nahezu vollständig beschrieben. Eine Besonderheit stellt die Annotation zu dem Genus *Coptotermes* im Datensatz der Probe 7n dar. Diese Organismen sind nicht aus der Literatur als Bewohner der Achselhöhle bekannt. Eine händische Abfrage einiger zufällig ausgewählter Sequenzen in der GenBank-Datenbank unter Verwendung des BLASTN-Algorithmus lieferte Sequenzhomologien zu Corynebakterien. Daher kann auch hier angenommen werden, dass es sich bei der Zuordnung um eine fehlerhafte Annotation in einer der zugrunde liegenden Datenbanken handelt.

In der funktionalen Taxonomie können neben den Corynebakterien und Staphylokokken analog zur rRNA-basierten Taxonomie Annotationen der Genera *Propionibacterium* sowie der Gram-positiven anaeroben Kokken *Anaerococcus*, *Fingoldia* und *Peptoniphilus* beschrieben werden (Abbildung 15 b). Weiterhin ist der Genus *Bacillus* (in hellblau dargestellt) ebenfalls in Profilen mit einem hohen Anteil an Staphylokokken dominant. Die Gattung *Bacillus* wurde jedoch bisher nicht in der Literatur als Mitglied der Haut-Mikrobiota beschrieben. Da die Genera *Bacillus* und *Staphylococcus* zu der Ordnung *Bacillales* gehören und daher nahe verwandt sind, stellen die Annotationen des Genus *Bacillus* wahrscheinlich eine falsche Zuordnung von Staphylokokken-Sequenzen dar. Weitere Bakteriengattungen, die hier zugeordnet werden, jedoch nicht in der Literatur im Zusammenhang mit der Haut-Mikrobiota beschrieben sind, sind die ebenfalls zum Bakterienstamm der *Firmicutes* gehörenden Genera: *Lactobacillus*, *Listeria*, *Clostridium*, *Pseudoramibacter*, *Heliobacterium* und *Candidatus Phytoplasma*. Weiterhin werden die Gattungen *Bifidobacterium*, *Actinomyces*, *Streptomyces* und *Kineococcus* annotiert, die Mitglieder des Bakterienstammes *Actinobacteria* darstellen. Da diese Bakteriengattungen wie die Gattungen *Staphylococcus* und *Corynebacterium* zu den Bakterienstämmen *Firmicutes* bzw. *Actinobacteria* gehören, handelt es sich bei der taxonomischen Annotation dieser funktionalen Sequenzen vermutlich um Fehlannotationen aufgrund relativ hoher Sequenzhomologie proteinkodierender Sequenzen innerhalb der Gattungen eines Stammes. Der Genus *Halomonas*, der zu dem Stamm der *Proteobacteria* gehört, wird im Datensatz der Probe 42n mit einem Anteil von 2 % der *Reads* annotiert. Bei dieser Gattung handelt es sich um halotolerante und halophile Organismen, die zunehmend als menschliche Krankheitserreger beschrieben werden (Kim et al., 2013; Stevens et al., 2009, 2013). Proteobakterien wurden ebenfalls in der Literatur im Zusammenhang mit der Achselhöhlen-Mikrobiota beschrieben (Grice et al., 2009).

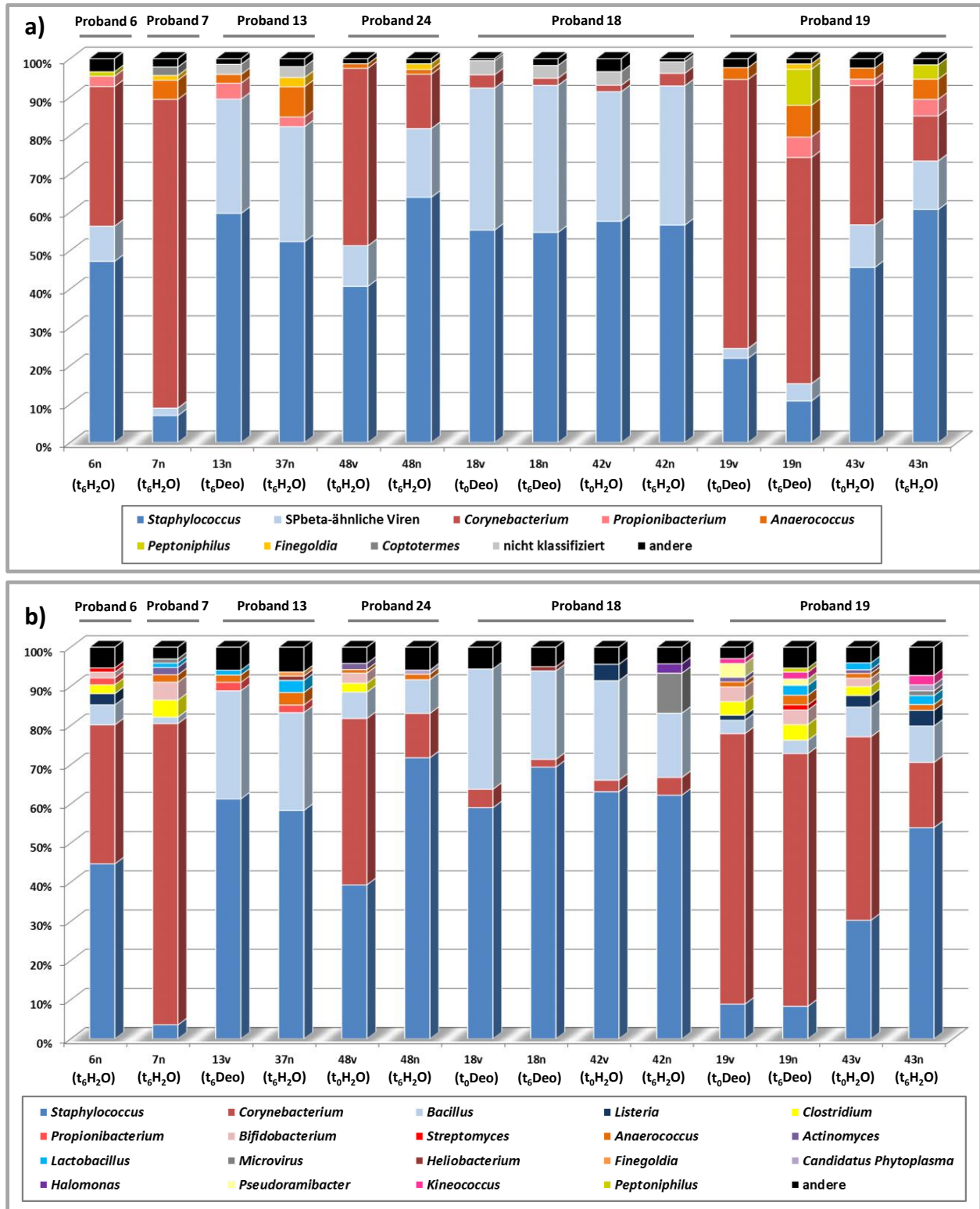


Abbildung 15: Taxonomische Darstellung der aktiven Achselhöhlen-Mikrobiota auf Genus-Ebene. a) rRNA-basierte taxonomische Profile auf Grundlage der *MSRNA*-Datenbank; b) funktionale taxonomische Profile auf Grundlage der *MSNR*-Datenbank.

Der Großteil der betrachteten taxonomischen Profile wird entweder von Sequenzen mit Homologie zu Staphylokokken-Sequenzen oder von solchen mit Homologie zu Corynebakterien-Sequenzen dominiert. Annotationen zum Genus *Staphylococcus* dominieren die taxonomischen Profile der acht Proben 13n, 37n, 48n, 18v, 18n, 42v, 42n und 43n. Die taxonomischen Profile der Datensätze der drei Proben 7n, 19v und 19n werden von Corynebakterien dominiert. Als dritter Typ lassen sich zusätzlich taxonomische Profile identifizieren, in denen sowohl Corynebakterien als auch Staphylokokken dominant sind. Die taxonomischen Zuordnungen der Reads in den drei Proben 6n, 48v und 43v sind relativ ausgeglichen. Somit ist die Achselhöhlen-Mikrobiota der Mehrheit der Proben und der Mehrheit der untersuchten Probanden von Staphylokokken dominiert. Besondere Beachtung verdient die Tatsache, dass die Probe 7n, die mit einem Wert von 3,5 den höchsten *Malodor Score* der gesamten Deo-Studie aufweist, ein eindeutig corynebakteriell dominiertes taxonomisches Profil aufweist. Weiterhin besitzen die Proben 6n und 19v mit ähnlich hohen Anteilen an Corynebakterien- und Staphylokokken-Annotationen überdurchschnittlich hohe *Malodor Scores*. Im Falle der Probe 48v liegt der *Malodor Score* jedoch mit einem Wert von 0,75 unter dem Durchschnitt. Weiterhin können z. T. überdurchschnittlich hohe *Malodor Scores* ebenfalls für Proben beschrieben werden, deren taxonomische Profile eindeutig durch Annotationen zu Staphylokokken dominiert sind.

Insgesamt lässt sich beobachten, dass taxonomische Profile aus Proben der selben Probanden eine im Vergleich zu Proben anderer Probanden höhere Ähnlichkeit aufweisen. Entsprechend ist eine sehr große Ähnlichkeit der Staphylokokken-dominierten taxonomischen Profile der Datensätze der Probanden 18 und 13 erkennbar. Taxonomische Profile von Datensätzen mit einem hohen Anteil an corynebakteriellen Sequenzen weisen jedoch untereinander eine deutlich geringere Ähnlichkeit auf. So nimmt bei einem Vergleich der Profile der Proben 48v und 48n des Probanden 24 der Anteil an Staphylokokken nach Behandlung mit Wasser zu. Gleichzeitig wird hingegen der Anteil an corynebakteriellen Sequenzen deutlich verringert. Im Falle der rRNA-basierten taxonomischen Profile der Proben des Probanden 19 lassen sich ebenfalls sehr deutliche Veränderungen beschreiben. Am ersten Probennahme-Tag vor Deo-Anwendung wird das taxonomische Profil von Corynebakterien (70 %) dominiert und der Anteil der Staphylokokken liegt bei 22 %. Nach der Deo-Behandlung wird sowohl der Anteil der Corynebakterien- als auch der Anteil der Staphylokokken-Annotationen reduziert (59 % *Corynebacterium* und 11 % *Staphylococcus*). Weiterhin wird der Anteil der Annotationen zu den Gram-positiven anaeroben Kokken *Peptoniphilus* (9 %), *Anaerococcus* (8 %) und *Fingoldia* (2 %) sowie zum Genus *Propionibacterium* (5 %) heraufgesetzt. Am zweiten Beprobungstag wird das taxonomische Profil mit einem Anteil von 36 % nicht mehr nur durch corynebakterielle Annotationen sondern auch durch Annotationen zum Genus *Staphylococcus* (46 %) dominiert.



Dies zeigt, dass sich die Zusammensetzung der Achselhöhlen-Mikrobiota eines Menschen im zeitlichen Verlauf ändern kann, was ebenfalls in der Studie von Costello et al. beschrieben ist. Nach Wasser-Anwendung kommt es analog zu dem für den Probanden 24 beschriebenen Phänomen zu einer Erhöhung des Anteils an Sequenzen mit Homologie zu Sequenzen aus Staphylokokken (61 %) und gleichzeitig zu einer Verringerung der Corynebakterien (12 %). Gleichzeitig steigt der Anteil an Sequenzen mit Homologie zu Sequenzen aus den Genera *Fingoldia*, *Propionibacterium* und *Peptoniphilus*. Dies zeigt, dass sich die Zusammensetzung der Achselhöhlen-Mikrobiota intraindividuell in Abhängigkeit von der Zeit und externen Faktoren verändern kann.

### **2.3.2 Hauptkomponentenanalyse der rRNA-basierten taxonomischen Zusammensetzung der Achselhöhlen-Mikrobiota auf Genus-Ebene**

Zur Beschreibung der Ähnlichkeit der taxonomischen Zusammensetzung der Achselhöhlen-Mikrobiota auf Genus-Ebene wurde eine Hauptkomponentenanalyse durchgeführt. Als Grundlage dienten dabei die rRNA-basierten Annotationen der Datensätze. Für die Analyse wurden die 12 Genera mit den höchsten Abundanzen über alle Datensätze herangezogen: *Staphylococcus* (1), *Corynebacterium* (2), *Propionibacterium* (3), *Anaerococcus* (4), *Peptoniphilus* (5), nicht klassifizierte Genera (6), *Coptotermes* (7), *Fingoldia* (8), *Rhodococcus* (9), *Enterococcus* (10), *Brachybacterium* (11) und *Micrococcus* (12).

Abbildung 16 a) zeigt den *Loading-Plot*, der den Zusammenhang zwischen den Hauptkomponenten 1 und 2 und den annotierten Bakteriengattungen, die als Vektoren dargestellt sind. Bei den Vektoren mit dem größten Betrag handelt es sich um die Annotationen der Genera *Staphylococcus* (1) und *Corynebacterium* (2), die in der rRNA-basierten taxonomischen Zuordnung auf Genus-Ebene den höchsten Anteil aufweisen. Keiner dieser Vektoren ist an einer der beiden Hauptkomponenten ausgerichtet, was bedeutet, dass die Varianz der Zusammensetzung der Datensätze auf der Varianz in den Abundanzen beider Genera beruht. Besonders auffällig ist, dass die beiden Vektoren nahezu orthogonal zueinander stehen, wobei der Vektor 1 (*Staphylococcus*) in den Quadranten I und der Vektor 2 (*Corynebacterium*) in den Quadranten IV zeigt. Dies steht im Einklang mit der Beobachtung, dass in der taxonomischen Zuordnung der Datensätze entweder der Genus *Staphylococcus* oder der Genus *Corynebacterium* dominiert und dass die beiden Genera negativ korrelieren (Taylor et al., 2003; Troccaz et al., 2015).

Im Gegensatz dazu sind die Vektoren der Annotationen der Gram-positiven anaeroben Kokken *Anaerococcus* (4) und *Peptoniphilus* (5) sowie dem anaeroben Actinobacterium

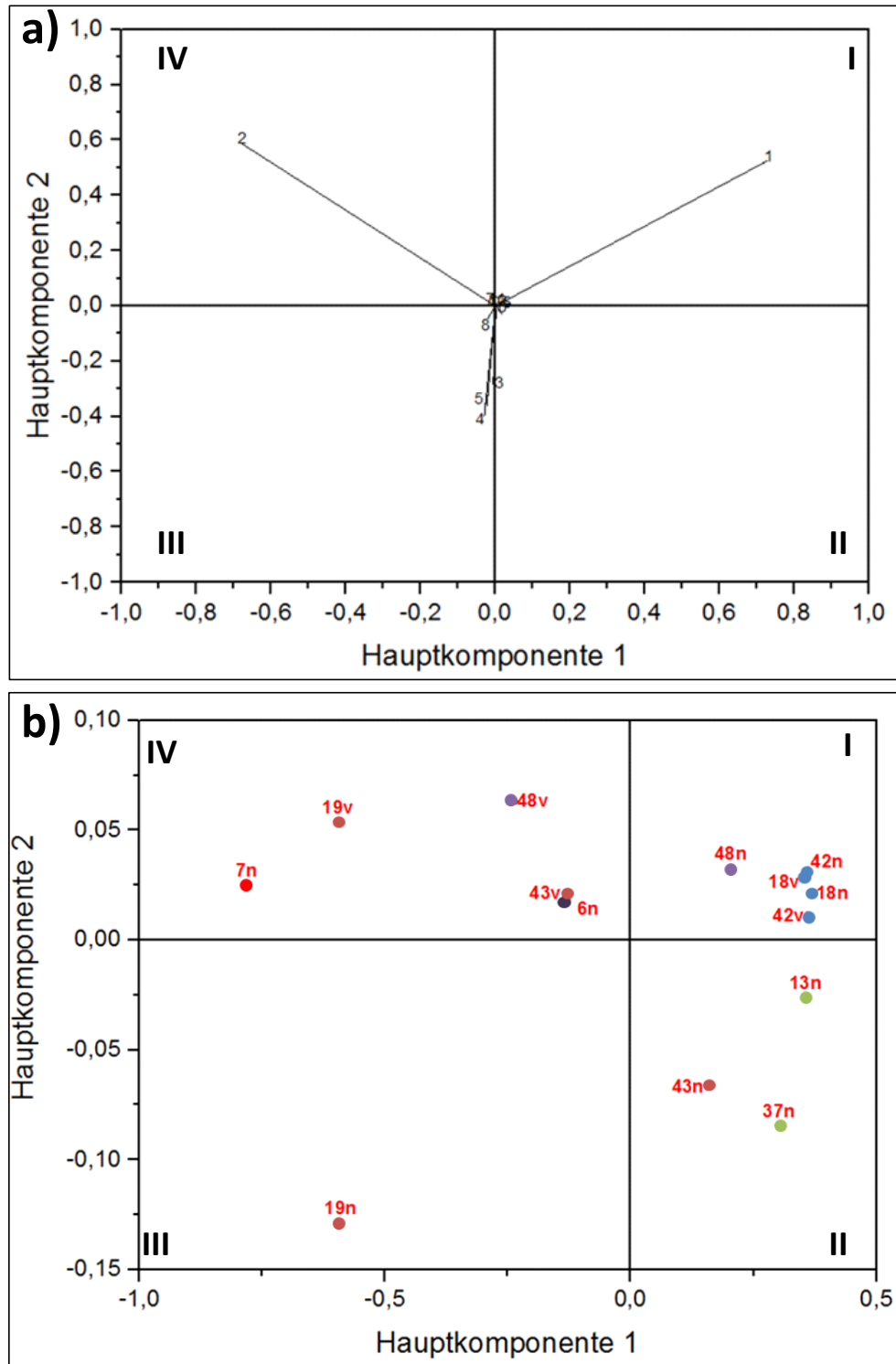


Abbildung 16: Hauptkomponentenanalyse (PCA) der rRNA-basierten taxonomischen Zusammensetzung der Achselhöhlen-Mikrobiota auf Genus-Ebene. a) *Loading-Plot* für die Genera *Staphylococcus* (1), *Corynebacterium* (2), *Propionibacterium* (3), *Anaerococcus* (4), *Peptoniphilus* (5), nicht klassifizierte Genera (6), *Coptotermes* (7), *Finegoldia* (8), *Rhodococcus* (9), *Enterococcus* (10), *Brachybacterium* (11) und *Micrococcus* (12); b) *PCA-Plot* mit den Proben der Probanden 6 (6n, schwarz), 7 (7n, hellrot), 13 (13n und 37n, grün), 24 (48v und 48n, lila), 18 (18v, 18n, 42v und 42n, blau) und 19 (19v, 19n, 43v und 43n, braun).

*Propionibacterium* an der Hauptkomponente 2 ausgerichtet. Dies zeigt, dass die zweitgrößte im Rahmen der Datentransformation erzielte Varianz durch Annotationen dieser Genera beschrieben wird. Weiterhin beschreibt der in den Quadranten III ragende Vektor 8 Annotationen zum Genus *Fingoldia*, der ebenfalls zu den Gram-positiven anaeroben Kokken (GPAC) zählt. Die übrigen 6 Genera werden durch sehr kurze Vektoren beschrieben und spielen somit bei der Varianz der Datensätze keine nennenswerte Rolle. Somit kann die gesamte inter- und intraindividuelle Variabilität der Achselhöhlen-Metatranskriptome durch Annotationen zu diesen sechs dominantesten Bakteriengattungen beschrieben werden, was im Einklang mit den in Kapitel 2.3.1 beschriebenen Beobachtungen steht.

Bei Betrachtung des PCA-Plots in Abbildung 16 b) ist erkennbar, dass der Datensatz 48n sowie die Datensätze 18v, 18n, 42v, 42n des Probanden 18 entsprechend ihres hohen Anteils an Sequenzen mit Homologie zu Sequenzen aus Staphylokokken im Quadranten I angesiedelt sind (vgl. Abbildung 15). Analog dazu befinden sich die Datenpunkte der Proben 7n, 19v, 48v, 43v und 6n, die einen hohen Anteil an Corynebakterien-Annotationen aufweisen, in dem Quadranten IV. Dabei stellt der Datenpunkt 7n aufgrund seines sehr hohen Anteils an Sequenzen mit Homologie zu Corynebakterien-Sequenzen den am weitesten vom Ursprung des Koordinatensystems entfernten Punkt im Quadranten IV dar. Die Datensätze der Proben 13n und 37n sowie 43n, die durch Staphylokokken-Annotationen dominiert werden, jedoch ebenfalls einen Anteil an Annotationen zu den GPAC und zum Genus *Propionibacterium* besitzen, sind in dem Quadranten II angesiedelt. Der Datensatz der Probe 19n ist als einziger im Quadranten III angesiedelt, da die taxonomische Zusammensetzung von Corynebakterien dominiert wird und Annotationen zu den GPAC und zum Genus *Propionibacterium* insgesamt etwa 24 % des Datensatzes umfassen (vgl. Abbildung 15 a).

Bei Betrachtung der intraindividuellen Variabilität zeigt sich ein deutliches *Clustering* der Datenpunkte der Proben 18v ( $t_0$  Deo), 18n ( $t_6$  Deo), 42v ( $t_0$  H<sub>2</sub>O) und 42n ( $t_6$  H<sub>2</sub>O) des Probanden 18 im Quadranten I (blau), was auf eine sehr hohe zeitliche intraindividuelle Ähnlichkeit der Achselhöhlen-Mikrobiota dieses Probanden hindeutet, sowie auf eine sehr hohe Stabilität der mikrobiellen Zusammensetzung unabhängig von der Behandlung mit Deodorant oder Wasser. Im Gegensatz dazu liegen die Datenpunkte der Datensätze 19v ( $t_0$  Deo), und 43v ( $t_0$  H<sub>2</sub>O), 43n ( $t_6$  H<sub>2</sub>O) und 19n ( $t_6$  Deo) des Probanden 19 sehr weit voneinander entfernt in den Quadranten IV, II und III und weisen daher eine sehr hohe intraindividuelle Variabilität bezogen auf die Zeit und die Behandlung auf. Der Datenpunkt der Probe 19n weist dabei eine besonders große Distanz zu den übrigen Datenpunkten auf, was auf eine sehr starke Veränderung in der Zusammensetzung der Achselhöhlen-Mikrobiota nach Deo-Anwendung hinweist. Diese Distanz ist zwischen den

Proben 43v und 43n weniger stark ausgeprägt, was darauf hindeutet, dass die Behandlung der Achselhöhle mit Wasser einen geringeren Einfluss auf die mikrobielle Gesellschaft dieses Probanden ausübt als die Deo-Behandlung. Auffallend ist, dass sich die Datenpunkte der Proben, die an einem Tag genommen wurden, d. h. 19v und 19n sowie 43v und 43n, hauptsächlich entlang der Hauptkomponente 2 verschieben. Da diese im Wesentlichen durch Annotationen der Genera *Anaerococcus*, *Peptoniphilus* und *Propionibacterium* beschrieben wird, ist naheliegend, dass diese Genera maßgeblich an der Veränderung der Zusammensetzung der Achselhöhlen-Mikrobiota nach Deo- bzw. Wasser-Anwendung beteiligt sind. Diese Verschiebung ist zwischen den Datenpunkten der Proben 19v ( $t_0$  Deo) und 19n ( $t_6$  Deo) besonders ausgeprägt, während die Verschiebung zwischen den Datenpunkten 43v ( $t_0$  H<sub>2</sub>O) und 43n ( $t_6$  H<sub>2</sub>O) entlang der Hauptkomponente 2 weniger stark ausfällt und gleichzeitig ebenfalls eine geringfügige Verschiebung entlang der Hauptkomponente 1 erfolgt. Im Gegensatz dazu lässt sich bei Vergleich der Unterschiede zwischen den Proben der einzelnen Beprobungstage (19v und 19n vs. 43v und 43n) eine starke Verschiebung entlang der Hauptkomponente 1 und eine schwächere Verschiebung entlang der Hauptkomponente 2 beobachten.

Zwischen den Datenpunkten der Proben 48v ( $t_0$  H<sub>2</sub>O) und 48n ( $t_6$  H<sub>2</sub>O) des Probanden 24 findet die größere Verschiebung entlang der Hauptkomponente 1 statt. Der Datenpunkt 48n clustert im Quadranten I mit den Datenpunkten des Probanden 18. Dies lässt auf eine höhere interindividuelle Ähnlichkeit dieser Proben im Vergleich zur intraindividuellen Ähnlichkeit der Proben des Probanden 24 schließen. Weiterhin verschieben sich die Datenpunkte 13n ( $t_6$  Deo) und 37n ( $t_6$  H<sub>2</sub>O) hauptsächlich entlang der Hauptkomponente 2. Der Datenpunkt der Probe 6n ( $t_6$  H<sub>2</sub>O) des Probanden 6 ist nahezu deckungsgleich mit dem Datenpunkt 43v des Probanden 19, was auf eine sehr hohe interindividuelle Ähnlichkeit dieser beiden Proben im Hinblick auf die Zusammensetzung der Achselhöhlen-Mikrobiota schließen lässt.

Zusammenfassend ist festzustellen, dass die durch Staphylokokken dominierten taxonomischen Profile weniger stark variieren als solche mit einem relativ hohen Anteil an Corynebakterien. Eine Staphylokokken-dominierte Mikrobiota könnte demnach als stabiler beschrieben werden als eine Corynebakterien-dominierte Mikrobiota im Hinblick auf Zeit und die Behandlung mit Deodorant oder Wasser.

### **2.3.3 Taxonomische Beschreibung der Achselhöhlen-Mikrobiota auf Spezies-Ebene**

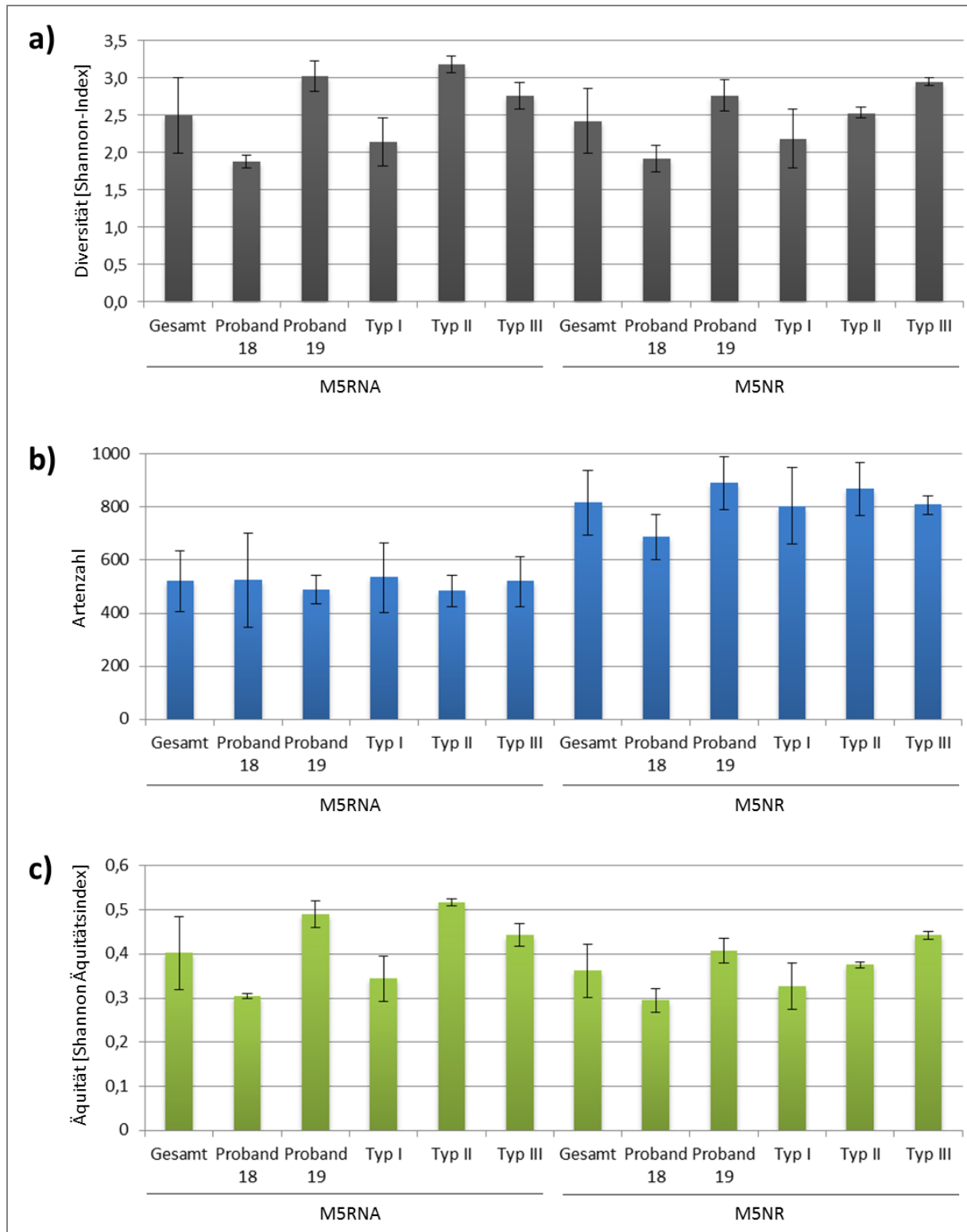
Auf Spezies-Ebene ist die Beschreibung der aktiven Lebensgemeinschaft der Achselhöhle ebenfalls auf Grundlage der *M5RNA*- und *M5NR*-Datenbank möglich. In Abbildung 17 sind die Diversitäten, Artenzahlen und Äquitäten für alle untersuchten Proben sowie für diverse

Subdatensätze dargestellt. Auf rRNA-Ebene liegt die mittlere Diversität aller untersuchten Datensätze bei einem Shannon-Index von  $2,5 \pm 0,51$ . Auffallend ist, dass die Lebensgemeinschaften des Probanden 18 und des Typ I, welche von Staphylokokken dominiert sind, unterdurchschnittliche Diversitäten aufweisen (1,9 bzw. 2,1). Diese sind darüber hinaus deutlich niedriger, als die Diversitäten der Corynebakterien-dominierten Lebensgemeinschaften des Probanden 19 (3,0) und des Typ II (3,2). Die Diversität der sowohl von Corynebakterien als auch von Staphylokokken dominierten Lebensgemeinschaften des Typen III ist mit einem Shannon-Index von 2,8 intermediär. Die Diversitäten auf funktionaler Ebene, d. h. auf Grundlage der *M5NR*-Datenbank weichen lediglich geringfügig von denen auf Grundlage der *M5RNA*-Datenbank basierten ab.

Bei Betrachtung der Artenzahlen fällt auf, dass diese auf funktionaler Ebene deutlich höher sind als auf rRNA-Ebene (s. Abbildung 17 b). Im Mittel werden auf Grundlage der *M5RNA*-Datenbank  $528 \pm 111$  Spezies und auf Grundlage der *M5NR*-Datenbank  $818 \pm 122$  Spezies beschrieben. Die Unterschiede zwischen den einzelnen Subdatensätzen sind sehr gering und liegen noch innerhalb der Streuung.

Die Äquität ist ein Maß für die Gleichmäßigkeit der Verteilung von Spezies innerhalb einer Lebensgemeinschaft. Der mittlere Äquitätsindex aller Proben auf rRNA-Grundlage beträgt  $0,4 \pm 0,08$  (s. Abbildung 17 c). Analog zur Diversität, besitzen Staphylokokken-dominierte Lebensgemeinschaften geringere Äquitäten als Corynebakterien-dominierte Lebensgemeinschaften, während gemischte Lebensgemeinschaften eine intermediäre Äquität besitzen. Die Äquitätsindizes auf der Ebene der funktionalen Taxonomie unterscheiden sich lediglich geringfügig von denen auf rRNA-Ebene.

Staphylokokken-dominierte Lebensgemeinschaften der Achselhöhle sind somit weniger divers als Corynebakterien-dominierte und besitzen darüber hinaus eine weniger ausgeglichene Verteilung der Spezies. Die Artenzahlen folgen hingegen keinem erkennbaren Trend. Das Phänomen, dass auf funktionaler Ebene mehr Spezies beschrieben werden, steht im Einklang mit der Beobachtung auf Genus-Ebene und stellt wahrscheinlich ein Artefakt der Annotationen dar (vgl. Kapitel 2.3.1). Daher werden im Folgenden lediglich die rRNA-basierten taxonomischen Profile näher betrachtet.



**Abbildung 17: Diversität (a), Artenzahl (b) und Äquität (c) der aktiven Lebensgemeinschaften der Achselhöhle. Dargestellt ist jeweils der Mittelwert +/- Standardabweichung für alle Datensätze (Gesamt), für die Probanden 18 und 19, für Staphylokokken-dominierte Lebensgemeinschaften (Typ I), Corynebakterien-dominierte Lebensgemeinschaften (Typ II) und gemischte Lebensgemeinschaften (Typ III) auf Grundlage der *M5RNA*- und der *M5NR*-Datenbank.**

In den Abbildungen 18 – 20 sind die Zusammensetzungen der Lebensgemeinschaften rRNA-basiert auf Spezies-Ebene dargestellt. Die Profile stehen im Einklang mit den Beschreibungen auf Genus-Ebene. Bei den dominanten Staphylokokken handelt es sich um die vier Spezies *S. epidermidis*, *S. aureus*, *S. hominis* und *S. haemolyticus*. Weiterhin sind in den Staphylokokken-dominierten Lebensgemeinschaften Zuordnungen zu *S. lugdunensis*, *S. capitis* und *S. cohnii* zu beobachten. Bei allen beschriebenen Staphylokokken handelt es sich um bekannte Vertreter der Haut- und Achselhöhlen-Mikrobiota. Der Zuordnung *Bacillus Phage SPbeta* liegt wahrscheinlich der fehlerhafte Eintrag in der *Greengenes*-Datenbank zugrunde. Dies wurde ebenfalls auf Genus-Ebene beschrieben (s. 2.3.1). Die Spezies-Zugehörigkeiten der Corynebakterien sind vor allem in Corynebakterien-dominierten Lebensgemeinschaften deutlich diverser als die der Staphylokokken. Bei den dominanten Corynebakterien-Spezies handelt es sich um *C. tuberculostearicum*, *C. amycolatum*, *C. striatum*, *C. lipophiloflavum* und *C. aurimucosum*. Insbesondere *C. striatum* und *C. tuberculostearicum* wurden im Zusammenhang mit Schweißgeruchsentstehung beschrieben (s. Einleitung). Als weitere weniger dominante Corynebakterien können die Spezies *C. genitalium*, *C. mucifaciens*, *C. imitans*, *C. appendicis*, *C. tuscaniense*, *C. pseudogenitalium*, *C. glaucum*, *C. coyleae*, *C. afermentans*, *C. ammoniagenes*, *Corynebacterium sp.* sowie *Corynebacterium sp.*NML00-0156, *Corynebacterium sp.*NML09-0341 und *Corynebacterium sp.*NML93-0612 beschrieben werden. Die Annotation zu *Coptotermes formosanus* in der Probe 7n stellt wahrscheinlich ebenfalls eine Fehlzuordnung dar und ist konsistent mit der entsprechenden Zuordnung auf Genus-Ebene. In einigen Datensätzen werden die Propionibakterien-Spezies *P. acnes*, *P. granulosum* und *Propionibacterium sp.* 434-HC2 beschrieben, wobei insbesondere der erste Organismus als klassisches Mitglied der Haut-Mikrobiota aus der Literatur bekannt ist. Als Mitglieder der sogenannten Gram-positiven anaeroben Kokken (GPAC) werden die *Anaerococcus*-Spezies *A. tetradius*, *A. hydrogenalis*, *A. prevotii*, *A. lactolyticus* sowie *Finegoldia magna*, *Peptoniphilus lacrimalis* und *Peptoniphilus sp. oral taxon t 386* gelistet. Die Anteile unbekannter Bakterienspezies befinden sich im einstelligen Prozentbereich und sind damit erstaunlich gering. Die Anteile anderer Spezies variieren zwischen 5 % (Proband 18, s. Abbildung 19 c und d) und 24 % (Proband 19, s. Abbildung 19 b) und scheinen mit den Anteilen der Corynebakterien und der Diversität der Lebensgemeinschaften im Zusammenhang zu stehen.

Auch auf Spezies-Ebene können inter- und intraindividuelle Ähnlichkeiten beschrieben werden, welche mit den Beobachtungen auf Genus-Ebene im Einklang stehen. Die Lebensgemeinschaften der Probanden 13 und 18 sind beispielsweise von Staphylokokken (*Bacillus Phage SPbeta*), *S. epidermidis* und *S. aureus* dominiert und verändern sich in ihrer

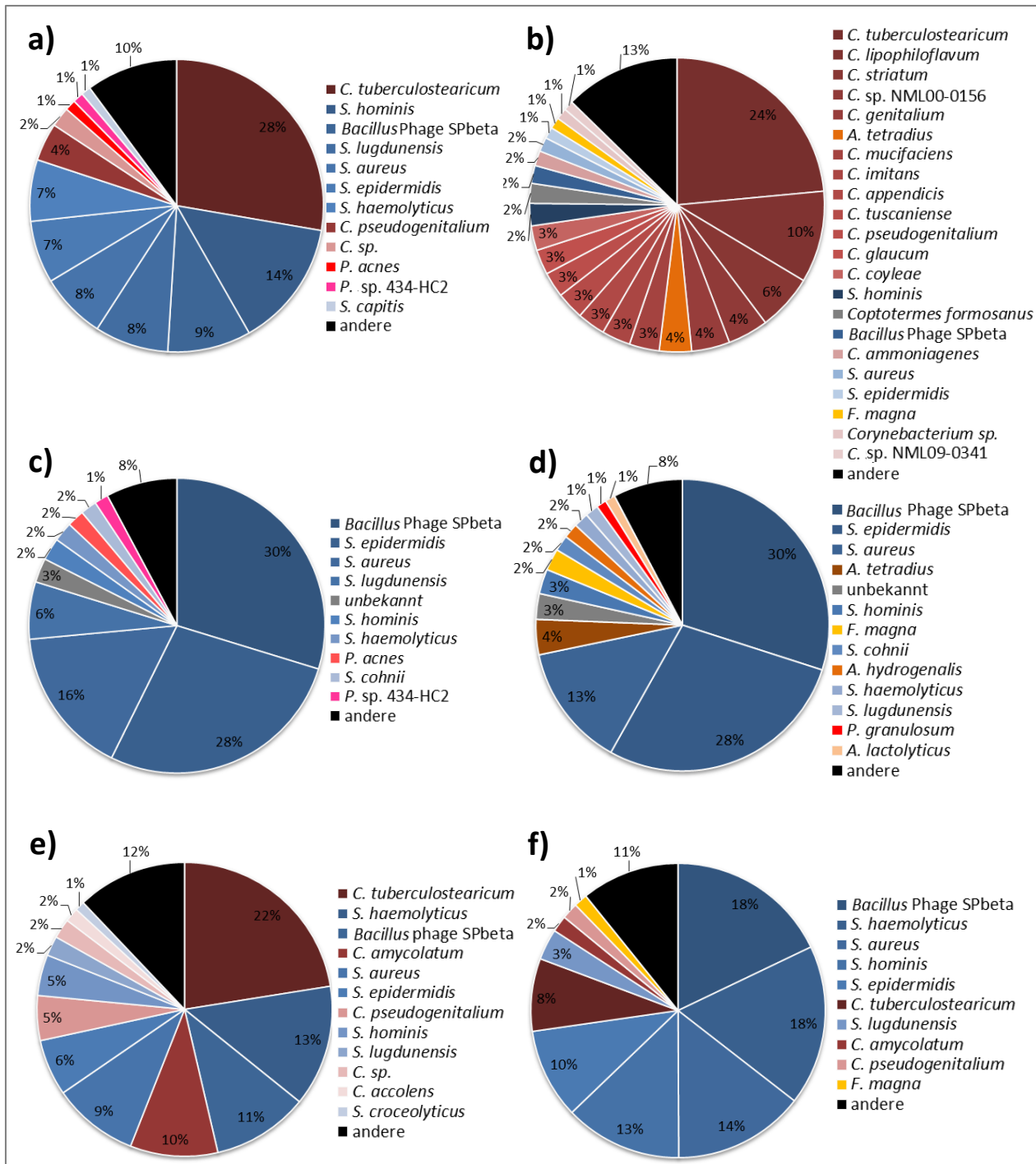


Abbildung 18: Zusammensetzung der Achselhöhlen-Mikrobiota auf Spezies-Ebene basierend auf der *M5RNA*-Datenbank. a) Probe 6n (Proband 6, t<sub>6</sub> H<sub>2</sub>O); b) Probe 7n (Proband 7, t<sub>6</sub> H<sub>2</sub>O); c) Probe 13n (Proband 13, t<sub>6</sub> Deo); d) Probe 37n (Proband 13, t<sub>6</sub> H<sub>2</sub>O); e) Probe 48v (Proband 24, t<sub>0</sub> H<sub>2</sub>O); f) Probe 48n (Proband 24, t<sub>6</sub> H<sub>2</sub>O). Die Gattungen werden wie folgt abgekürzt: S.: *Staphylococcus*, C.: *Corynebacterium*, P.: *Propionibacterium*; F.: *Fingoldia*, A.: *Anaerococcus*. Die Spezies-Listen sind absteigend nach den relativen Abundanz in den jeweiligen Lebensgemeinschaften angeordnet.



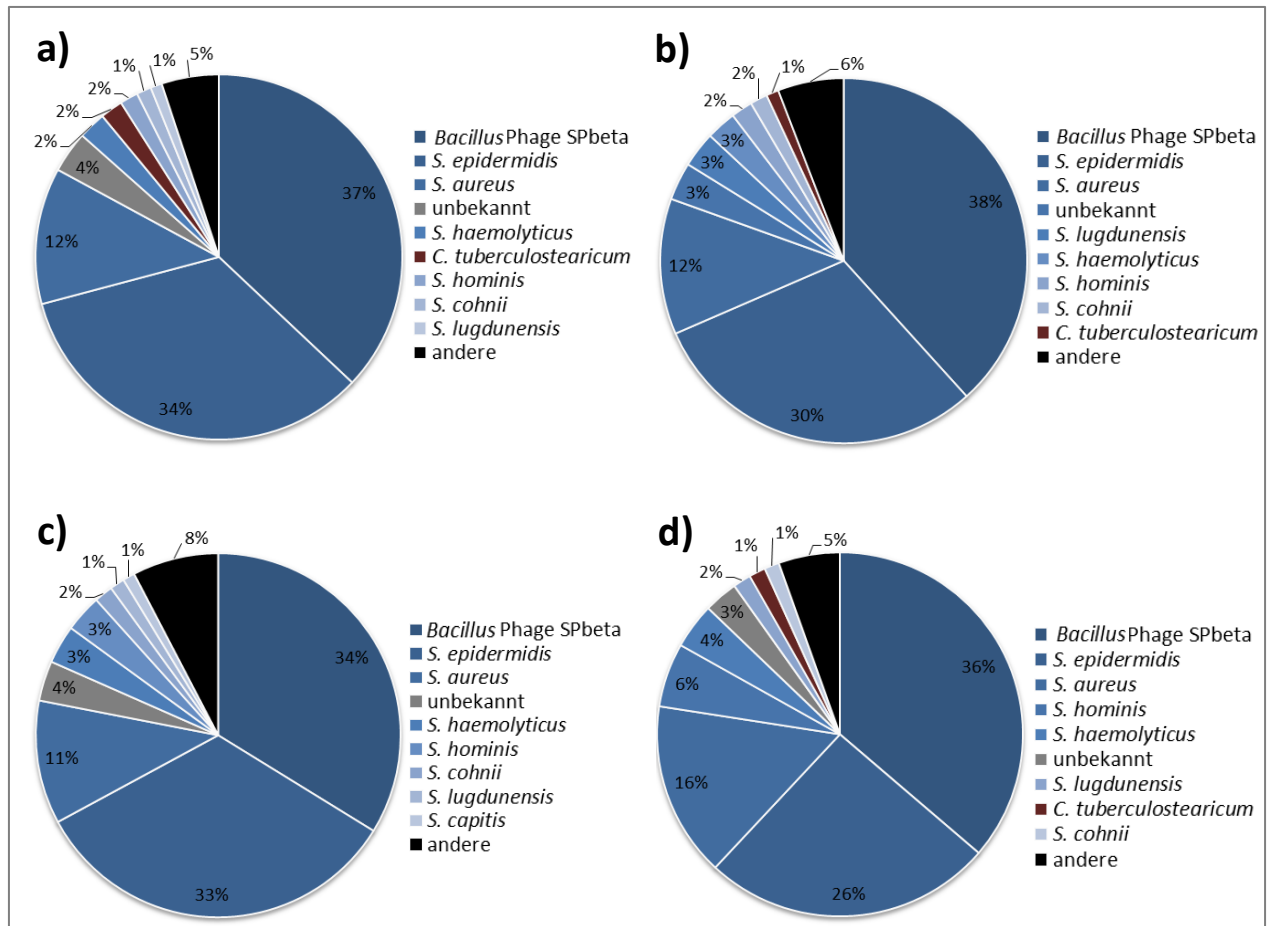
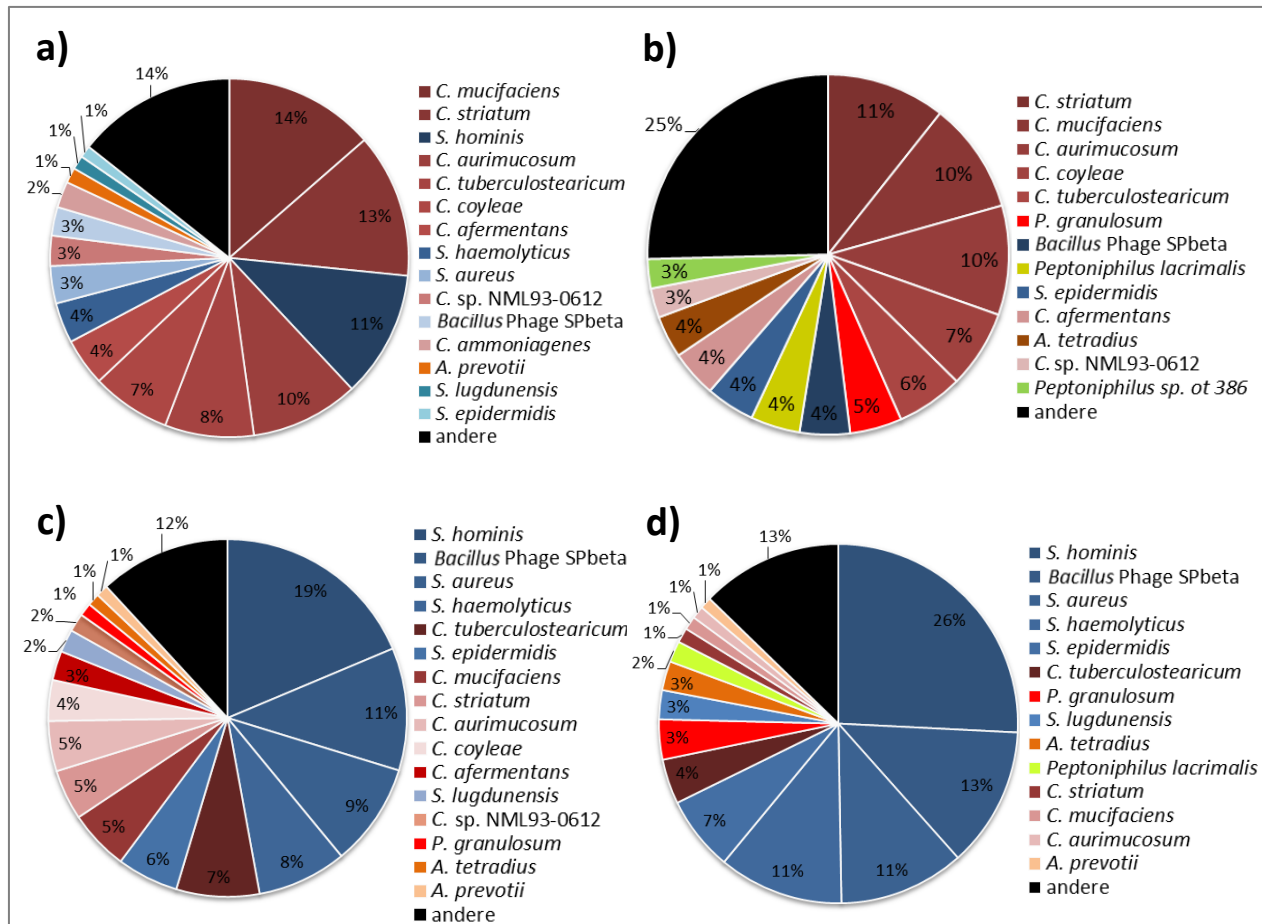


Abbildung 19: Zusammensetzung der Achselhöhlen-Mikrobiota des Probanden 18 auf Spezies-Ebene basierend auf der MSRNA-Datenbank. a) Probe 18v ( $t_0$  Deo); b) Probe 18n ( $t_6$  Deo); c) Probe 42v ( $t_0$  H<sub>2</sub>O); d) Probe 42n ( $t_6$  H<sub>2</sub>O). Die Gattungen werden wie folgt abgekürzt: S.: *Staphylococcus*, C.: *Corynebacterium*. Die Spezies-Listen sind absteigend nach den relativen Abundanz in den jeweiligen Lebensgemeinschaften angeordnet.

Zusammensetzung kaum in Abhängigkeit von der Behandlung. Dahingegen verschiebt sich die von *C. tuberculostearicum* und *S. haemolyticus* dominierte Mikrobiota des Probanden 24 hin zu einer Staphylokokken-dominierten. Besonders variabel ist die Mikrobiota des Probanden 19, welche am ersten Probennahmetag von Corynebakterien, insbesondere *C. striatum*, *C. aurimucosum* und *C. mucifaciens*, dominiert ist. Nach Deo-Anwendung verschiebt sich die Lebensgemeinschaft durch die starke Reduktion von *S. hominis* sogar weiter hin zu den Corynebakterien. Am zweiten Probennahmetag dominieren hingegen *S. hominis*, *S. aureus*, *S. haemolyticus* und andere Staphylokokken. Die Mikrobiota verschiebt sich nach Wasser-Anwendung weiter hin zu Staphylokokken. Auffallend bei den Probanden 13, 24 und 19 ist, dass die  $t_6$ -Proben höhere Anteile an GPAC enthalten. Dies ist bei der Lebensgemeinschaft des Probanden 18 nicht der Fall.



**Abbildung 20: Zusammensetzung der Achselhöhlen-Mikrobiota des Probanden 19 auf Spezies-Ebene basierend auf der MSRNA-Datenbank. a) Probe 19v (t<sub>0</sub> Deo); b) Probe 19n (t<sub>6</sub> Deo); c) Probe 43v (t<sub>0</sub> H<sub>2</sub>O); d) Probe 43n (t<sub>6</sub> H<sub>2</sub>O). Die Gattungen werden wie folgt abgekürzt: S.: *Staphylococcus*, C.: *Corynebacterium*, P.: *Propionibacterium*; F.: *Fingoldia*, A.: *Anaerococcus*. Die Spezies-Listen sind absteigend nach den relativen Abundanzen in den jeweiligen Lebensgemeinschaften angeordnet.**

Die Korrelation zwischen einzelnen Bakterienspezies und dem Körpergeruch ist in Tabelle 16 dargestellt. Es zeigt sich, dass corynebakterielle Spezies überwiegend positiv und Staphylokokken überwiegend negativ mit dem Körpergeruch korrelieren. *C. lipophiloflavum* und *C. tuscaniense* weisen dabei die beiden höchsten positiven Korrelationen mit Korrelationskoeffizienten größer 0,5 auf. Lediglich *C. accolens* und *C. amycolatum* weisen mit Korrelationskoeffizienten von -0,20 und -0,08 schwache negative Korrelationen auf. Alle übrigen beschriebenen Corynebakterien mit Abundanzen von mindestens 1 % in mindestens einem Datensatz korrelieren schwach bis mäßig positiv mit dem *Malodor Score*. Die dominanten Staphylokokken *S. epidermidis*, *S. aureus*, *S. haemolyticus* und *S. hominis* korrelieren negativ, wobei die Korrelation von *S. hominis* mit einem Korrelationskoeffizient von -0,03 sehr gering ausgeprägt ist. Hingegen korreliert *S. croceolyticus* mit einem relativ starken Korrelationskoeffizient von -0,56. Besonders starke negative Korrelationen zum Körpergeruch

lassen sich für *Anaerococcus lactolyticus* (-0,72) und *A. prevotii* (-0,57) beschreiben, während *A. tetradius* (0,46) und *A. hydrogenalis* (0,36) positiv korrelieren. Alle beschriebenen Propionibakterien korrelieren schwach bis mäßig positiv mit Körpergeruch. *Peptoniphilus*-Spezies sowie *Finegoldia magna* weisen lediglich schwache negative und positive Korrelationen zum Körpergeruch auf.

**Tabelle 16: Korrelation einzelner Bakterienspezies mit dem Körpergeruch.** Dargestellt sind auf Grundlage der *M5RNA*-Datenbank zugeordnete Spezies mit Sequenzabundanzen von mindestens 1 % in mindestens einem Datensatz. Die Korrelation ist als Korrelationskoeffizient zwischen den jeweiligen Abundanzen nach Normalisierung und Transformation und dem *Malodor Score* angegeben.

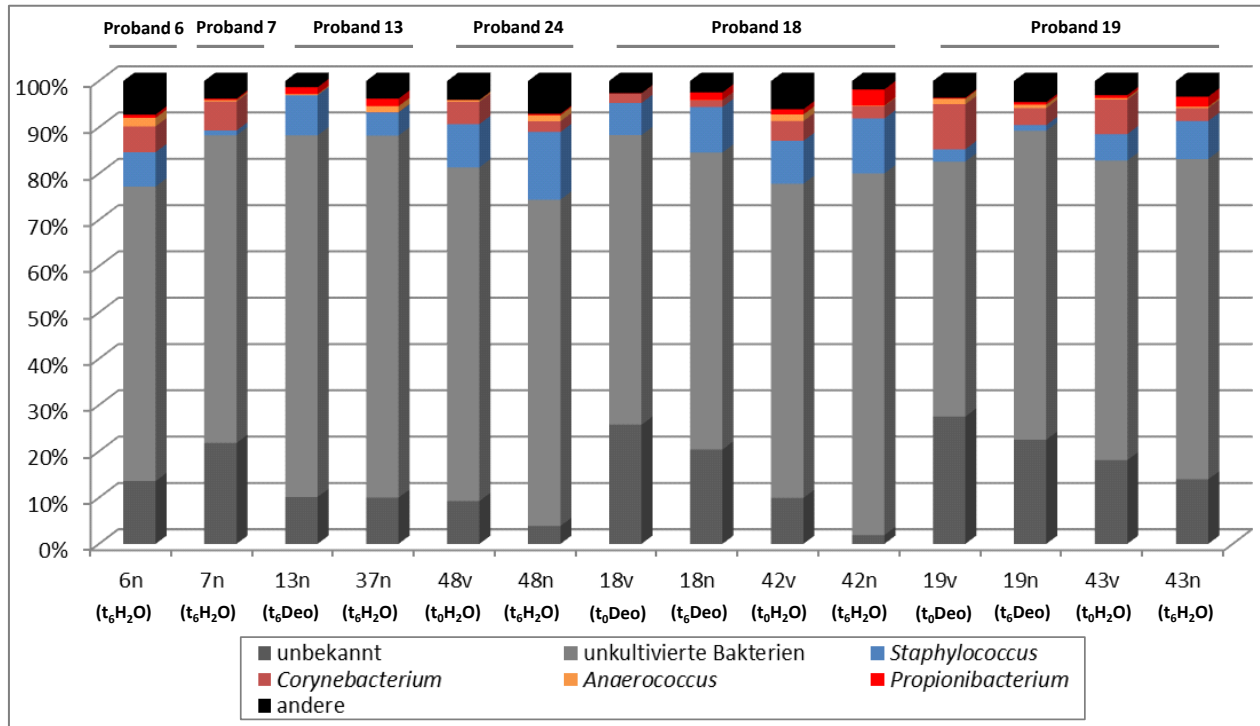
| Positiv korrelierte Spezies               |             | Negativ korrelierte Spezies         |             |
|---|-------------|-------------------------------------|-------------|
| Spezies                                   | Korrelation | Spezies                             | Korrelation |
| <i>Corynebacterium liophiloflavum</i>     | 0,61        | <i>Anaerococcus lactolyticus</i>    | -0,72       |
| <i>Corynebacterium tuscaniense</i>        | 0,52        | <i>Anaerococcus prevotii</i>        | -0,57       |
| <i>Corynebacterium sp. NML00-0156</i>     | 0,48        | <i>Staphylococcus croceolyticus</i> | -0,56       |
| <i>Corynebacterium sp. NML09-0341</i>     | 0,48        | <i>Bacillus Phage SPbeta</i>        | -0,49       |
| <i>Anaerococcus tetradius</i>             | 0,46        | <i>Staphylococcus epidermidis</i>   | -0,47       |
| <i>Corynebacterium imitans</i>            | 0,45        | <i>Staphylococcus aureus</i>        | -0,42       |
| <i>Corynebacterium appendicis</i>         | 0,44        | <i>Staphylococcus cohnii</i>        | -0,38       |
| <i>Propionibacterium acnes</i>            | 0,43        | <i>Staphylococcus capitis</i>       | -0,37       |
| <i>Corynebacterium glaucum</i>            | 0,42        | <i>Staphylococcus haemolyticus</i>  | -0,35       |
| <i>Corynebacterium genitalium</i>         | 0,40        | <i>Staphylococcus lugdunensis</i>   | -0,21       |
| <i>Propionibacterium granulosum</i>       | 0,39        | <i>Corynebacterium accolens</i>     | -0,20       |
| <i>Anaerococcus hydrogenalis</i>          | 0,36        | <i>Peptoniphilus lacrimalis</i>     | -0,17       |
| <i>Corynebacterium mucifaciens</i>        | 0,36        | <i>Corynebacterium amycolatum</i>   | -0,08       |
| <i>Corynebacterium coyleae</i>            | 0,30        | <i>Staphylococcus hominis</i>       | -0,03       |
| <i>Corynebacterium striatum</i>           | 0,29        |                                     |             |
| <i>Corynebacterium ammoniagenes</i>       | 0,29        |                                     |             |
| <i>Propionibacterium sp. 434-HC2</i>      | 0,26        |                                     |             |
| <i>Corynebacterium tuberculostearicum</i> | 0,24        |                                     |             |
| <i>Corynebacterium aurimucosum</i>        | 0,14        |                                     |             |
| <i>Corynebacterium sp.</i>                | 0,13        |                                     |             |
| <i>Corynebacterium pseudigenitalium</i>   | 0,09        |                                     |             |
| <i>Peptoniphilus sp. oral taxon 386</i>   | 0,08        |                                     |             |
| <i>Finegoldia magna</i>                   | 0,07        |                                     |             |

### 2.3.4 16S-rRNA-basierte Darstellung der Achselhöhlen-Mikrobiota auf Grundlage der *GenBank*-Datenbank

Eine alternative Möglichkeit der taxonomischen Klassifizierung der Metatranskriptomik-Datensätze mit MG-RAST stellt die 16S-rRNA-basierte Klassifizierung auf Grundlage der *GenBank*-Datenbank außerhalb von MG-RAST dar. Eine rRNA-basierte taxonomische Klassifizierung der Datensätze auf Grundlage der *GenBank*-Datenbank ist in MG-RAST nicht möglich. Da die *GenBank*-Datenbank jedoch die mit Abstand größte und aktuellste Datenbank darstellt, in der auch Datensätze nicht in Fachzeitschriften publizierter Haut-Mikrobiom-Studien enthalten sind, kann diese Methode der Klassifizierung genutzt werden, um Sequenzen mit Homologie zu rRNA-Sequenzen nicht kultivierter Hautbewohner darzustellen. Weiterhin kann der Benutzer die Aktualität der *M5RNA*- und *M5NR*-Datenbanken in MG-RAST nicht steuern. Daher bietet die Methode ebenfalls die Möglichkeit gegebenenfalls ein aktuelleres taxonomisches Profil zu generieren.

Im Rahmen dieser Arbeit wurden diejenigen Sequenzen, die in MG-RAST als Annotationen zu *SSU*-rRNA-Sequenzen identifiziert wurden, heruntergeladen und mit einem Schwellenwert von 97% Sequenzähnlichkeit geclustert. Die repräsentativen Sequenzen der Cluster wurden anschließend in der *GenBank*-Datenbank abgefragt und jeweils der erste Treffer zur Auswertung herangezogen. Bei der Auswertung wurde zusätzlich auf die jeweilige *Read*-Abundanz zurückgerechnet. Die entsprechenden Skripte wurden von Daniel Wibberg am CeBiTec zur Verfügung gestellt. In Abbildung 21 ist die Verteilung der relativen *Read*-Abundanzen für jeden Datensatz auf der Genus-Ebene dargestellt. Auffallend ist, dass der größte Anteil der Sequenzen (68,4 % +/- 6,4 %) eine Homologie zu rRNA-Sequenzen unkultivierter Bakterien (in hellgrau dargestellt) aufweist. Darüber hinaus wurden im Mittel 14,9 % +/- 7,6 % der Sequenzen mit dieser Methode nicht klassifiziert (in dunkelgrau dargestellt). Insgesamt werden 74 % (48n) bis 88 % (18v) der Sequenzen durch Annotationen zu unbekanntem oder unkultivierten Bakterien beschrieben (Abbildung 21). Dieses Phänomen ist in der Literatur beschrieben und begründet im Allgemeinen die Durchführung kultivierungsunabhängiger auf *High-Throughput*-Sequenziermethoden beruhender Ansätze zur Beschreibung der Haut-Mikrobiota auf der Ebene des Metagenoms und Metatranskriptoms (Fredrich et al., 2013). Der übrige Anteil der Sequenzen kann Homologie-basiert im Wesentlichen durch Annotationen zu den vier Genera *Staphylococcus*, *Corynebacterium*, *Anaerococcus* und *Propionibacterium* beschrieben werden. Sequenzen mit Homologie zu anderen Bakteriengattungen besitzen einen Anteil von etwa 1 % (13n) bis 7 % (6n) der *Reads* (Abbildung 21). Bei näherer Betrachtung des Anteils der Sequenzen, der einer Bakteriengattung zugeordnet werden konnte, sind Ähnlichkeiten zur rRNA-basierten Zuordnung in MG-RAST erkennbar (vgl. Abbildung 15). Neben den unbekanntem und

unkultivierten Bakterienspezies werden die Datensätze der Proben 13n, 37n, 48n, 18v, 18n, 42v und 42n von *Staphylokokken* dominiert (Typ I). In den Datensätzen der Proben 7n, 19v und 19n dominieren *Corynebakterien* (Typ II) und in den Datensätzen der Proben 6n, 48v und 43v dominieren sowohl *Staphylokokken* als auch *Corynebakterien* den Anteil kultivierbarer Bakteriengattungen (Typ III).



**Abbildung 21:** 16S-rRNA-basierte taxonomische Darstellung der Metatranskriptomik-Datensätze der Achselhöhle auf Genus-Ebene unter Verwendung der *GenBank*-Datenbank. Dargestellt sind die Verteilungen der besten Treffer für den Probanden 6 (6n), 7 (7n), 13 (13n und 37n), 24 (48v und 48n) in a), 18 (18v, 18n, 42v und 42n) und 19 (19v, 19n, 42v, 42n).

Insgesamt kann durch die *SSU*-rRNA-basierte Analyse auf Grundlage der *GenBank*-Datenbank bestätigt werden, dass die im Rahmen kultivierungsbasierter Studien in der Literatur beschriebenen Mitglieder nur einen geringen Anteil der Haut-Mikrobiota beschreiben. Der Großteil der aktiven Mitglieder der Achselhöhlen-Mikrobiota ist somit lediglich auf Grundlage von 16S-rDNA-Homologien beschrieben und klassifiziert. Funktionen, die gegebenenfalls mit der Entstehung von Körpergeruch im Zusammenhang stehen könnten, sind für diesen überwältigenden Großteil der mikrobiellen Lebensgemeinschaft vermutlich nicht beschrieben.

## 2.4 Funktionale Analyse des Metatranskriptoms der Achselhöhlen-Mikrobiota

Mithilfe der *M5NR*-Datenbank in MG-RAST können die Metatranskriptomik-Datensätze der Achselhöhlen-Mikrobiota auf funktionaler Ebene beschrieben werden. Dazu können die unterschiedlichen Ontologien der Datenbanken *GenBank*, *SEED*, *IMG*, *UniProt*, *KEGG* und *eggNOGs* verwendet werden, um die Funktionen der Datensätze auf unterschiedliche Weise zu betrachten. Während die Datenbanken *SEED* und *KEGG* beispielsweise Enzymfunktionen abbilden, kann die Datenbank *eggNOGs* mit der *COG*-Ontologie verwendet werden, um physiologische Vorgänge zu beschreiben. Im Folgenden werden zunächst die funktionalen Profile der Metatranskriptomik-Datensätze allgemein mithilfe der *COG*-Ontologie beschrieben und vergleichend dargestellt. Anschließend werden Funktionen, die im Zusammenhang mit Körpergeruchsentstehung stehen, auf Grundlage unterschiedlicher Ontologien beschrieben und auf taxonomischer Ebene klassifiziert. Schließlich werden Funktionen beschrieben, die mit dem Körpergeruch korrelieren.

### 2.4.1 Allgemeine funktionale Analyse der Achselhöhlen-Mikrobiota

Zur Beschreibung der Metatranskriptomik-Datensätze im Hinblick auf physiologische Funktionen der aktiven Achselhöhlen-Mikrobiota wurden *COG*-Profile der einzelnen Datensätze erstellt (s. Abbildung 22). Dazu wurden die jeweiligen Datensätze mithilfe der *Hierarchical Classification* auf Grundlage der *COG*-Ontologie funktional klassifiziert und die Profile auf der Hierarchieebene 2 (*level 2*) dargestellt. Im Mittel konnten auf diese Weise 178.193 Sequenzen pro Datensatz klassifiziert werden. Jedoch variierten die Anzahlen klassifizierter Sequenzen sehr stark zwischen 522.892 Sequenzen für den Datensatz der Probe 6n und lediglich 223 Sequenzen für den Datensatz der Probe 7n und 3.656 Sequenzen für den Datensatz der Probe 13n. Bei Betrachtung der in Abbildung 22 dargestellten funktionalen Profile heben sich die Profile der Datensätze 13n und 7n besonders stark von allen anderen Profilen ab.

In den *COG*-Profilen der zellphysiologischen Funktionen stellen die Funktionen Energiegewinnung und –umwandlung (orange) sowie Translation, ribosomale Struktur und Biogenese (hellgrün) die *COG*-Klassen mit den höchsten *Read*-Abundanzen in nahezu allen Datensätzen dar (s. Abbildung 22). Die *COG*-Klassen mit dem geringsten Anteil der *Read*-Abundanzen werden durch die Funktionen Zellkernstruktur und Zytoskelett (blau) beschrieben, denen in den meisten Datensätzen keine Sequenzen zugeordnet werden. Diese Beobachtung steht in Einklang mit der überwältigenden Dominanz der Datensätze durch bakterielle Sequenzen und der daraus abgeleiteten Dominanz bakterieller

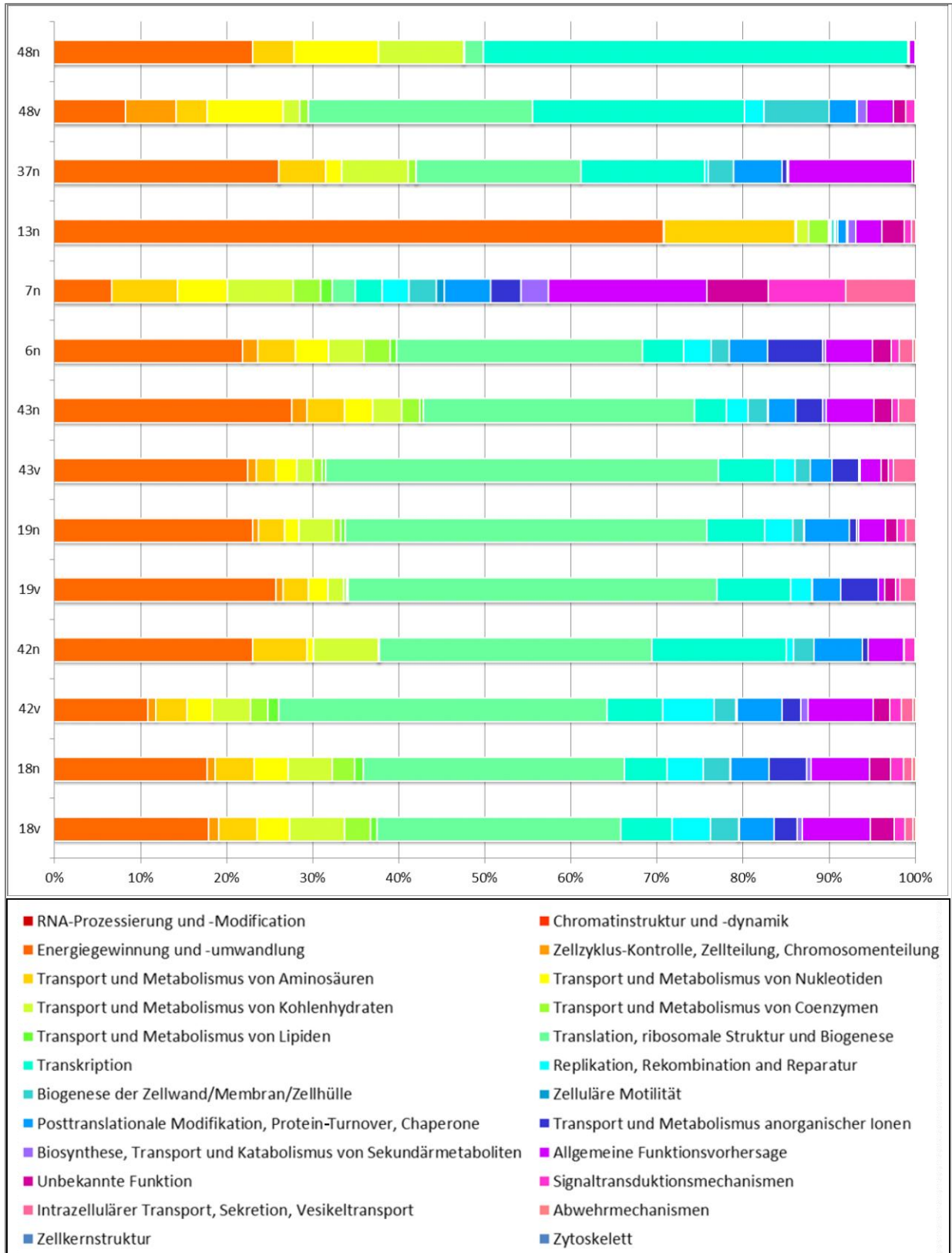


Abbildung 22: Funktionale Profile der Metatranskriptomik-Datensätze der Achselhöhle nach der COG-Klasseneinteilung. Dargestellt sind die biologischen Funktionen für jeden Metatranskriptomik-Datensatz.

Mitglieder der transkriptionell aktiven Achselhöhlen-Mikrobiota. Bei Betrachtung der funktionalen Profile einzelner Probanden lassen sich keine deutlichen intraindividuellen Ähnlichkeiten erkennen (von den Profilen der Proben 7n und 13n abgesehen). Somit erscheint die intraindividuelle Ähnlichkeit auf funktionaler Ebene gering im Vergleich zu den teilweise sehr deutlichen Ähnlichkeiten Staphylokokken-dominiierter Proben auf taxonomischer Ebene (vgl. Abbildung 15).

Um Ähnlichkeiten und Unterschiede auf funktionaler Ebene besser darstellen zu können und gegebenenfalls besonders variable Funktionen zu beschreiben, wurde eine Hauptkomponenten-Analyse auf Grundlage der 12 abundantesten COG-Klassen durchgeführt (s. Abbildung 23). Im *Loading-Plot* in Abbildung 23 a) ist zu sehen, dass die beiden abundantesten Funktionen Translation (1) und Energiemetabolismus (2) die Vektoren mit den größten Beträgen darstellen. Jedoch ist keiner der beiden Vektoren mit einer der beiden Hauptkomponenten kongruent. Die Vektoren der übrigen zehn COG-Kategorien besitzen deutlich geringere Beträge. Die Vektoren der COG-Klassen anorganische Ionen (4) und unbekannte Funktionen (10) weisen dabei in die Richtung der Hauptkomponente 1, während die Vektoren der Funktionsklassen Kohlenhydrate (6), Chaperone (8) und allgemeine Funktionsvorhersage (3) in Richtung der Hauptkomponente 2 weisen.

Der *PCA-Plot* ist in Abbildung 23 b) dargestellt und zeigt die Lage der 14 Datenpunkte relativ zu einander. Auffällig ist, dass die meisten Datenpunkte in den Quadranten I, III und IV liegen, während sich lediglich der Datenpunkt der Probe 43v im Quadranten II befindet. Im Vergleich zur Hauptkomponentenanalyse auf taxonomischer Ebene ist auf funktionaler Ebene kein starkes *Clustering* der Datenpunkte 18v, 18n, 42v und 42n des Probanden 18 erkennbar (vgl. Abbildung 16). Stattdessen liegen Datenpunkte von Proben unterschiedlicher Probanden sehr nahe beieinander, und es lassen sich keine Cluster erkennen, die auf die intraindividuelle Ähnlichkeit von Proben eines Probanden schließen lassen. Ebenso wenig clustern Datenpunkte von Proben vor bzw. nach Deo- oder Wasser-Behandlung. Dies lässt schließen, dass die durch das Experiment bedingten funktionalen Unterschiede der einzelnen Proben auf der hier gewählten Betrachtungsebene sehr gering sind. Weiterhin sind auf dieser Betrachtungsebene keine funktionalen Kategorien erkennbar, die im Zusammenhang mit einer *Staphylokokken*- oder *Corynebakterien*-dominierten Achselhöhlen-Mikrobiota stehen könnten. Auffallend sind jedoch die Positionen der Datenpunkte 42v und 43v, die eine besonders große Distanz zu allen übrigen Datenpunkten aufweisen, was auf sehr deutliche Unterschiede dieser beiden funktionalen Profile im Vergleich zu denen anderer Proben hindeutet. Bezeichnend ist weiterhin, dass die



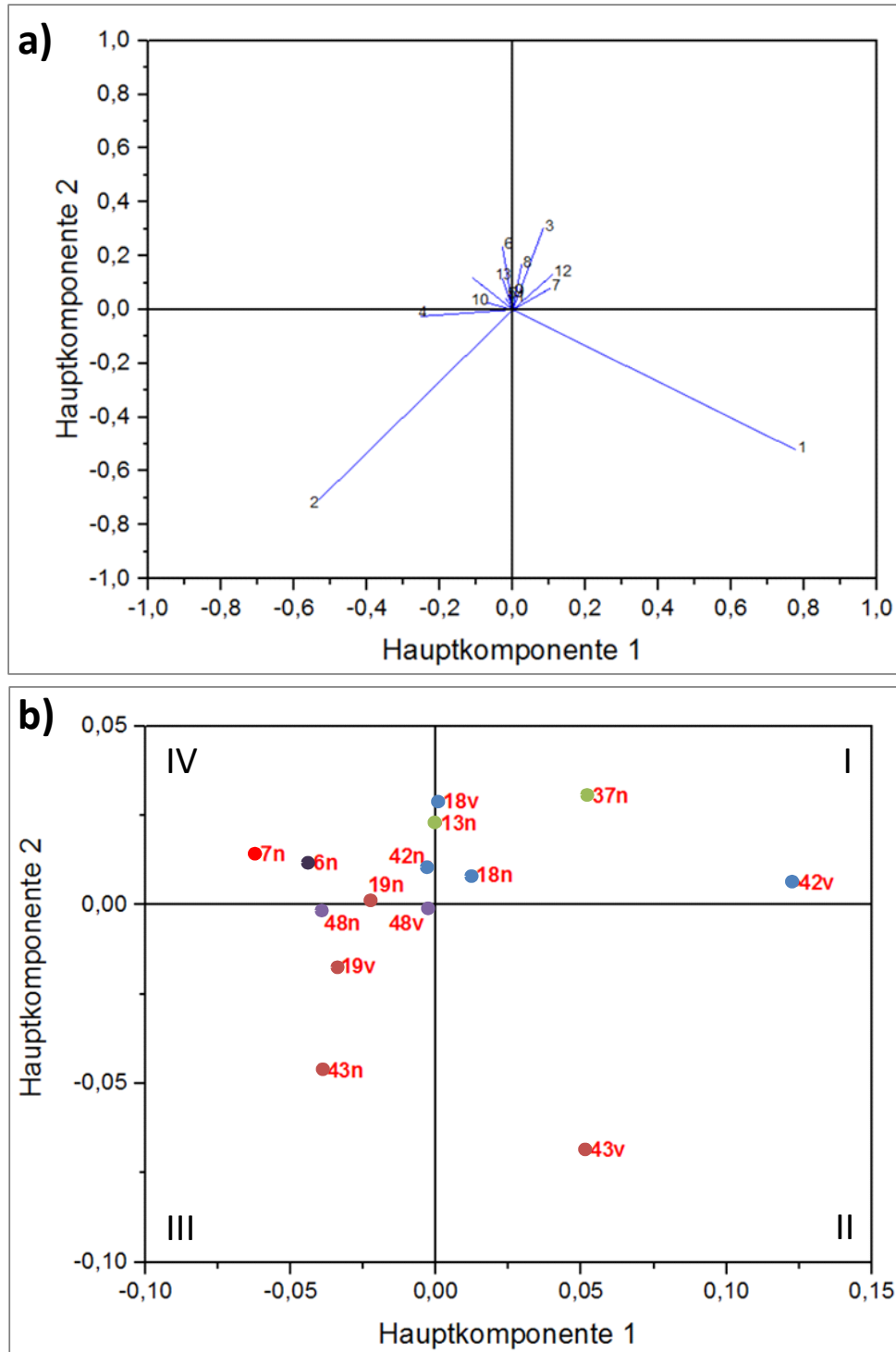


Abbildung 23: Hauptkomponentenanalyse (PCA) der funktionalen Zuordnung der Metatranskriptomik-Datensätze auf Grundlage der COG-Klasseneinteilung. a) *Loading-Plot* für die funktionalen Klassen Translation (1), Energiemetabolismus (2), allgemeine Funktionsvorhersage (3), anorganische Ionen (4), Aminosäuren (5), Kohlenhydrate (6), Transkription (7), Chaperone (8), Nukleotide (9), unbekannte Funktion (10), Zellhülle (11), Replikation (12), Coenzyme (13); b) *PCA-Plot* mit den Proben der Probanden 6 (6n, schwarz), 7 (7n, hellrot), 13 (13n und 37n, grün), 24 (48v und 48n, lila), 18 (18v, 18n, 42v und 42n, blau) und 19 (19v, 19n, 43v und 43n, braun).

Verschiebung der Datenpunkte 18v ( $t_0$  Deo) und 18n ( $t_6$  Deo) des Probanden 18 bzw. 19v ( $t_0$  Deo) und 19n ( $t_6$  Deo) des Probanden 19, die im Zusammenhang mit Deodorant-Anwendung stehen, überwiegend entlang der Hauptkomponente 2 (Kohlenhydrate, Chaperone und allgemeine Funktionsvorhersage) erfolgt. Dahingegen erfolgt die Verschiebung der Datenpunkte 42v ( $t_0$  H<sub>2</sub>O) und 42n ( $t_6$  H<sub>2</sub>O) bzw. 43v und 43n, die im Zusammenhang mit Wasser-Behandlung stehen, überwiegend entlang der Hauptkomponente 1 (anorganische Ionen und unbekannte Funktionen). Die Verschiebung der Datensätze im Zusammenhang mit Wasser-Behandlung fällt insgesamt auch deutlich stärker aus als die Verschiebung der Datenpunkte der Proben der selben Probanden im Zusammenhang mit Deodorant-Anwendung. Dieses Phänomen ist auf taxonomischer Ebene in dieser Form nicht zu beobachten und legt somit eine starke Veränderung der Funktionen der Achselhöhlen-Mikrobiota auf RNA-Ebene nach Wasser-Behandlung nahe.

In den folgenden Kapiteln werden funktionale Zuordnungen innerhalb der Metatranskriptomik-Datensätze beschrieben, die in der Literatur im Zusammenhang mit Körpergeruchsentstehung beschrieben sind.

#### **2.4.2 Analyse von Transkripten aus *Corynebacterium striatum* im Zusammenhang mit Körpergeruchsentstehung**

In der Literatur wurden spezifische Funktionen aus *Corynebacterium striatum* Ax20 im Zusammenhang mit Körpergeruchsentstehung beschrieben (s. Einleitung). Im Rahmen dieser Arbeit wurden ebenfalls in einigen Proben Sequenzen mit Homologie zu rRNA-Sequenzen aus *C. striatum* mithilfe von MG-RAST beschrieben (s. Kapitel 2.3.3). Daher wurden zunächst auf Grundlage der *GenBank*-Annotationen der *M5NR*-Datenbank Sequenzen mit Ähnlichkeit zu den Sequenzen der N<sup>α</sup>-Acylglutamin-Aminoacylase (N-AGA), Cystathionin-β-Lyase und *Thiol Precursor Dipeptidase* TpdA aus *C. striatum* untersucht.

In Tabelle 17 ist für jeden Datensatz die Anzahl der identifizierten Sequenzen sowie die mittlere Sequenzidentität zu den Referenzsequenzen aus *C. striatum* dargestellt. Es zeigt sich, dass die Datensätze der Proben 18v, 18n und 42v des Probanden 18 sowie die Proben der Datensätze 13n und 37n des Probanden 13, die auf rRNA-Ebene von Annotationen zu Staphylokokken dominiert werden, keine Sequenzen mit Ähnlichkeit zu den drei Funktionen aus *C. striatum* aufweisen. Auch in den Datensätzen der Proben 48v und 48n des Probanden 24 sowie 6n des Probanden 6 und 42n des Probanden 18 befinden sich keine Sequenzen mit Ähnlichkeit zu N<sup>α</sup>-Acylglutamin-Aminoacylase und Cystathionin-β-Lyase. Lediglich 1 (48v), 3 (48n), 1 (6n) bzw. 2 (42n) Sequenzen weisen Ähnlichkeiten zu der TpdA-Dipeptidase aus *C. striatum* auf. Diese

**Tabelle 17:** Analyse der für *Corynebacterium striatum* im Zusammenhang mit Geruchsentstehung beschriebenen Funktionen in den Metatranskriptomik-Datensätzen der Achselhöhle. Dargestellt ist die identifizierte Anzahl der Reads und die mittlere Sequenzidentität zu den Sequenzen mit Homologie zu der N<sup>α</sup>-Acylglutamin-Aminoacylase, Cystathionin-β-Lyase und TpdA-Dipeptidase.

| Probe | N <sup>α</sup> -Acylglutamin-Aminoacylase |             | Cystathionin-β-Lyase |             | TpdA Dipeptidase |             |
|-------|---|-------------|----------------------|-------------|------------------|-------------|
|       | # Reads                                   | % Identität | # Reads              | % Identität | # Reads          | % Identität |
| 6n    | -   | -           | -                    | -           | 1                | 81,0        |
| 7n    | 4   | 72,7        | 8                    | 78,0        | 62               | 83,4        |
| 13n   | -   | -           | -                    | -           | -                | -           |
| 37n   | -   | -           | -                    | -           | -                | -           |
| 48v   | -   | -           | -                    | -           | 1                | 97,9        |
| 48n   | -   | -           | -                    | -           | 3                | 80,6        |
| 18v   | -   | -           | -                    | -           | -                | -           |
| 18n   | -   | -           | -                    | -           | -                | -           |
| 42v   | -   | -           | -                    | -           | -                | -           |
| 42n   | -   | -           | -                    | -           | 2                | 91,6        |
| 19v   | 10  | 76,8        | -                    | -           | 29               | 75,5        |
| 19n   | 6   | 75,1        | -                    | -           | 10               | 75,0        |
| 43v   | 8   | 76,7        | 3                    | 84,7        | 6                | 79,2        |
| 43n   | 5   | 74,5        | 2                    | 80,4        | 4                | 75,0        |

Datensätze sind auf rRNA-Ebene ebenfalls durch Annotationen von Staphylokokken oder von Staphylokokken und Corynebakterien dominiert (vgl. Kapitel 2.3.1). Der relative geringe Anteil an Sequenzen mit Homologie zu Sequenzen aus Corynebakterien auf RNA-Ebene steht somit im Einklang mit dem geringen Anteil oder dem Fehlen corynebakterieller funktionaler Annotationen in diesen Datensätzen. Die Datensätze der Proben 7n des Probanden 7 sowie 19v, 19n, 43v und 43n des Probanden 19, die, mit Ausnahme von 43n, auf rRNA-Ebene von corynebakteriellen Annotationen dominiert werden, weisen Sequenzähnlichkeiten zu den an der Geruchsentstehung beteiligten Funktionen aus *C. striatum* auf. Jedoch enthalten die Datensätze 19v und 19n keine Sequenzen mit Ähnlichkeit zur Cystathionin-β-Lyase. Insgesamt können pro Datensatz für die N<sup>α</sup>-Acylglutamin-Aminoacylase maximal 10 Sequenzen (19v), für die Cystathionin-β-Lyase maximal 8 Sequenzen (7n) und für die TpdA-Dipeptidase maximal 62 Sequenzen (7n) mit Ähnlichkeit zur jeweiligen Referenzsequenz von *C. striatum* beschrieben werden. Die Sequenzähnlichkeiten liegen dabei für die N<sup>α</sup>-Acylglutamin-Aminoacylase zwischen 72,7 % und 76,8 %, für die Cystathionin-β-Lyase zwischen 78,0 % und 84,7 % und für die

Dipeptidase zwischen 75,0 % und 97,9 %. Es wurden keine Sequenzen mit einer Sequenzähnlichkeit von 100 % zu einer der drei Funktionen aus *C. striatum* identifiziert.

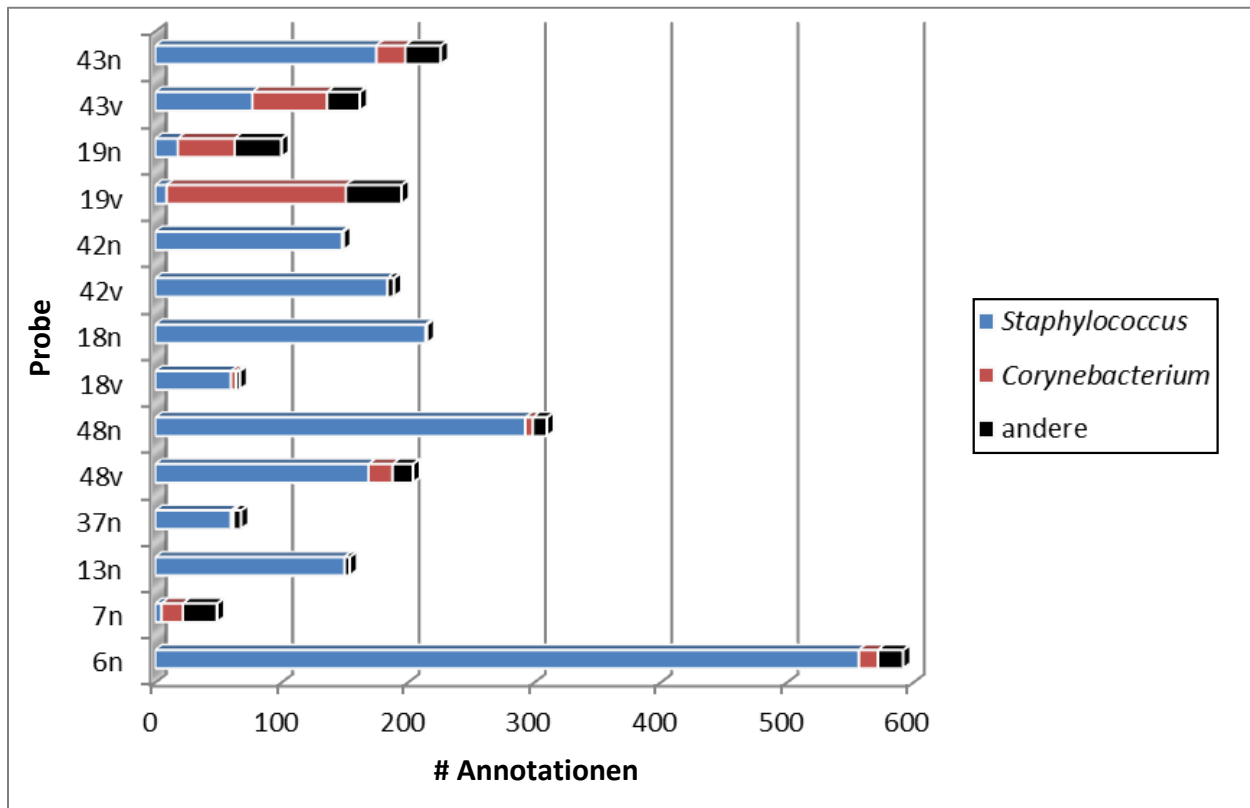
Die Anzahlen der identifizierten Sequenzen variieren zwar z. T. sehr stark, jedoch ist kein Zusammenhang mit dem *Malodor Score* erkennbar. Die im Vergleich zur Gesamtzahl der *Reads* geringe Anzahl sowie das Fehlen von Sequenzähnlichkeiten von 100 % sprechen dafür, dass die von Natsch et al. beschriebenen Funktionen aus *Corynebacterium striatum* in den in dieser Arbeit untersuchten Proben nicht vorhanden sind. Somit ist zu vermuten, dass diese Funktionen zu anderen Organismen annotiert werden. Daher wird im Folgenden jede dieser Funktionen in den Datensätzen näher beschrieben.

### 2.4.3 Analyse von Aminoacylase-Transkripten

Die von Natsch et al. beschriebene Proteinsequenz der N<sup>α</sup>-Acylglutamin-Aminoacylase besitzt die *GenBank-Accession* AAN77164.1 und wird in der *CDD*-Datenbank der Gruppe *M20\_Acy1\_like6* und auf unspezifischerer Ebene der *COG*-Gruppe COG1473 mit der Annotation Metall-abhängige Amidase/Aminoacylase/Carboxypeptidase zugeordnet. Die Familie dieser M20 Zink-abhängigen Peptidasen ist sehr groß, divers und wenig charakterisiert (Marchler-Bauer et al., 2015). Die M20-Peptidasen zeichnen sich im Allgemeinen durch geringe Homologie auf Sequenzebene bei z. T. sehr hoher funktionaler Konvergenz aus (Makarova und Grishin, 1999). Typisch für diese Familie ist jedoch ein Zinkbindemotiv mit der Aminosäure-Sequenz HEXXH, HXXEH oder HXH (Gomis-Rüth, 2008; Hooper, 1994). Jedoch konnte in der Sequenz der Aminoacylase von *C. striatum* sowie anderen Homologen aus anderen Organismen kein solches Motiv identifiziert werden (Natsch et al., 2003). Aufgrund der Tatsache, dass keine Sequenzen mit einer 100%igen Ähnlichkeit zu der von Natsch et al. beschriebenen Sequenz der Aminoacylase aus *C. striatum* identifiziert werden konnten, wurde geprüft, ob Sequenzen mit Ähnlichkeit zu anderen Aminoacylasen in den Datensätzen enthalten sind.

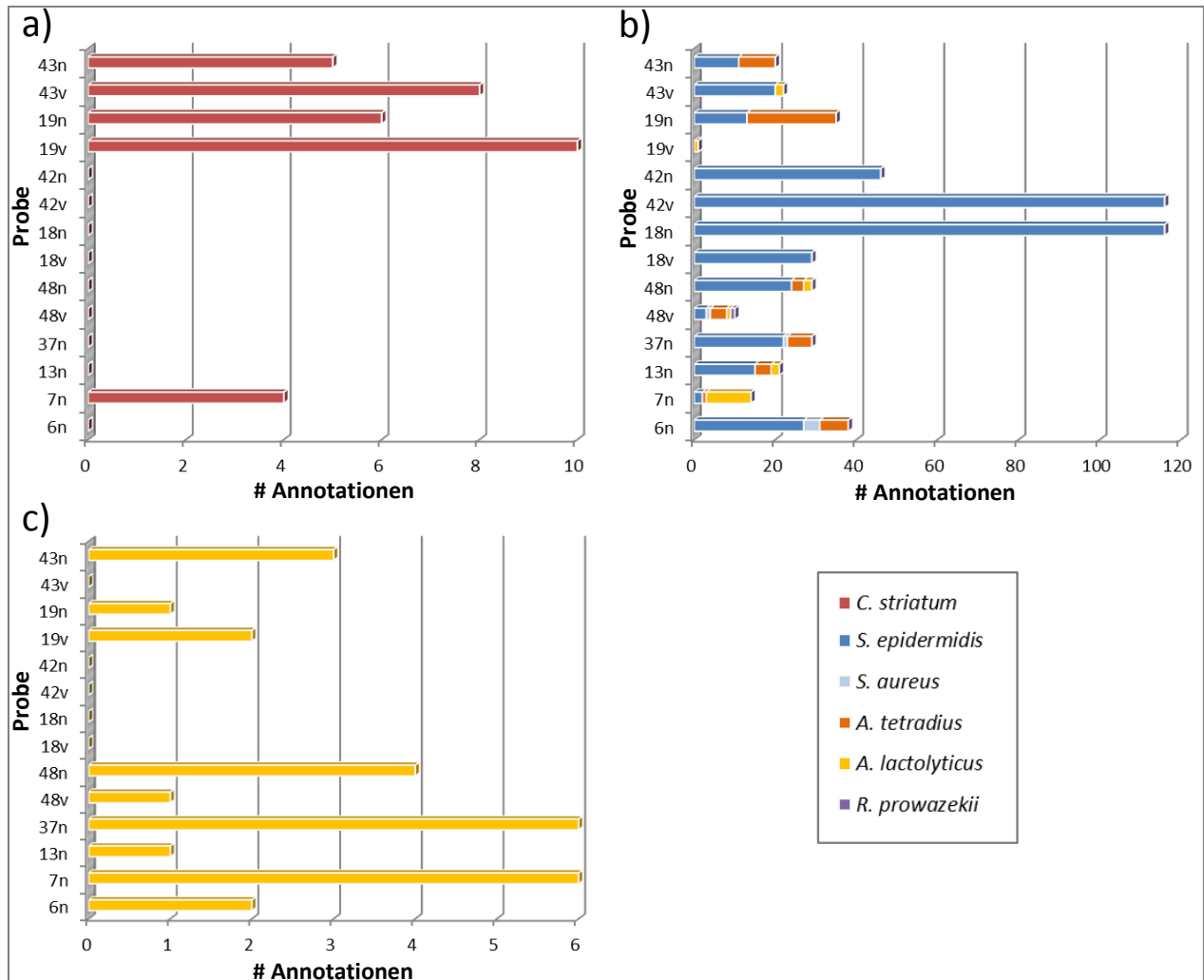
Zunächst wurden die Annotationen zur Gruppe COG1473 untersucht. Im Mittel wurden  $191 \pm 130$  Sequenzen pro Datensatz dieser COG-Funktion zugeordnet. Das Maximum lag bei 588 Sequenzen für Probe 6n und das Minimum bei 49 Sequenzen für Probe 7n. Diese Sequenzen wurden anschließend mithilfe der *Best Hit Classification* auf Genus-Ebene klassifiziert. Das entsprechende taxonomische Profil ist in Abbildung 24 dargestellt. Hier wurden im Mittel  $191 \pm 131$  Annotationen pro Datensatz zugeordnet, sodass davon ausgegangen werden kann, dass die *Best Hit Classification* die Anzahlen annotierter Sequenzen nicht übermäßig verzerrt. Mit Ausnahme der Datensätze 7n, 19v, 19n und 43v werden die meisten Sequenzen in den jeweiligen Datensätzen dem Genus *Staphylococcus* zugeordnet. In den anderen Datensätzen sind ebenfalls Zuordnungen zum Genus *Corynebacterium* dominant. Zuordnungen zu anderen

Genera sind mit Ausnahme des Datensatzes 7n eher selten. Somit wird die Funktion COG1473 ähnlich wie die gesamten Metatranskriptomik-Datensätze von Staphylokokken- und Corynebakterien-Sequenzen dominiert. Die Analyse der Korrelation zwischen den Annotationen zur Funktion COG1473 und dem *Malodor Score* lieferte einen leicht negativen Korrelationskoeffizienten von -0,37.



**Abbildung 24:** Funktionales taxonomisches Profil der Gruppe COG1473 (Metall-abhängige Amidase/Aminoacylase/Carboxypeptidase) auf Genus-Ebene. Dargestellt ist die Anzahl der Annotationen für jede untersuchte Probe zu den Genera *Staphylococcus* und *Corynebacterium* sowie zu anderen Genera mittels *Best Hit Classification*.

Im nächsten Schritt wurden Annotationen zur Funktion Aminoacylase auf Grundlage des Genbank-Annotationensystems in MG-RAST analysiert. Dabei konnten insgesamt 5 verschiedene Annotationen identifiziert werden: N<sup>α</sup>-Acylglutamin-Aminoacylase (33 Sequenzen aus 5 Datensätzen), Aminoacylase (420 Sequenzen aus allen 14 Datensätzen), mögliche Aminoacylase (26 Sequenzen aus 9 Datensätzen) und Metall-abhängige Amidase/Aminoacylase/Carboxypeptidase und putative Metall-abhängige Amidase/Aminoacylase/Carboxypeptidase (jeweils zwei Sequenzen im Datensatz der Probe 37n). Die letzteren Annotationen wurden nicht weiterführend analysiert. Die ersten drei



**Abbildung 25:** Funktionale taxonomische Profile der GenBank-Annotationen N<sup>α</sup>-Acylglutamin-Aminoacylase (a), Aminoacylase (b) und mögliche Aminoacylase (c) auf Spezies-Ebene. Dargestellt sind die Anzahlen der Annotationen für jede untersuchte Probe zu den Spezies *Corynebacterium striatum*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Anaerococcus tetradius*, *Anaerococcus lactolyticus* und *Rickettsia prowazekii* mittels *Best Hit Classification*.

Annotationen wurden mithilfe der *Best Hit Classification* auf Spezies-Ebene klassifiziert. Die entsprechenden taxonomischen Profile sind in Abbildung 25 dargestellt. Die Annotationen können den fünf Spezies *Corynebacterium striatum*, *Staphylococcus epidermidis*, *Staphylococcus aureus*, *Anaerococcus tetradius*, *Anaerococcus lactolyticus* und *Rickettsia prowazekii* (ein *Alpha-Proteobacterium*) zugeordnet werden. Annotationen zur N<sup>α</sup>-Acylglutamin-Aminoacylase liegen lediglich in den Datensätzen der Proben 7n, 19v, 19n, 43v und 43n vor, welche auf taxonomischer Ebene im Vergleich zu den anderen Datensätzen einen großen Anteil an Corynebakterien aufweisen. Die Annotationen werden ausschließlich der Spezies *C. striatum* zugeordnet (s. Abbildung 25 a). Die Annotationen zur Aminoacylase werden hauptsächlich den

Spezies *S. epidermidis* zugeordnet. Darüber hinaus gibt es Anntationen zu *S. aureus*, *A. tetradius*, *A. lactolyticus* und, im Falle von einer Sequenz im Datensatz der Probe 48v, *R. prowazeki* (s. Abbildung 25 b). Annotationen zur möglichen Aminoacylase werden ausschließlich *A. lactolyticus* zugeordnet und fehlen in den Datensätzen der Proben 18v, 18n, 42v und 4n des Probanden 18 sowie im Datensatz 43n des Probanden 19 vollständig (s. Abbildung 25 c).

|            | 80 | *  | 100   | *  | 120          | *                   | 140                 |            |      |
|------------|----|--|-------|--|--------------|---------------------|---------------------|------------|------|
| 7n_3       | :  | -----                                      | ----- | DYAADPEALGRMHACGHC                         | MHT          | TALIGAAEALMQHKDAWSG | TFIALFQPGE          | : 49       |      |
| 19n_4      | :  | -----                                      | ----- | DPELGRMHACGHC                              | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 45       |      |
| 43n_4      | :  | -----                                      | ----- | TGLDYAADPELGRMHACGHC                       | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 52       |      |
| 43n_2      | :  | -----                                      | ----- | EATGLDYAADPELGRMHACGHC                     | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 54       |      |
| 19v_1      | :  | -----                                      | ----- | LDYAADPELGRMHACGHC                         | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 50       |      |
| 19n_1      | :  | -----                                      | ----- | LDYAADPELGRMHACGHC                         | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 50       |      |
| 19n_2      | :  | -----                                      | ----- | -----                                      | PELGRMHACGHC | MHT                 | AALLGAAEALAQHKDAWSG | TFIAPFQ--- | : 41 |
| 43v_1      | :  | -----                                      | ----- | FRADFDALPITEATGLDYAADPELGRMHACGHC          | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 65       |      |
| Cure_hp    | :  | NGEGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT   | AALLGAAEALAQHKDAWSG                        | TFIALFQPGE   |                     |                     | : 129      |      |
| 43n_3      | :  | -----                                      | ----- | CFRADFDALPITEATGLDYAADPELGRMHACGHC         | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 66       |      |
| 19n_6      | :  | -----                                      | ----- | ALPITEATGLDYAADPELGRMHACGHC                | MHT          | AALLGAAEALAQHKDA    | -----               | : 46       |      |
| 19n_5      | :  | -----                                      | ----- | TGLDYAADPELGRMHACGHC                       | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 52       |      |
| 19n_3      | :  | -----                                      | ----- | ITEATGLDYAADPELGRMHACGHC                   | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 56       |      |
| 19v_4      | :  | -----                                      | ----- | VVCFRADFDALPITEATGLDYAADPELGRMHACGHC       | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 68       |      |
| 19v_7      | :  | -----                                      | ----- | ITEATGLDYAADPELGRMHACGHC                   | MHT          | AALLGAAEALAQHKDAWS  | -----               | : 45       |      |
| 7n_2       | :  | -----                                      | ----- | NGDGPTVLFRADFDALPLAEEETGLDYAADPELGRMHACGHC | MHT          | TAL                 | -----               | : 66       |      |
| 7n_1       | :  | -----                                      | ----- | NGDGPTVLFRADFDALPLAEEETGLDYAADPELGRMHACGHC | MHT          | TALIGAAEALMQHKD     | -----               | : 70       |      |
| 7n_4       | :  | -----                                      | ----- | NGDGPTVLFRADFDALPLAEEETGLDYAADPELGRMHACGHC | MHT          | TALIGAAEALMQHKDAW   | -----               | : 64       |      |
| Cimi_hp    | :  | NGDGPTVCFRADFDALPLAEEETGLDYAADPELGRMHACGHC | MHT   | AALLGATKLLTEHRDAWSG                        | TFIALFQPAE   |                     |                     | : 146      |      |
| Cstr_N-AGA | :  | NGEGPSVAFRADFDALPITENTGLDYAADPELGRMHACGHC  | DLHT  | TALLGAVRALVENKDLWSG                        | TFIAPHQPGE   |                     |                     | : 141      |      |
| 43v_2      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEA                      | -----        | -----               | -----               | : 26       |      |
| 43v_3      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITQATGLDYAADPELGRMHACGHC  | MHT          | -----               | -----               | : 49       |      |
| 43n_1      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPEL           | -----        | -----               | -----               | : 68       |      |
| 43n_5      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | -----               | -----               | : 69       |      |
| 43v_6      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGR         | -----        | -----               | -----               | : 55       |      |
| 43v_4      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | AALLGAAEALAQHKDAWSG | TFI                 | : 82       |      |
| 19v_10     | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | AALLGAAEALAQHKDAWSG | TFI                 | : 68       |      |
| 19v_9      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | -----               | -----               | : 68       |      |
| 19v_6      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | -----               | -----               | : 67       |      |
| 19v_2      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | -----               | -----               | : 48       |      |
| 19v_5      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | AALLGAAEALAQHKDAWSG | TFIALFQPGE          | : 87       |      |
| 19v_8      | :  | -----                                      | ----- | NGDGPVVCFRADFDALPITEATGLDYAADPELGRMHACGHC  | MHT          | AALLGAAEALAQHKDAWSG | TFIALF              | : 82       |      |

tgldyaadpelgrmhacghd ht

Abbildung 26: Ausschnitt des *Multiple Sequence Alignment* der Sequenzen mit der *GenBank*-Annotation zur  $N^{\alpha}$ -Acylglutamin-Aminoacylase in *MG-RAST* sowie der Referenzsequenzen aus der händischen Abfrage in der *GenBank*-Datenbank: hypothetisches Protein aus *Corynebacterium ureicelerivorans* (*Cure\_hp*), hypothetisches Protein aus *Corynebacterium imitans* (*Cimi\_hp*) und  $N^{\alpha}$ -Acylglutamin-Aminoacylase aus *Corynebacterium striatum* (*Cstr\_N-AGA*). Die Sequenzen aus den Metatranskriptomik-Datensätzen sind den jeweiligen Proben zugeordnet und fortlaufend nummeriert. Nicht dargestellt sind die Sequenzen 19v\_3, 43v\_5, 43v\_7 und 43v\_8, da es sich hierbei vermutlich um Chimären handelt.

Die Sequenzen mit Annotation zur  $N^{\alpha}$ -Acylglutamin-Aminoacylase aus *C. striatum* wurden anschließend händisch in der Online-Version der *GenBank*-Datenbank unter Verwendung des *BLASTX*-Algorithmus abgefragt. Bei den identifizierten Referenzsequenzen (es wurde jeweils lediglich der erste Treffer mit der jeweils höchsten Sequenzidentität registriert) handelt es sich

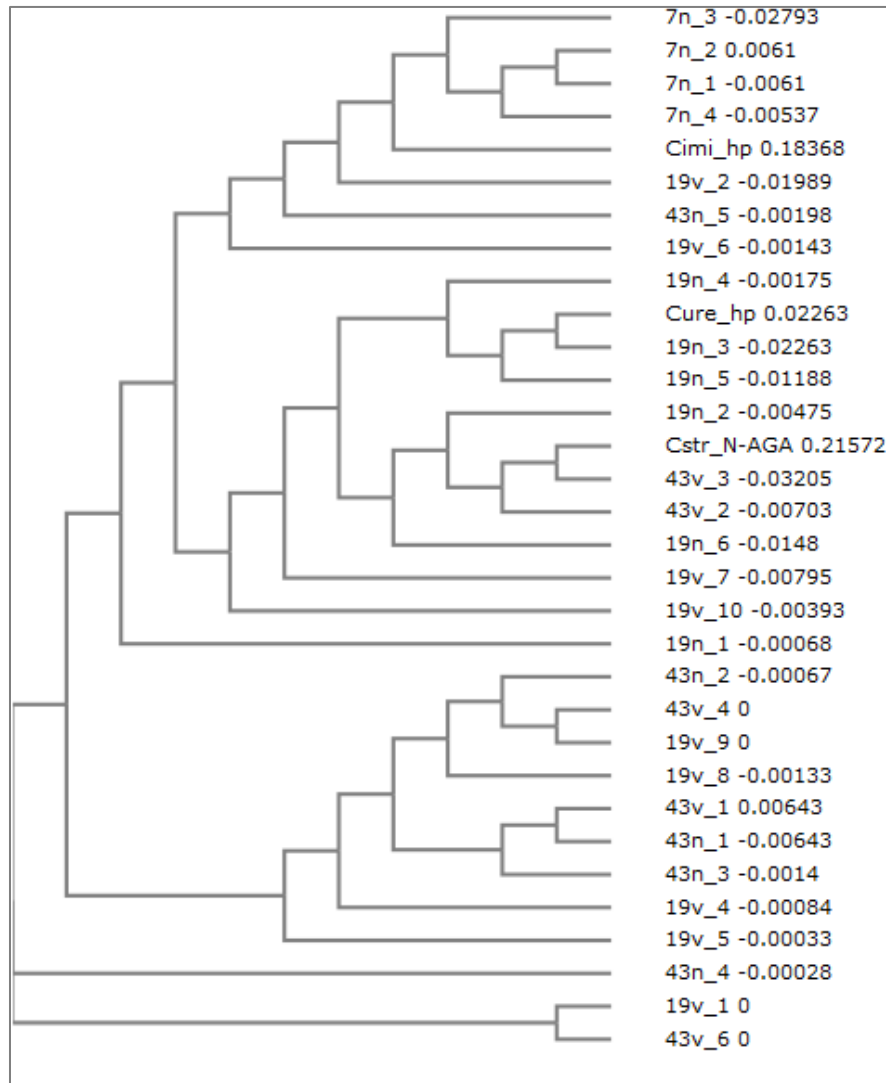


Abbildung 27: Kladogramm der in MG-RAST identifizierten Sequenzen mit Homologie zur N<sup>α</sup>-Acylglutamin-Aminoacylase und der Referenzsequenzen aus der händischen Abfrage in der GenBank-Datenbank: hypothetisches Protein aus *Corynebacterium ureicelerivorans* (Cure\_hp), hypothetisches Protein aus *Corynebacterium imitans* (Cimi\_hp) und N<sup>α</sup>-Acylglutamin-Aminoacylase aus *Corynebacterium striatum* (Cstr\_N-AGA).

um jeweils ein hypothetisches Protein aus *Coynebacterium ureicelerivorans* (GenBank-ID WP\_038613795.1) und *Corynebacterium imitans* (GenBank-ID WP\_051904756.1). Beide Organismen wurden aus dem Menschen isoliert (Funke et al., 1997; Yassin, 2007). Dabei wurde die Sequenz aus *C. imitans* lediglich den Sequenzen 7n\_1, 7n\_2 und 7n\_4 zugeordnet. Alle übrigen Sequenzen wiesen eine höhere Homologie zur Sequenz aus *C. ureicelerivorans* auf. Die mittlere Squenzidentität beträgt 92,8 %. Die mittlere Squenzidentität zu N-AGA aus *C. striatum* beträgt 73,8 % und liegt in den meisten Fällen unterhalb von 80 %. Ein *Multiple Sequence Alignment* zeigt, dass die identifizierten Sequenzen in der N-terminalen Hälfte der



Referenzsequenzen, welche der Zinkbindedomäne entspricht, alignieren (s. Abbildung 26). Dabei ist ein Abschnitt von etwa 70 Aminosäureresten stark konserviert. Dieser umfasst unter anderem auch die von Natsch et al. identifizierten konservierten Sequenzabschnitte MHACGHDXH, welches in den hier identifizierten Sequenzen die Form MHACGHDMH besitzt, und RADXDLPXXE, hier RADFDALPITE (Natsch et al., 2003).

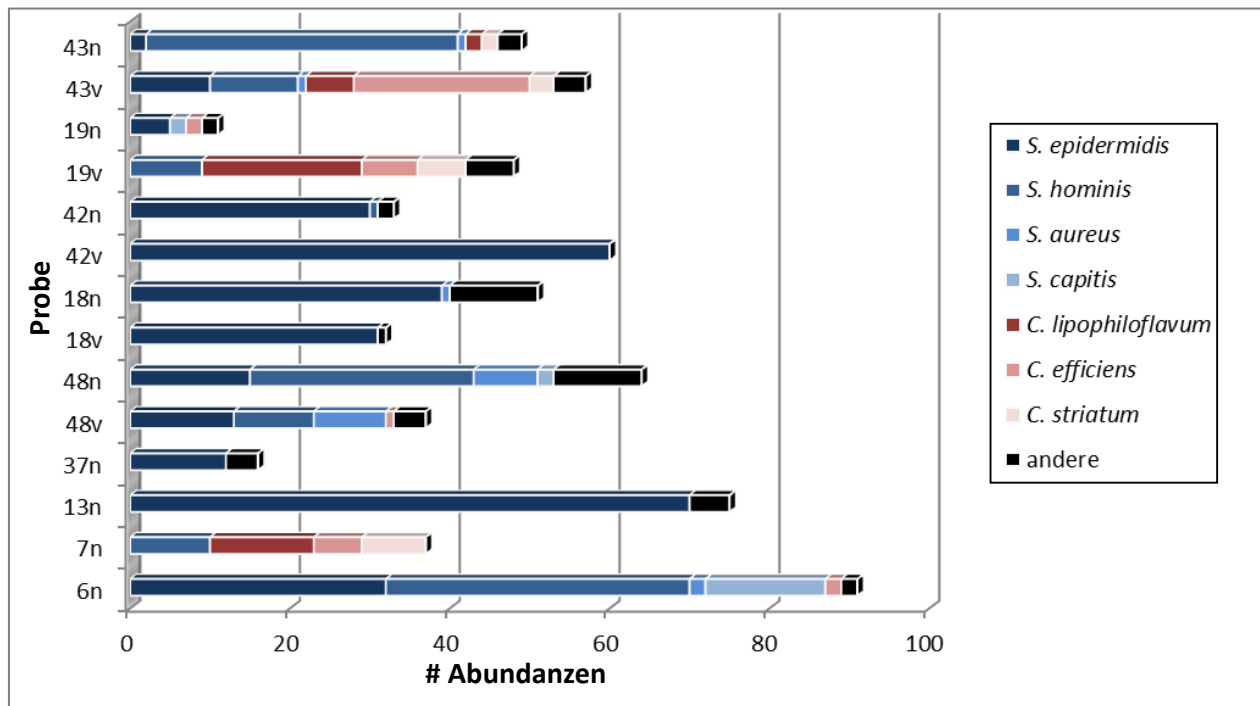
Ein mithilfe des *Online-Tools* CLUSTALOMEGA erstelltes Kladogramm zeigt die homologiebasierten Ähnlichkeiten zwischen den in den Metatranskriptomen identifizierten Sequenzen sowie den Referenzsequenzen. Auffallend ist, dass die Sequenzen aus dem Datensatz 7n des Probanden 7 mit der Referenzsequenz aus *C. imitans* clustern. Innerhalb der Sequenzen aus den Datensätzen der Proben 19v, 19n, 43v und 43n des Probanden 19 ist kein eindeutiges *Clustering* zu erkennen. Auffallend sind auch die hohen Distanzen der Referenzsequenzen aus *C. striatum* (~0,216) und *C. imitans* (~0,184) im Vergleich zu den Distanzen der übrigen Sequenzen. Dies deutet darauf hin, dass die in den vorliegenden Datensätzen identifizierten Sequenzen aus bisher uncharakterisierten Organismen stammen.

#### 2.4.4 Analyse von Cystathionin- $\beta$ -Lyase-Transkripten

Zur Identifizierung von Sequenzen, die Enzymfunktionen zur Freisetzung von Sulfanylalkanolen kodieren, wurde zunächst das *GenBank*-Annotationssystem in MG-RAST ausgewählt. In der Literatur sind die zwei Enzymfunktionen Cystein-S-Konjugat- $\beta$ -Lyase (EC 4.4.1.13) aus *Staphylococcus haemolyticus* und Cystathionin- $\beta$ -Lyase (EC 4.4.1.8) aus *Corynebacterium striatum* zur Abspaltung von körpergeruchsverursachenden Sulfanylalkanolen beschrieben (Natsch et al., 2004; Starkenmann et al., 2005). Die Suche nach Annotationen zur Cystein-S-Konjugat- $\beta$ -Lyase lieferte lediglich zwei Sequenzen mit der Annotation *possible cysteine-S-conjugate beta-lyase* im Datensatz der Probe 7n. Alle übrigen Datensätze lieferten keine Treffer. Daher wurden diese Sequenzen nicht weiterführend analysiert.

Die Suche nach Annotationen zur Cystathionin- $\beta$ -Lyase lieferte mehrere Treffer mit den folgenden funktionalen Annotationen in allen Metatranskriptomik-Datensätzen der Achselhöhlen-Proben: Cystathionin Beta Lyase (unterschiedliche Schreibweisen), Putative Cystathionin Beta Lyase, Cystathionin Beta Lyase (CBL) (Beta Cystathioninase) und Cystathionin Beta/Gamma Lyase. Im Mittel wurden 47 *Read*-Abundanzen pro Probe zugeordnet mit einem Minimum von 11 für die Probe 19n und einem Maximum von 91 für die Probe 6n. Die Klassifizierung dieser Funktionen mittels *Best Hit Classification* zeigte, dass es sich bei den dominanten Spezies um die Staphylokokken *S. epidermidis*, *S. hominis*, *S. aureus* und *S. capitis* und um die Corynebakterien *C. lipophiloflavum*, *C. efficiens* und *C. striatum* sowie um andere Spezies handelt (s. Abbildung 28). Dabei werden die Datensätze der Proben 7n des Probanden 7

sowie 19v und 43v des Probanden 19 von corynebakteriellen Funktionen dominiert. Dahingegen werden die Datensätze der Proben 13n und 37n des Probanden 13 sowie der Proben 18v, 18n, 42v und 42n des Probanden 18 von Zuordnungen zu *S. epidermidis* dominiert. Dies steht im Einklang mit dem gesamten taxonomischen Profil dieser Proben auf Genus-Ebene (vgl. Abschnitt 2.3.1). Die Datensätze der Proben 6n des Probanden 6, 48n des Probanden 24 und 43n des Probanden 19 werden von Zuordnungen zu *S. hominis* dominiert. Die Read-Abundanzen der Annotation korrelieren leicht negativ mit dem *Malodor Score* ( $r^2 = -0,26$ ).



**Abbildung 28:** Funktionales taxonomisches Profil der Annotationen zur Cystathionin- $\beta$ -Lyase auf Spezies-Ebene. Dargestellt ist die Anzahl der Annotationen für jede untersuchte Probe zu den Spezies *Staphylococcus epidermidis*, *Staphylococcus hominis*, *Staphylococcus aureus*, *Staphylococcus capitis*, *Corynebacterium lipophiloflavum*, *Corynebacterium efficiens*, *Corynebacterium striatum* sowie zu anderen Spezies mittels *Best Hit Classification*.

Die Zuordnung der jeweiligen Referenzsequenzen der in MG-RAST identifizierten Sequenzen mit Homologie zur Cystathionin- $\beta$ -Lyase wurde in der Online-Version der *Genbank*-Datenbank und in der *CDD*-Datenbank untersucht. Die taxonomische sowie die funktionale Zuordnung dieser Sequenzen in die Gruppen MetC und MalY ist in Tabelle 18 dargestellt. Auf taxonomischer Ebene wurden die meisten Sequenzen *Staphylococcus*- und *Corynebacterium*-Spezies zugeordnet. Einige wenige *Read*-Abundanzen wurden zusätzlich *Escherichia*- *Shigella*-, *Anaerococcus*- und *Propionibacterium*-Spezies zugeordnet. Diese sind unter dem Begriff „andere“ in Tabelle 18 zusammengefasst. Auf funktionaler Ebene erfolgten Zuordnungen der

Referenzsequenzen in der CDD-Datenbank hauptsächlich zu den unspezifischen Gruppen *MetC* (COG0626) und *MalY* (COG1168). Sehr wenige Annotationen werden weiterhin den Gruppen *CysK*, *CBS-like* und *Met\_gamma\_lyase* zugeordnet. Diese Annotationen werden nicht weiterführend betrachtet und sind nicht in Tabelle 18 dargestellt. Bei Betrachtung der in Tabelle 18 dargestellten Zuordnungen zeigt sich, dass Referenzsequenzen von Staphylokokken ausschließlich der Gruppe *MetC* und Referenzsequenzen der Corynebakterien ausschließlich der Gruppe *MalY* zugeordnet werden. Ein ähnliches Phänomen wurde bereits von James et al. beschrieben, jedoch wurden dazu keine Primärdaten publiziert (James et al., 2013). Die Einteilung der Cystathionin- $\beta$ -Lysen in die beiden Gruppen *MetC* und *MalY* basiert auf der Beschreibung zweier Enzyme mit Cystathionin- $\beta$ -Lyase-Aktivität im Methionin-Stoffwechsel von *E. coli*, wobei das *MalY*-Protein zusätzlich eine regulatorische Funktion bei der Aufnahme von Maltose besitzt (Zdych et al., 1995).

**Tabelle 18: Identifizierung von Sequenzen mit Homologie zu Sequenzen der Cystathionin- $\beta$ -Lyase in der GenBank-Datenbank. Die Sequenzen werden funktional den Familien *MetC* und *MalY* und taxonomisch den Genera *Staphylococcus* und *Corynebacterium* sowie der Gruppe „andere“ zugeordnet. Die Zuordnung erfolgt auf Grundlage der Referenzsequenzen. Dargestellt ist jeweils die Anzahl der Annotationen sowie die mittlere Sequenzidentität zur Referenzsequenz in Klammern.**

| Probe | <i>Staphylococcus</i> |      | <i>Corynebacterium</i> |             | andere      |            |
|-------|-----------------------|------|------------------------|-------------|-------------|------------|
|       | MetC                  | MalY | MetC                   | MalY        | MetC        | MalY       |
| 6n    | 187 (95,5 %)          | -    | -                      | -           | -           | -          |
| 7n    | 10 (98 %)             | -    | -                      | 21 (75,7 %) | -           | -          |
| 13n   | 352 (99 %)            | -    | -                      | -           | -           | -          |
| 37n   | 56 (97,9 %)           | -    | -                      | -           | -           | -          |
| 48v   | 40 (89,6 %)           | -    | -                      | -           | -           | -          |
| 48n   | 104 (89,4 %)          | -    | -                      | -           | -           | -          |
| 18v   | 147 (99,2 %)          | -    | -                      | -           | -           | -          |
| 18n   | 166 (98 %)            | -    | -                      | -           | -           | 13 (100 %) |
| 19v   | 31 (84,3 %)           | -    | -                      | 20 (73,3 %) | -           | 13 (100 %) |
| 19n   | 13 (85,5 %)           | -    | -                      | -           | -           | -          |
| 42v   | 260 (99,2 %)          | -    | -                      | -           | -           | -          |
| 42n   | 133 (99,6 %)          | -    | -                      | -           | -           | -          |
| 43v   | 34 (95,7 %)           | -    | -                      | 9 (77,7 %)  | 33 (76,7 %) | -          |
| 43n   | 41 (98,2 %)           | --   | -                      | 2 (80,4 %)  | 1 (86,1 %)  | -          |

Für die im Rahmen dieser Arbeit untersuchten Datensätze wurden deutlich mehr Annotationen zu den MetC-Proteinen aus Staphylokokken beschrieben als zu den MalY-Proteinen aus Corynebakterien (s. Tabelle 18). Annotationen zu den MalY-Proteinen aus Corynebakterien wurden lediglich für die Datensätze der Proben 19v, 43v und 43n des Probanden 19 und für den Datensatz 7n des Probanden 7 beschrieben. Diese Datensätze weisen im Rahmen der in Kapitel 2.3.1 beschriebenen rRNA-basierten Taxonomie einen hohen Anteil an corynebakteriellen Sequenzen auf. Die Datensätze der Proben 13n und 37n des Probanden 13 sowie 18v, 18n, 42v und 42n des Probanden 18 sind auf rRNA-Ebene von Staphylokokken-Sequenzen dominiert und weisen relativ viele Annotationen zu MetC-Proteinen von Staphylokokken auf. Bei Betrachtung der mittleren Identitäten der Sequenzen aus den Metatranskriptomik-Datensätzen zu den jeweiligen Referenzsequenzen werden Unterschiede zwischen den MetC- und MalY-Annotationen deutlich. Die Sequenzidentitäten der MetC-Annotationen liegen zwischen 84,3 % und 99,6 %, während die Sequenzidentitäten der MalY-Annotationen aus Corynebakterien mit 73,3 % - 80,4 % geringer sind. Annotationen zu MetC-Proteinen anderer Bakterienarten werden lediglich für die Datensätze der Proben 43v und 43n beschrieben. Jeweils 13 Annotationen zu MalY-Proteinen mit 100 % Sequenzidentität können für die Datensätze 18n und 19v beschrieben werden. Dabei handelt es sich um Annotationen zu MalY-Proteinen unterschiedlicher Stämme von *Escherichia coli* und somit um eine Annotation, welche wahrscheinlich ohne biologische Bedeutung für die Achselhöhlen-Mikrobiota ist, da *E. coli* nicht als Mitglied der Lebensgemeinschaften der Achselhöhle beschrieben ist.

Die Proteine der MalY-Gruppe wurden mit der Entstehung von Körpergeruch in Verbindung gebracht. Zum einen handelt es sich bei der von Natsch et al. im Zusammenhang mit der Abspaltung von Sulfanylalkanolen von Vorläufer-Molekülen beschriebenen Cystathionin- $\beta$ -Lyase aus *C. striatum* Ax20 um ein Protein der MalY-Gruppe (Natsch et al., 2004). Weiterhin konnten James et al. die gleiche Aktivität für das MalY-Protein mit der Annotation Cystathionin- $\beta$ -Lyase aus *C. jeikeium* NCIB 40928 zeigen (James et al., 2013). In der selben Studie konnte diese Aktivität für die von Starkenmann et al. beschriebene Cystathionin- $\beta$ -Lyase vom MetC-Typ aus *Staphylococcus haemolyticus* nicht demonstriert werden. Daher werden im nächsten Schritt der hier vorliegenden Arbeit alle Sequenzen der Metatranskriptomik-Datensätze mit einer Zuordnung zur Gruppe der MalY-Proteine weiterführend analysiert.

Zur Beschreibung aller MalY-Annotationen in den Metatranskriptomik-Datensätzen der Achselhöhlen-Proben wurden die Datensätze auf Grundlage des COG-Annotationssystems auf funktionaler Ebene klassifiziert und Sequenzen, die zur Gruppe COG1168 (MalY) annotiert wurden, mithilfe einer Abfrage der *GenBank*-Datenbank unter Verwendung des *BLASTX*-

Algorithmus weiterführend untersucht. Bei der taxonomischen Zuordnung der Referenzsequenzen handelt es sich im Wesentlichen um Zuordnungen zu *Staphylococcus*- und *Corynebacterium*-Spezies (s. Abbildung 29). Ein geringer Anteil der Sequenzen liefert Treffer gegen Sequenzen anderer Bakterienspezies wie z. B. *Escherichia coli* und wird im Folgenden nicht weiterführend betrachtet. Auffallend ist die besonders hohe Anzahl an Sequenzen mit Homologie zu MaLY-Sequenzen in den Datensätzen der Proben 48v (50) und 48n (86), die auf taxonomischer Ebene überwiegend Staphylokokken zugeordnet werden. Dies ist ebenfalls für die Probe 43n des Probanden 19 sowie für die Probe 6n des Probanden 6 der Fall. Zuordnungen zu corynebakteriellen MaLY-Sequenzen dominieren in den Datensätzen 19v und 43v des Probanden 19 sowie in den Datensätzen der Probe 7n des Probanden 7. Diese Datensätze weisen auf der Ebene der rRNA ebenfalls einen hohen Anteil an corynebakteriellen Sequenzen auf. Die Datensätze der Proben 13n und 37n des Probanden 13 sowie die Datensätze 18v, 18n, 42v und 42n des Probanden 18 weisen sehr wenige Sequenzen ( $\leq 4$ ) mit Homologie zu MaLY-Sequenzen auf. Diese Datensätze sind auf rRNA-Ebene stark durch Sequenzen mit Homologie zu Staphylokokken-Sequenzen dominiert. Die Read-Abundanzen der Annotation MaLY (COG1168) korrelieren leicht positiv mit dem *Malodor Score* ( $r^2 = 0,34$ ).

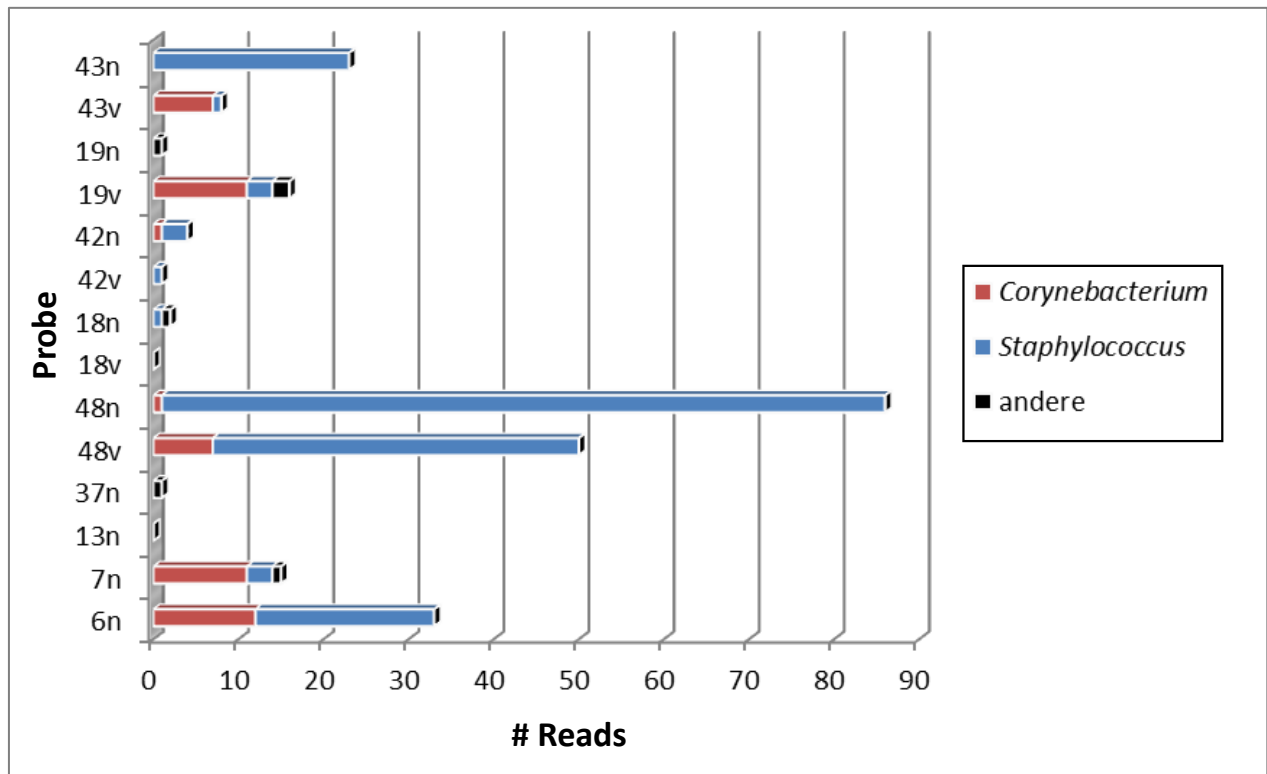


Abbildung 29: Taxonomisches Profil der Sequenzen mit Homologie zur Gruppe der MaLY-Proteine (COG1168). Dargestellt ist die Anzahl der Reads (kumulativ) mit Zuordnung zu den Genera *Corynebacterium* (rot) und *Staphylococcus* (blau) sowie zu anderen Genera (schwarz).

Tabelle 19: Referenzsequenzen der MalY-Sequenzen in der *GenBank*-Datenbank. Dargestellt sind jeweils die *GenBank-ID*, die taxonomische Annotation (Organismus) und die funktionale Annotation (Funktion) zu Referenzsequenzen aus Corynebakterien und Staphylokokken.

| GenBank-ID     | Organismus   | Funktion   |
|----------------|--|--|
| WP_023020984.1 | <i>Corynebacterium</i> (Multispezies)                      | Hypothetisches Protein                           |
| WP_018016626.1 | <i>Corynebacterium capitovis</i>                           | Hypothetisches Protein                           |
| WP_005289287.1 | <i>Corynebacterium genitalium</i>                          | Plastocyanin                                     |
| AIJ34036.1     | <i>Corynebacterium imitans</i>                             | Aminotransferase                                 |
| WP_018296070.1 | <i>Corynebacterium lubricantis</i>                         | Hypothetisches Protein                           |
| WP_018580763.1 | <i>Corynebacterium pilosum</i>                             | Hypothetisches Protein                           |
| WP_005324171.1 | <i>Corynebacterium pseudogenitalium</i>                    | Plastocyanin                                     |
| AAU06200.1     | <i>Corynebacterium striatum</i>                            | Cystathionin- $\beta$ -Lyase                     |
| WP_005326419.1 | <i>Corynebacterium tuberculostearicum</i>                  | Plastocyanin                                     |
| AIL97355.1     | <i>Corynebacterium ureicelerivorans</i>                    | Aminotransferase                                 |
| KDP47812.1     | <i>Staphylococcus aureus</i> subsp. <i>aureus</i><br>CO-98 | Cystathionin- $\beta$ -Lyase,<br>Domänen-Protein |
| KDP47826.1     | <i>Staphylococcus aureus</i> subsp. <i>aureus</i><br>CO-99 | Aminotransferase,<br>Domänen-Protein             |
| WP_011274550.1 | <i>Staphylococcus haemolyticus</i>                         | Aminotransferase                                 |
| WP_033079824.1 | <i>Staphylococcus haemolyticus</i>                         | Aminotransferase                                 |
| WP_016930549.1 | <i>Staphylococcus haemolyticus</i>                         | Aminotransferase                                 |
| WP_002460749.1 | <i>Staphylococcus haemolyticus</i>                         | Aminotransferase                                 |
| WP_017174763.1 | <i>Staphylococcus hominis</i>                              | Aminotransferase                                 |
| WP_002449246.1 | <i>Staphylococcus hominis</i>                              | Aminotransferase                                 |
| WP_002488459.1 | <i>Staphylococcus hominis</i>                              | Aminotransferase                                 |
| WP_002454642.1 | <i>Staphylococcus hominis</i>                              | Aminotransferase                                 |
| WP_002492387.1 | <i>Staphylococcus lugdunensis</i>                          | Aminotransferase                                 |
| KAK58237.1     | <i>Staphylococcus lugdunensis</i> VCU150                   | Putative C-S-Lyase                               |
| WP_019835080.1 | <i>Staphylococcus</i> sp. MDS7B                            | Aminotransferase                                 |
| WP_019835081.1 | <i>Staphylococcus</i> sp. MDS7B                            | Hypothetisches Protein                           |

In Tabelle 19 sind die identifizierten Referenzsequenzen aus Corynebakterien und Staphylokokken dargestellt. Andere Referenzsequenzen werden den Spezies *Actinomyces neuii*, *Pantholops hodgsonii*, *Escherichia coli* und *Cupridavidus taiwanensis* zugeordnet. Diese werden

nicht weiterführend betrachtet. Bei den häufigsten funktionalen Annotationen handelt es sich um Aminotransferase, hypothetisches Protein und Plastocyanin.

Innerhalb der Zuordnungen zu corynebakteriellen MalY-Sequenzen wurden zehn unterschiedliche Referenzsequenzen identifiziert. Zu den Referenzsequenzen mit den *GenBank-IDs* WP\_023020984.1 (*Corynebacterium* (Multispecies) *hypothetisches Protein*), WP\_005324171.1 (*C. pseudogenitalium*, Plastocyanin) und WP\_005326419.1 (*C. tuberculostearicum*, Plastocyanin) konnten innerhalb der Metatranskriptomik-Datensätze Sequenzen mit 100%iger Sequenzidentität identifiziert werden. Darüber hinaus wurden die Referenzsequenzen WP\_018580763.1 (*C. pilosum*, hypothetisches Protein), WP\_005289287.1 (*C. genitalium*, Plastocyanin), AAU06200.1 (*C. striatum*, Cystathionin- $\beta$ -Lyase), WP\_018296070.1 (*C. lubricantis*, hypothetisches Protein), AIL97355.1 (*C. ureicelerivorans*, Aminotransferase), AIJ34036.1 (*C. imitans*, Aminotransferase) und WP\_018016626.1 (*C. capitovis*, hypothetisches Protein) mit jeweils weniger als 100 % Sequenzidentität innerhalb der Metatranskriptomik-Datensätze identifiziert. Bei den Spezies *C. pilosum*, *C. lubricantis* und *C. capitovis* handelt es sich um Organismen, welche abweichend zu allen anderen in diesem Zusammenhang beschriebenen Corynebakterien nicht aus dem Menschen isoliert wurden (Collins et al., 2001; Kämpfer et al., 2009; Yanagawa, 1986). Innerhalb der *CDD*-Datenbank werden diese Referenzsequenzen der Proteinfamilie der AAT-ähnlichen Aspartat-Aminotransferasen (*cd0609*) zugeordnet. Dabei handelt es sich um Pyridoxalphosphat-abhängige Enzyme, die Transaminierungs-, Racemisierungs-, Decarboxylierungs- sowie diverse andere Reaktionen katalysieren (Marchler-Bauer et al., 2015). Pyridoxalphosphat bindet kovalent die Seitenkette des katalytischen Lysinrestes und bildet eine Schiff'sche Base aus. Es handelt sich um eine Gruppe von Proteinsequenzen mit hoher Sequenzähnlichkeit. Der katalytische Lysinrest befindet sich an Position 233 innerhalb einer stark konservierten Region. Aufgrund der hohen Sequenzidentitäten zwischen den einzelnen Referenzsequenzen ist anzunehmen, dass es sich um eine hoch konservierte Gruppe von Proteinen handelt. In dieser Gruppe der stark konservierten Proteine befindet sich auch die von Natsch et al. beschriebene Cystathionin- $\beta$ -Lyase, die die Abspaltung geruchsbildender Sulfanylalkanole von Cysteinkonjugaten katalysiert (Natsch et al., 2004). Die identifizierten Sequenzen innerhalb der Metatranskriptomik-Datensätze mit Homologie zu corynebakteriellen MalY-Sequenzen stellen daher potentielle Kandidaten für Transkripte mit Cystathionin- $\beta$ -Lyase-Aktivität in der Achselhöhle dar.

Innerhalb der Zuordnungen zu MalY-Referenzsequenzen von Staphylokokken wurden insgesamt 16 unterschiedliche Referenzsequenzen identifiziert. Bei zwei Referenzsequenzen (*GenBank-ID* KDP47812.1 und KDP47826.1) handelt es sich um funktionale Annotationen zu *S. aureus* subsp.

*aureus* CO-98. Die Referenzsequenzen sind im Vergleich zu den anderen im Rahmen dieser Analyse identifizierten Referenzsequenzen sehr kurz und stellen vermutlich lediglich Sequenzfragmente dar. Daher werden diese Sequenzen im Folgenden nicht weiterführend betrachtet. Bei den übrigen 14 Referenzsequenzen handelt es sich um funktionale Annotationen der vier *Staphylococcus*-Spezies *S. lugdunensis* (GenBank-IDs WP\_002460749.1, KAK58237.1, WP\_002492387.1, WP\_002479040.1, YP\_005759351.1 und YP\_003470502.1), *S. hominis* (GenBank IDs WP\_017174763.1, WP\_002488459.1, WP\_002454642.1 und WP\_002449246.1), *Staphylococcus sp.* MDS7B (GenBank-ID WP\_019835080.1) sowie *S. haemolyticus* (GenBank-IDs WP\_011274550.1, WP\_033079824.1 und WP\_016930549.1), welcher von Starkenmann et al. bereits im Zusammenhang mit C-S-Lyase-Aktivität und der Spaltung von Cystein-Konjugaten beschrieben wurde (Starkenmann et al., 2005). Bei den funktionalen Annotationen der Sequenzen handelt es sich überwiegend um Aminotransferasen, putative C-S-Lyasen sowie um ein hypothetisches Protein. Dabei wurde für nahezu jede der Referenzsequenzen mindestens eine Sequenz mit 100%iger Sequenzidentität in mindestens einem der Metatranskriptomik-Datensätze identifiziert. Lediglich für die Sequenzen von zwei Aminotransferasen aus *S. haemolyticus* (WP\_033079824.1 und WP\_011274550.1) sowie einer Aminotransferase aus *S. lugdunensis* (WP\_002460749.1) wurden keine Sequenzen mit 100%iger Sequenzidentität identifiziert.

Die Längen der Referenzsequenzen liegen zwischen 393 Aminosäuren und 399 Aminosäuren und die Sequenzen weisen eine sehr hohe Sequenzähnlichkeit über die gesamte Länge der Sequenz auf (s. *Multiple Sequence Alignment* im Anhang). Aufgrund der hohen Sequenzhomologie innerhalb der Gruppe dieser MalY-Sequenzen und aufgrund der Tatsache, dass Sequenzen mit 100%iger Sequenzidentität zu 11 dieser Sequenzen innerhalb der Metatranskriptomik-Datensätze identifiziert wurden, stellen die identifizierten Sequenzen mögliche Kandidaten für Transkripte mit C-S-Lyase-Aktivität in der Achselhöhlen-Mikrobiota dar, welche an der Entstehung von Körpergeruch beteiligt sind.

Insgesamt konnten elf Sequenzen aus Staphylokokken und drei Sequenzen aus Corynebakterien aus der Gruppe der MalY-Proteine in den Metatranskriptomik-Datensätzen identifiziert werden, die eine 100%ige Sequenzidentität zu den entsprechenden Abschnitten der Referenzsequenzen aufweisen und an der Geruchsentstehung durch Abspaltung von körpergeruchsverursachenden Sulfanylalkanolen von Vorläufer-Molekülen beteiligt sein könnten.



### 2.4.5 Analyse von Dipeptidase-Transkripten

Für die Identifizierung von Sequenzen mit Homologie zur von Natsch et al. beschriebenen Sequenz der TpdA-Dipeptidase wurde ein Vorgehen gewählt, welches ähnlich zur Analyse der Cystathionin- $\beta$ -Lyase ist (s. oben). Dabei wurden die Metatranskriptomik-Datensätze mithilfe des *GenBank*-Annotationssystems in MG-RAST nach der Annotation *Thiol Precursor Dipeptidase* durchsucht. Die identifizierten Annotationen wurden anschließend mithilfe der *Best Hit Classification* auf Spezies-Ebene klassifiziert (s. Abbildung 30). Die Sequenzen werden ausschließlich den Spezies *Corynebacterium lipophiloflavum*, *Corynebacterium striatum* und *Corynebacterium accolens* zugeordnet. Dabei wurden keine Sequenzhomologien in den Datensätzen der Proben 13n und 37n des Probanden 13 sowie 18v, 18n und 42v des Probanden 18 gefunden. Diese Proben werden in der rRNA-basierten Taxonomie von Staphylokokken-Sequenzen dominiert. Daher steht die Abwesenheit von Sequenzen mit Homologie zu corynebakteriellen Sequenzen auf funktionaler Ebene mit dieser Beobachtung im Einklang. Die höchste Anzahl an Sequenzen mit Homologie zu corynebakteriellen TpdA-Sequenzen wurde im Datensatz der Probe 7n des Probanden 7 identifiziert, die auf taxonomischer Ebene ebenfalls stark von Sequenzen mit Homologie zu corynebakteriellen Sequenzen dominiert wird. In diesem

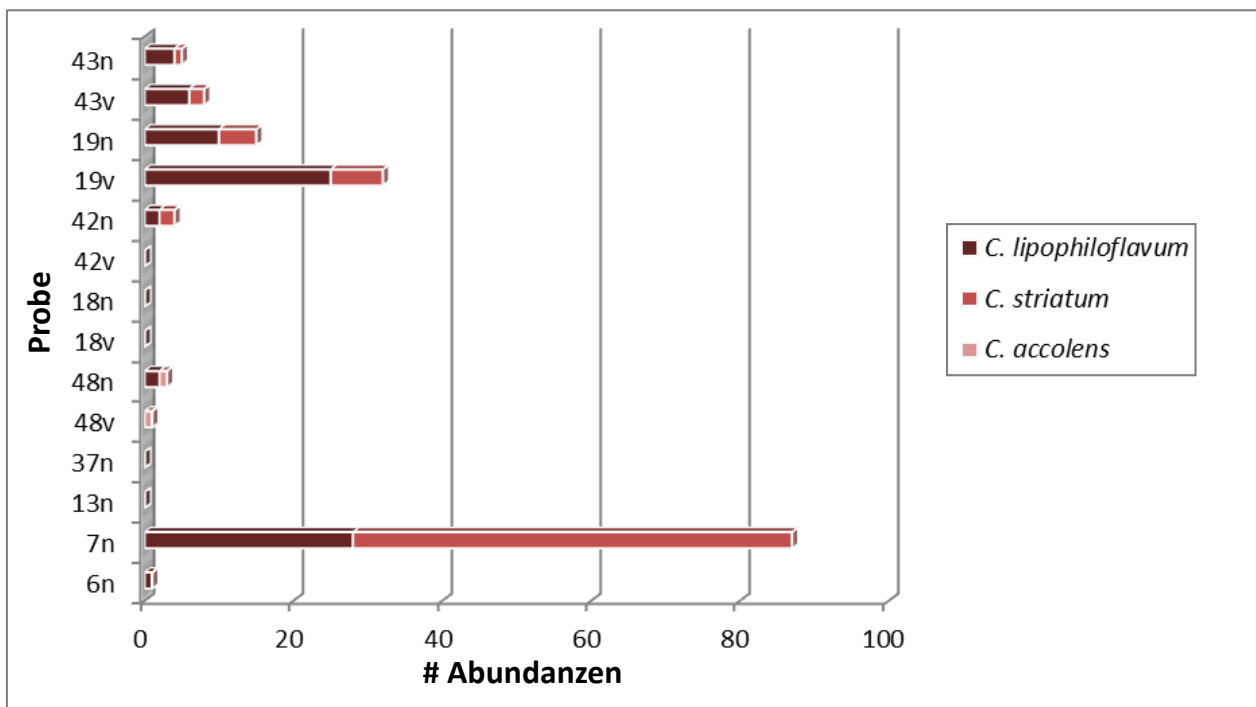


Abbildung 30: Funktionales taxonomisches Profil der Annotationen zur *Thiol Precursor Dipeptidase* (TpdA) auf Spezies-Ebene. Dargestellt ist die Anzahl der Annotationen für jede untersuchte Probe zu den Spezies *Corynebacterium lipophiloflavum*, *Corynebacterium striatum* und *Corynebacterium accolens* mittels *Best Hit Classification*.

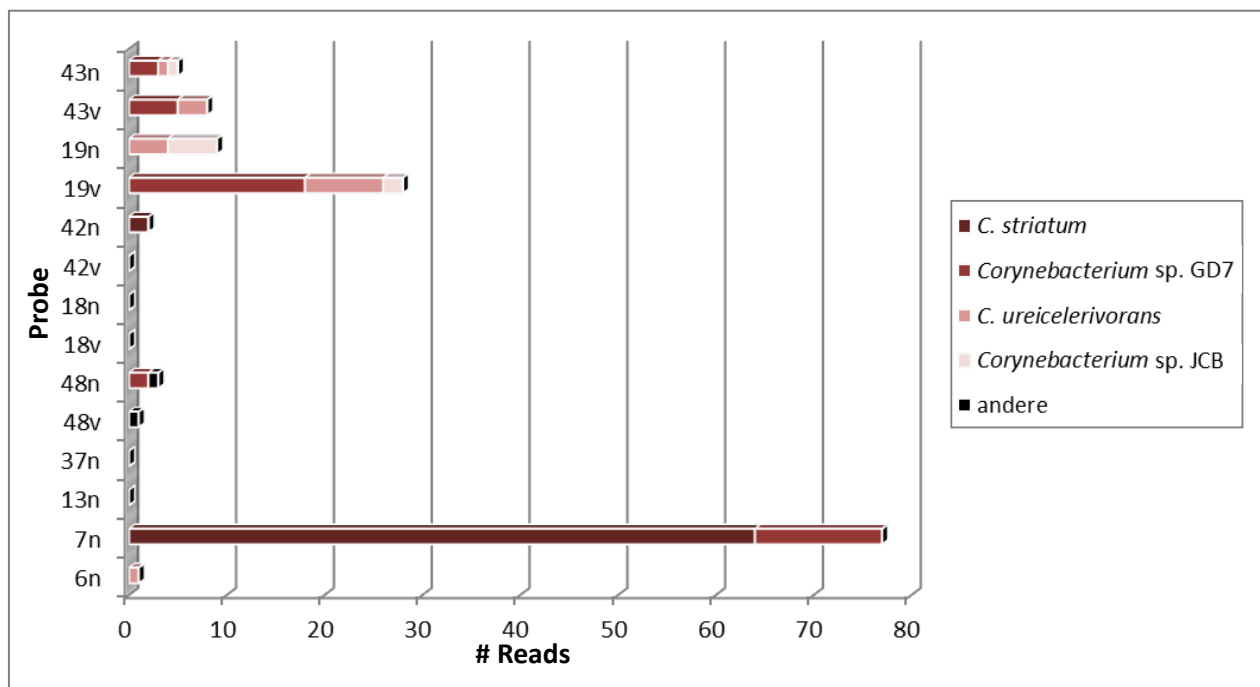
Datensatz dominieren abweichend zu den anderen Datensätzen Sequenzhomologien zu *C. striatum*.

Die auf diese Weise identifizierten Sequenzen wurden anschließend in der Online-Version der *GenBank*-Datenbank abgefragt. Bei dieser Analyse wurden in der Online-Abfrage nicht die gleichen Referenzsequenzen getroffen wie in den *GenBank*-Annotationen in MG-RAST. Die Sequenzidentitäten zu den Referenzsequenzen waren sehr variabel und lagen im Bereich von 64 % - 99 %. Es wurde keine Sequenz mit einer 100%igen Identität zu einer Referenzsequenz identifiziert. Insgesamt wurden im Rahmen dieser Analyse acht corynebakterielle Referenzsequenzen identifiziert (s. Tabelle 20). Bei fünf dieser Sequenzen handelt es sich um Sequenzen mit der funktionalen Annotation „*Peptidase M20*“ und den *GenBank*-IDs WP\_026157213.1 (*C. capitovis*), WP\_035000114.1 (*Corynebacterium* sp. GD7), WP\_034662125.1 (*C. tuscaniense*), WP\_038613045.1 (*C. ureicelerivorans*), WP\_042408999.1 (*Corynebacterium* sp. JCB), WP\_005329875.1 (*C. tuberculostearicum*) und WP\_035108558.1 (*C. accolens*). Die Sequenz mit der *GenBank*-ID WP\_015652155.1 (*C. callunae*) ist als hypothetisches Protein annotiert und bei der Referenzsequenz mit der *GenBank*-ID ACA03770.1 handelt es sich um die Sequenz der von Natsch et al. beschriebenen TpdA-Dipeptidase. In der *CCD*-Datenbank werden alle Sequenzen der Gruppe der M20\_Dipept\_like (cd03893) zugeordnet. Diese Gruppe

**Tabelle 20: Referenzsequenzen der in MG-RAST identifizierten Sequenzen mit Homologie zur TpdA Dipeptidase. Dargestellt sind jeweils die *GenBank*-ID, die taxonomische Annotation (Organismus), die funktionale Annotation (Funktion) sowie die Zuordnung in der *CDD*-Datenbank (*CDD*-Hit).**

| <b>GenBank-ID</b> | <b>Organismus</b>  | <b>Funktion</b>        | <b>CDD-Hit</b>  |
|-------------------|--|------------------------|-----------------|
| ACA03770.1        | <i>C. striatum</i>   | TpdA-Dipeptidase       | M20_Dipept_like |
| WP_015652155.1    | <i>C. callunae</i>   | Hypothetisches Protein | M20_Dipept_like |
| WP_034662125.1    | <i>C. tuscaniense</i>                                      | Peptidase M20          | M20_Dipept_like |
| WP_026157213.1    | <i>C. capitovis</i>  | Peptidase M20          | M20_Dipept_like |
| WP_038613045.1    | <i>C. ureicelerivorans</i>                                 | Peptidase M20          | M20_Dipept_like |
| WP_042408999.1    | <i>Corynebacterium</i> sp. JCB<br>( <i>C. jeddahense</i> ) | Peptidase M20          | M20_Dipept_like |
| WP_035108558.1    | <i>C. accolens</i>   | Peptidase M20          | M20_Dipept_like |
| WP_005329875.1    | <i>C. tuberculostearicum</i>                               | Peptidase M20          | M20_Dipept_like |
| WP_035000114.1    | <i>Corynebacterium</i> sp. GD7<br>( <i>C. ihumii</i> )     | Peptidase M20          | M20_Dipept_like |

umfasst viele unterschiedliche Dipeptidase-Funktionen und es sind keine für M20-Peptidasen typischen Sequenzmotive erkennbar. Bei Betrachtung der Referenzsequenzen zeigt sich, dass viele Bereiche innerhalb der Proteinsequenzen stark konserviert sind (s. *Multiple Sequence Alignment* im Anhang). Insbesondere der Sequenzabschnitt mit den Aminosäureresten RWYgRGAADCKGN mit dem katalytischen Aspartatrest ist hoch konserviert. Bei den im Rahmen dieser Analyse identifizierten Sequenzen handelt es sich demnach um Sequenzen mit Homologie zu corynebakteriellen Sequenzen der Gruppe M20\_Dipept\_like. Der taxonomische Ursprung der Sequenzen in den Metatranskriptomik-Datensätzen kann jedoch aufgrund des Fehlens 100%iger Sequenzidentitäten nicht exakt auf auf Spezies-Ebene beschrieben werden. Es ist anzunehmen, dass es sich bei dem taxonomischen Ursprung dieser Sequenzen um nicht kultivierte Organismen handelt.



**Abbildung 31:** Funktionales taxonomisches Profil der Referenzsequenzen mit Annotation zur TpdA Dipeptidase in der Online-Version der *GenBank*-Datenbank. Dargestellt sind die Spezies *Corynebacterium striatum*, *Corynebacterium sp. GD7* (*C. ihumii*), *Corynebacterium ureicelerivorans*, *Corynebacterium sp. JCB* (*C. jeddahense*) sowie andere Spezies.

Die Häufigkeitsverteilung von Sequenzen mit Homologie zu den beschriebenen Referenzsequenzen in den Metatranskriptomik-Datensätzen ist in Abbildung 31 dargestellt. Auch bei dieser Betrachtung werden im Datensatz der Probe 7n die meisten Sequenzen mit Homologie zur TpdA Dipeptidase beschrieben. Die meisten Sequenzen in diesem Datensatz werden *C. striatum* zugeordnet. Die Datensätze der Proben 19v, 19n, 43v und 43n werden im

Wesentlichen durch Sequenzen mit Homologie zu *Corynebacterium* sp. GD7 (*C. ihumii*), *Corynebacterium ureicelerivorans* und *Corynebacterium* sp. JCB (*C. jeddahense*) beschrieben (Edouard et al., 2014; Padmanabhan et al., 2014).

Zur Identifizierung weiterer Sequenzen mit Homologie zu corynebakteriellen Sequenzen der Gruppe M20\_Dipept\_like wurde weiterhin Sequenzen mit Homologie zu der entsprechenden COG-Gruppe COG2195 mithilfe einer Abfrage der *GenBank*-Datenbank durchgeführt. Im Rahmen dieser Analyse wurden keine Sequenzen mit Homologie zu corynebakteriellen Sequenzen gefunden. Die meisten der identifizierten Sequenzen weisen Homologie zu Sequenzen der Peptidase T aus Staphylokokken auf. Bei der Peptidase T handelt es sich um eine Tripeptidase der Gruppe M20\_pepT (cd03892 in der *CDD*-Datenbank), die N-terminale Aminosäuren von Tripeptiden spaltet. Die physiologische Rolle sowie die Rolle dieser Proteine im Zusammenhang mit Körpergeruchsentstehung ist nicht bekannt. Weiterhin werden u. a. Sequenzhomologien zu Peptidase-T-Proteinen von Anaerokokken sowie Sequenzen mit der Annotation Aminoacyl-Histidin-Dipeptidase aus *Fingoldia magna* beschrieben. Die Rolle dieser Sequenzen im Zusammenhang mit Körpergeruchsentstehung ist ebenfalls nicht beschrieben.

#### **2.4.6 Analyse anderer im Zusammenhang mit Körpergeruchsentstehung beschriebener Funktionen**

Für die Analyse anderer in der Fachliteratur zur Körpergeruchsentstehung beschriebener Funktionen wurde das *Subsystems*-Annotationssystem in MG-RAST verwendet. Dabei wurden die entsprechenden Funktionen entweder auf Grundlage der Annotation oder der EC-Nomenklatur identifiziert und anschließend mittels der *Best Hit Classification* funktional auf Genus-Ebene charakterisiert.

Die Suche nach Annotationen zu Enzymen mit Beteiligung an der Entstehung geruchsbildender Steroide ergab keine Treffer für die Steroid-5- $\alpha$ -Reduktase (EC 1.3.99.5), 5- $\alpha$ -( $\beta$ )-Sterol Dehydrogenase (EC 1.1.1.-) und Steroid-4,5-Isomerase (EC 5.3.3.1). Insgesamt konnten lediglich 16 Annotationen zu Sequenzen der  $\beta$ -Glucuronidase (EC 3.2.1.31) in sechs Datensätzen identifiziert werden. Mithilfe der *Best Hit Classification* wurden diese Sequenzen den Genera *Clostridium*, *Propionibacterium*, *Staphylococcus*, *Shigella* und *Congregibacter* zugeordnet. Aufgrund der relativ geringen Anzahl an Sequenzen, die zur  $\beta$ -Glucuronidase annotiert wurden, ist anzunehmen, dass diese Enzymfunktion weder bei der Funktion der Achselhöhlen-Mikrobiota noch im Zusammenhang mit Körpergeruchsentstehung eine bedeutende Rolle spielt.

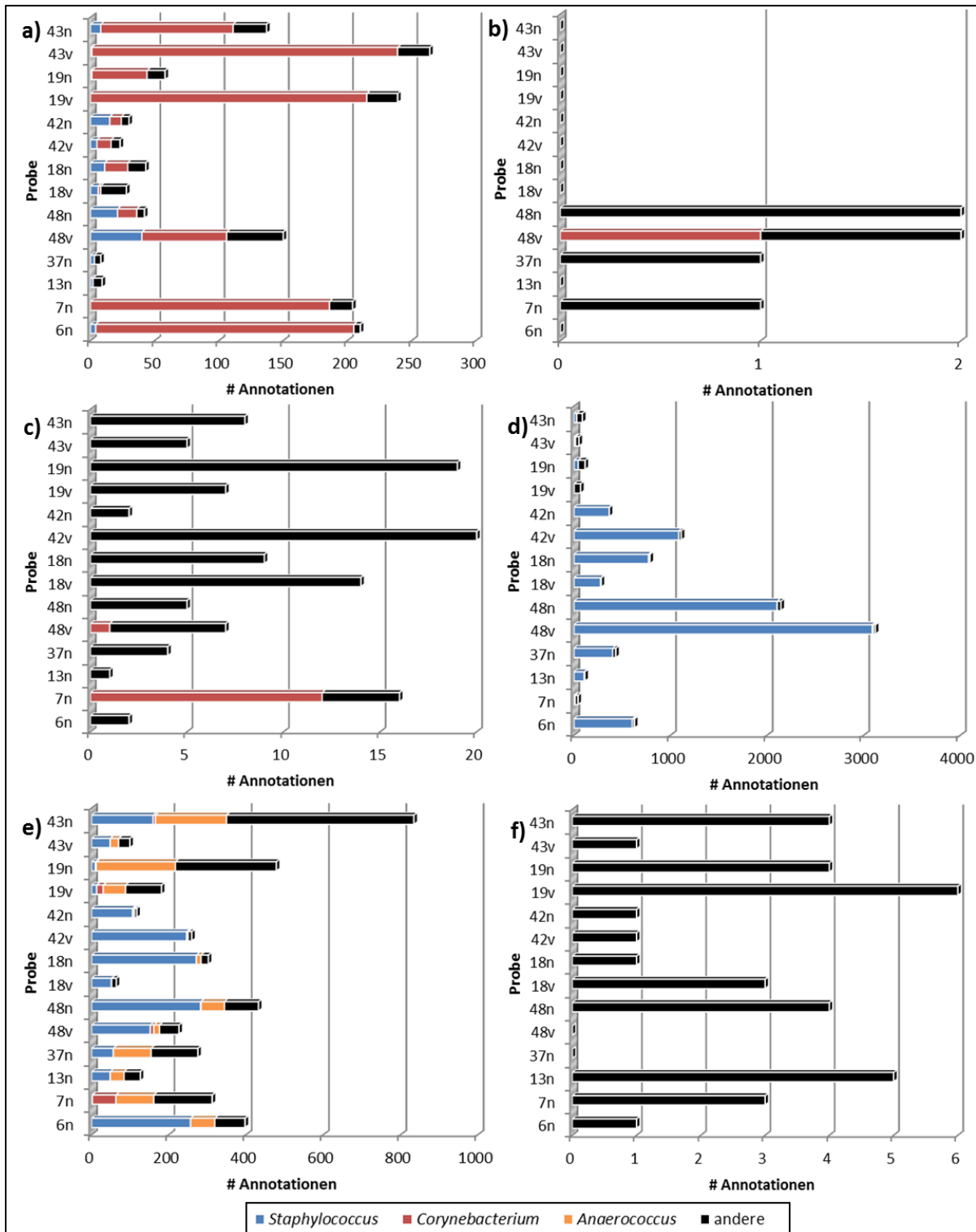
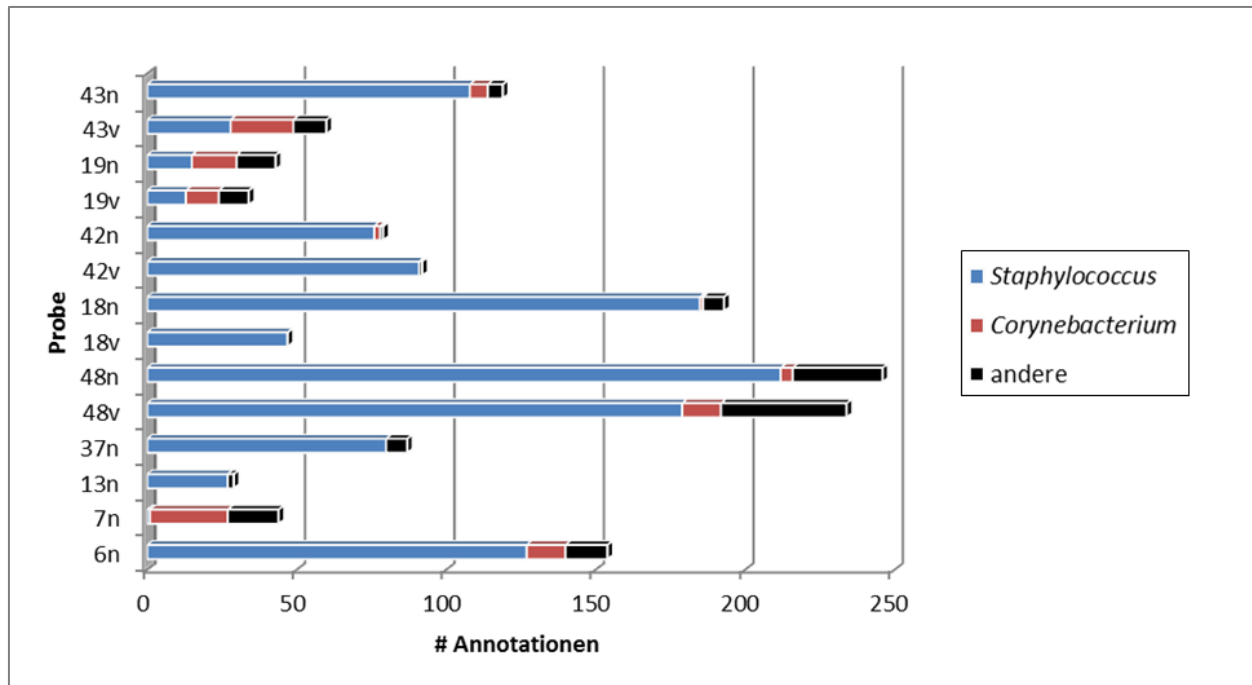


Abbildung 32: Funktionales taxonomisches Profil der Annotationen mit Beteiligung am Fettsäureabbau auf Genus-Ebene mithilfe der *Best Hit Classification* in MG-RAST; a): Acyl-CoA-Synthetase (EC 6.2.1.3), b): Acyl-CoA-Dehydrogenase (EC 1.3.99.3), c): Enoyl-CoA-Hydratase (EC 4.2.1.17), d): Hydroxyacyl-CoA-Dehydrogenase (EC 1.1.1.35), e): Ketoacyl-CoA-Thiolase (EC 2.3.1.16) und 2,4-Dienoyl-CoA-Reduktase (EC 1.3.1.34). Dargestellt sind die Annotationen zu den Genera *Staphylococcus* (blau), *Corynebacterium* (rot) und *Anaerococcus* (orange) sowie zu der Gruppe „andere“ (schwarz).

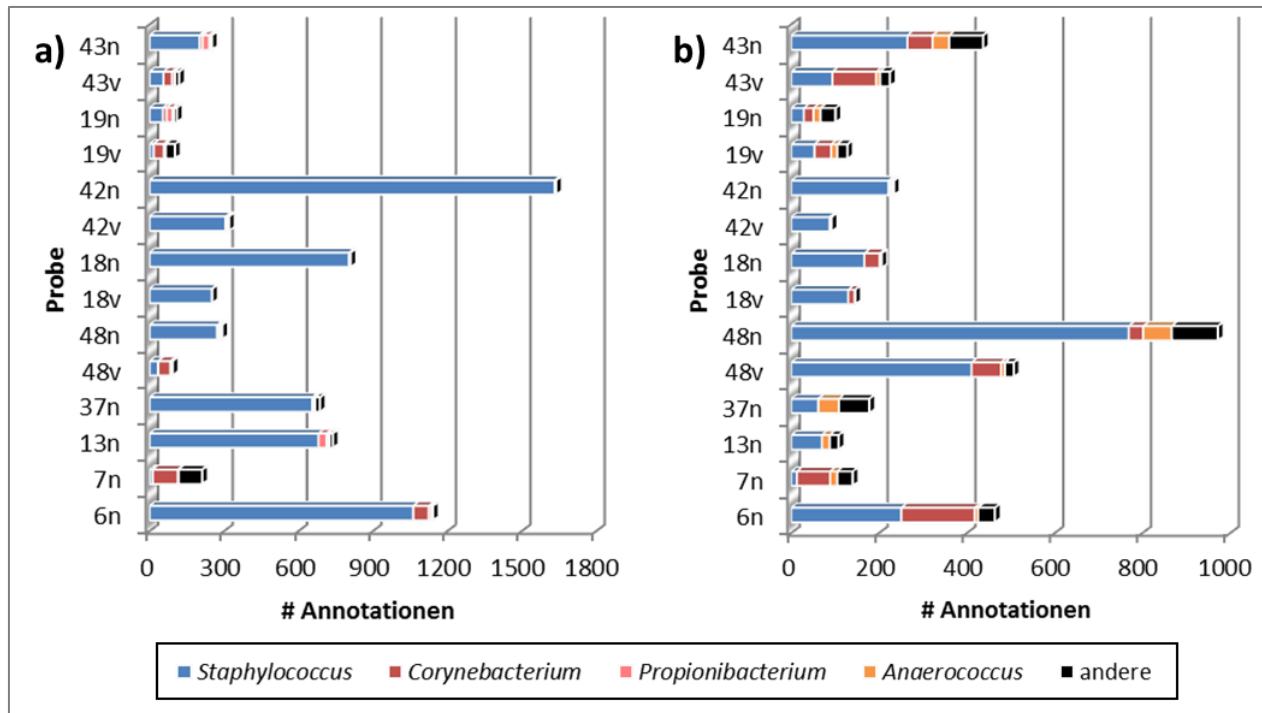
Weiterhin wurden Funktionen mit möglicher Beteiligung an der Entstehung geruchsbildender Fettsäuren untersucht. Dazu wurden zunächst Annotationen zu den am Fettsäureabbau beteiligten Enzymfunktionen Acyl-CoA-Synthetase (EC 6.2.1.3), Acyl-CoA Dehydrogenase (EC 1.3.99.3), Enoyl-CoA-Hydratase (EC 4.2.1.17), Hydroxyacyl-CoA-Dehydrogenase (EC 1.1.1.35), Ketoacyl-CoA-Thiolase (EC 2.3.1.16) und 2,4-Dienoyl-CoA-Reduktase (EC 1.3.1.34) identifiziert und taxonomisch klassifiziert. Die funktionale Taxonomie der Annotationen zu *fad*-Genen sind in Abbildung 32 auf Genus-Ebene dargestellt. Sequenzen mit Homologie zu Acyl-CoA-Synthetase-Genen werden überwiegend den Corynebakterien zugeordnet (s. Abbildung 32 a). Dabei werden besonders viele Annotationen (150 – 264) in den Datensätzen der Proben 6n, 7n und 48n sowie 19v, 43v und 43n identifiziert. Diese Datensätze weisen auf der Ebene der rRNA-basierten Taxonomie ebenfalls viele Annotation zu Corynebakterien auf. Die Datensätze der Proben 13n und 37n des Probanden 13 und 18v, 18n, 42v und 42n des Probanden 18, welche auf rRNA-Ebene von Annotationen zu Staphylokokken dominiert werden, weisen nur sehr wenige Annotationen (8 – 43) zu Acyl-CoA-Synthetase-Genen auf. Die Anzahl der Annotationen zur Acyl-CoA-Dehydrogenase (EC 1.3.99.3) sind mit ein (7n und 37n) bis zwei (48v und 48n) Annotationen pro Datensatz sehr gering (s. Abbildung 32 b). Mit Ausnahme einer Annotation zum Genus *Corynebacterium* im Datensatz der Probe 48v werden diese Sequenzen der Gruppe „andere“ zugeordnet. Die Anzahl der Annotationen zur Enoyl-CoA-Hydratase (EC 4.2.1.17) liegt zwischen 1 (13n) und 20 (42v) pro Datensatz und wird auf taxonomischer Ebene ebenfalls der Gruppe „andere“ zugeordnet (s. Abbildung 32 c). Auffällig ist die relativ hohe Anzahl der Annotationen zur Gattung *Corynebacterium* im Datensatz der Probe 7n, welcher auf rRNA-Ebene ebenfalls von Annotationen zu Corynebakterien dominiert wird. Zuordnungen zu Genen der Hydroxyacyl-CoA-Dehydrogenase (EC 1.1.1.35) sind mit 43 (7n) bis 3120 (48v) Annotationen vor allem in den Datensätzen 48v, 48n, 18v, 18n, 42v, 42n, 37n und 6n sehr häufig (s. Abbildung 32 d). Diese Datensätze werden auf rRNA-Ebene von Zuordnungen zu Staphylokokken dominiert. Der überwiegende Anteil der Annotationen wird Staphylokokken zugeordnet. Nur relativ wenige Annotationen werden zu Corynebakterien und anderen Genera gemacht. Annotationen zur Ketoacyl-CoA-Thiolase (EC 2.3.1.16) werden den Genera *Staphylococcus*, *Anaerococcus*, *Corynebacterium* und anderen Bakteriengattungen zugeordnet (s. Abbildung 32 e). Annotationen zur 2,4-Dienoyl-CoA-Reduktase (EC 1.3.1.34) sind ebenfalls verhältnismäßig selten und werden der Gruppe der anderen Bakteriengattungen zugeordnet (s. Abbildung 32 f). Die überwiegende Mehrheit dieser Bakteriengattungen ist bisher nicht in der Literatur im Zusammenhang mit der Achselhöhlen-Mikrobiota und Körpergeruchsentstehung beschrieben worden.



**Abbildung 33:** Funktionales taxonomisches Profil der Annotationen zu BCAA-Aminotransferasen (EC 2.6.1.42) auf Genus-Ebene mithilfe der *Best-Hit-Classification* in MG-RAST. Dargestellt sind die Annotationen zu den Genera *Staphylococcus* (blau), *Corynebacterium* (rot) sowie zu der Gruppe „andere“ (schwarz).

Die taxonomische Zuordnung von Annotationen zu BCAA-Aminotransferasen auf Genus-Ebene ist in Abbildung 33 dargestellt. Insgesamt können mindestens 29 (13n) und maximal 246 (48n) Annotationen pro Datensatz beschrieben werden. Die überwiegende Mehrheit dieser Annotationen wird dabei den Staphylokokken zugeordnet. In den Datensätzen der Proben 6n, 7n, 48v, 19v, 19n, 43v und 43n, welche auf rRNA-Ebene ebenfalls einen dominanten Corynebakterien-Anteil aufweisen, können auch Zuordnungen zu Corynebakterien beschrieben werden. Im Datensatz der Proben 7n des Probanden 7 dominieren abweichend zu allen anderen Profilen Zuordnungen zu Corynebakterien.

Die taxonomischen Profile der Zuordnungen zur L-Lactat-Dehydrogenase und Acetat-Kinase sind in Abbildung 34 dargestellt. Insgesamt werden bis zu 1639 *Read*-Abundanzen (Datensatz 42n, Lactat-Dehydrogenase) in einer Probe beschrieben. Der überwiegende Anteil der Annotationen zu beiden Funktionen wird den Staphylokokken zugeordnet. Relativ wenige Annotationen werden den Genera *Corynebacterium* und *Anaerococcus* zugeordnet. Im Falle der Laktat-Dehydrogenase finden sich darüber hinaus noch Zuordnungen zu Propionibakterien (s. Abbildung 34 b). In beiden Fällen stellt *Corynebacterium* im Datensatz der Probe 7 die dominanteste Bakteriengattung dieser Funktionen dar.



**Abbildung 34: Funktionales taxonomisches Profil der Annotationen zu fermentativen Stoffwechselwegen auf Genus-Ebene mithilfe der Best Hit Classification in MG-RAST; a): L-Lactat-Dehydrogenase (EC 1.1.1.27); b): Acetat-Kinase (EC 2.7.2.1).** Dargestellt sind die Annotationen zu den Genera *Staphylococcus* (blau), *Corynebacterium* (rot), *Propionibacterium* (rosa) und *Anaerococcus* (orange) sowie zu der Gruppe „andere“ (schwarz).

#### 2.4.7 Beschreibung von Funktionen mit Korrelation zum Körpergeruch

Auf Grundlage des KO-Annotationssystems und der entsprechend zugeordneten *Read*-Abundanzen wurde die Korrelation einzelner Funktionen mit dem *Malodor Score* ermittelt. Auf diese Weise konnten Korrelationskoeffizienten für insgesamt 2359 Funktionen beschrieben werden. Genau 263 (11 %) Funktionen wiesen dabei einen Korrelationskoeffizienten  $\geq 0,5$  auf und wurden als positiv korreliert gewertet. Weitere 270 (11 %) Funktionen wiesen einen Korrelationskoeffizienten  $\leq -0,5$  auf und wurden als negativ korreliert gewertet. Die übrigen 1826 (77 %) Funktionen wiesen einen Korrelationskoeffizienten zwischen 0,5 und -0,5 auf. Diese wurden als nicht mit dem *Malodor Score* korreliert interpretiert.

Die jeweils 10 am stärksten positiv bzw. negativ korrelierten Funktionen sind in Tabelle 21 dargestellt. Auffallend ist, dass die positiv korrelierten Funktionen auf taxonomischer Ebene überwiegend den Corynebakterien, während die negativ korrelierten nahezu ausschließlich den Staphylokokken zugeordnet werden. Mit einem Korrelationskoeffizienten von 0,88 stellt die Hämoxxygenase die am stärksten positiv korrelierte aller auf diese Weise beschriebenen Funktionen dar. Auf taxonomischer Ebene wird diese Annotation überwiegend den



Corynebakterien und im Falle der Probe 6n auch den Staphylokokken zugeordnet. Die Hämoxygenase spielt insbesondere bei prokaryotischen Pathogenen eine bedeutende Rolle bei der Eisenaufnahme (Frankenberg-Dinkel, 2004). Darüber hinaus weisen die zwei Annotationen aus der Tryptophan-Biosynthese TrpCF (Indol-3-Glycerinphosphat-Synthase/Phosphoribosylantranilat-Isomerase) und TrpD (Antranilat-Phosphoribosyltransferase) mit 0,86 und 0,79 sehr hohe positive Korrelationskoeffizienten auf. Die Enoyl-CoA-Hydratase,

**Tabelle 21: Funktionen mit den stärksten positiven bzw. negativen Korrelationen zum Körpergeruch und die dominanten Genera innerhalb der jeweiligen funktionalen taxonomischen Profile;  $r^2$ : Korrelationskoeffizient; graue Schrift: Funktionen mit sehr geringen maximalen *Read*-Abundanzen (< 10).**

| Funktion   | $r^2$ | Dominante Genera  |
|--|-------|---|
| HmuO: Hämoxygenase (EC:1.14.99.3)  | 0,88  | <i>Corynebacterium</i> , ( <i>Staphylococcus</i> )                            |
| 2-Enoat-Reduktase (EC:1.3.1.31)  | 0,87  | <i>Propionibacterium</i> , <i>Clostridium</i>                                 |
| Sialidase 1 (EC:3.2.1.18)  | 0,86  | <i>Corynebacterium</i> , <i>Propionibacterium</i> ,<br><i>Bifidobacterium</i> |
| Rubredoxin- NAD <sup>+</sup> Reduktase (EC:1.18.1.1)   | 0,86  | <i>Pseudomonas</i>  |
| TrpCF: Indol-3-Glycerinphosphat-Synthase/Phosphoribosylantranilat Isomerase (EC:4.1.1.48/5.3.1.24)       | 0,86  | <i>Corynebacterium</i>  |
| GabT: 4-Aminobutyrat Aminotransferase/(S)-3-amino-2-methylpropionat Transaminase (EC:2.6.1.19/ 2.6.1.22) | 0,82  | <i>Corynebacterium</i> , <i>Mycobacterium</i>                                 |
| Endonklease VIII (EC:3.2.2.- / 4.2.99.18)  | 0,82  | <i>Corynebacterium</i>  |
| GlcE: Glycolat-Oxidase (FAD-bindende Untereinheit)   | 0,81  | <i>Ralstonia</i> , <i>Escherichia</i> ,<br><i>Pseudomonas</i> , unbekannt     |
| PaaF, EchA: Enoyl-CoA Hydratase (EC:4.2.1.17)  | 0,80  | <i>Corynebacterium</i>  |
| TrpD: Anthranilat-Phosphoribosyltransferase (EC:2.4.2.18)  | 0,79  | <i>Corynebacterium</i>  |
| GlpA, GlpD: Glycerin-3-Phosphat Dehydrogenase (EC:1.1.5.3)   | -0,75 | <i>Staphylococcus</i> , ( <i>Corynebacterium</i> )                            |
| TagH: Teichonsäure-Transportsystem (ATP-bindendes Protein) (EC:3.6.3.40)                                 | 0,75  | <i>Staphylococcus</i>   |
| AgrA: Zwei-Komponentensystem, AgrA-Familie, <i>Response Regulator AgrA</i>                               | 0,77  | <i>Staphylococcus</i>   |
| K08303: putative Protease (EC:3.4.-.-)   | 0,77  | <i>Staphylococcus</i>   |
| SgtA: Transglycosylase (EC:2.4.1.-)  | 0,77  | <i>Staphylococcus</i>   |
| HrtA: Hämin-Transportsystem (ATP-bindendes Protein) (EC:3.6.3.-)   | 0,77  | <i>Staphylococcus</i>   |
| lpdC: Indolpyruvat Decarboxylase (EC:4.1.1.74)   | 0,78  | <i>Staphylococcus</i>   |
| RuvA: Holliday Junction DNA Helicase (EC:3.6.4.12)   | 0,78  | <i>Staphylococcus</i>   |
| Ribonuclease Z (EC:3.1.26.11)  | 0,78  | <i>Staphylococcus</i>   |
| LacG: 6-Phospho- $\beta$ -Galactosidase (EC:3.2.1.85)  | 0,85  | <i>Staphylococcus</i>   |

welche an der  $\beta$ -Oxidation und in vielen anderen katabolen Stoffwechselwegen durch die Hydratisierung der Doppelbindung zwischen C2 und C3 in Fettsäuren an der Generierung von 3-Hydroxysäuren beteiligt ist, korreliert ebenfalls stark positiv mit dem *Malodor Score* ( $r^2 = 0,80$ ) und wird den Corynebakterien zugeordnet. Aus den Arbeiten von Natsch et al. sind diverse 3-Hydroxysäuren aus frischem Schweiß bekannt, welche nach Spaltung durch N-AGA und Freisetzung mit starkem Körpergeruch assoziiert sind. Darunter befinden sich die besonders häufig beschriebenen Komponenten 3-Hydroxy-3-Methylhexansäure (HMHA), 3-Hydroxy-4-Methylhexansäure sowie die entsprechenden Heptan-, Octan-, Nonan- und Decansäuren (Natsch et al., 2006). Weiterhin korreliert die Sialidase 1 stark positiv mit dem Körpergeruch. Diese Funktion wird auf taxonomischer Ebene den Genera *Corynebacterium*, *Propionibacterium* und *Bifidobacterium* zugeordnet. Es handelt sich um Enzyme, die terminale Sialinsäuren komplexer Kohlenhydrat-Strukturen von Glycoproteinen und Glycolipiden abspalten. Diese werden dann von pathogenen oder kommensalen Bakterien als Nahrungsquelle genutzt. Dieses Phänomen ist auf physiologischer Ebene besonders gut für die Mikrobiota des Darmes beschrieben (Juge et al., 2016). Weitere Funktionen mit hohen positiven Korrelationskoeffizienten sind 2-Enoat-Reduktase, Rubredoxin-NAD<sup>+</sup>-Reduktase, 4-Aminobutyrat-Aminotransferase/(S)-3-Amino-2-methylpropionat-Transaminase, Endonuklease VIII und die FAD-bindende Untereinheit der Glykolat-Oxidase. Die entsprechenden Annotationen besitzen jedoch sehr geringe absolute *Read*-Abundanzen (< 10) und sind in nur wenigen der untersuchten Datensätze vorhanden. Daher werden diese Funktionen für die weitere Beurteilung von mit Körpergeruch im Zusammenhang stehender Funktionen nicht weiterführend betrachtet.

Bei den am stärksten mit Körpergeruch negativ korrelierten Funktionen aus Staphylokokken handelt es sich um diverse Enzymfunktionen aus unterschiedlichen physiologischen Kontexten. Mit einem Korrelationskoeffizienten von -0,85 weist die 6-Phospho- $\beta$ -Galactosidase (LacG, EC 3.2.1.85) die stärkste negative Korrelation mit dem *Malodor Score* auf. Das Enzym katalysiert im Allgemeinen die hydrolytische Spaltung von 6-Phospho- $\beta$ -D-Galactosiden. Allerdings wurde für das entsprechende Enzym aus *S. aureus* auch die Spaltung aromatischer  $\beta$ -Galactosid- und  $\beta$ -Glucosid-Phosphate gezeigt (Witt et al., 1993). Darüber hinaus stellen die RNaseZ ( $r^2 = -0,78$ ) und die putative Protease (K08303;  $r^2 = -0,77$ ) sehr allgemeine Polymer-abbauende Funktionen dar. Die DNA Helicase RuvA ( $r^2 = -0,78$ ; EC 3.6.4.12) ist an der Strukturmodifikation von DNA beteiligt und nimmt damit eine Rolle in der Regulierung der DNA-Replikation und Transkription ein. Die Indolpyruvat Decarboxylase (EC 4.1.1.74) ist am Tryptophan-Stoffwechsel beteiligt. Die SgtA-Transglycosylase ( $r^2 = -0,77$ ; EC 2.4.1.-) und das Teichonsäure-Transportsystem TagH ( $r^2 = -0,75$ ; EC 3.6.3.40) spielen eine wichtige Rolle bei der Zellwand-Biosynthese. Teichonsäuren sind

polymere Zellwandbestandteile Gram-negativer Bakterien und spielen eine wichtige Rolle bei der Zellteilung und bei der Pathogenese (Brown et al., 2013). Der Response-Regulator AgrA ( $r^2=0,77$ ) wurde im Zusammenhang mit der *Quorum-Sensing*-vermittelten Virulenz von *S. aureus* beschrieben (Singh und Ray, 2014). Diese Funktion könnte einen Hinweis auf intraspezies-Interaktionen innerhalb der Staphylokokken und auf Mechanismen der Besiedlung des Wirtsorganismus darstellen. Schließlich ist die Glycerinaldehyd-3-Dehydrogenase ( $r^2=-0,75$ ; EC 1.1.5.3) am Energiestoffwechsel beteiligt. Diese wird auf taxonomischer Ebene neben den Staphylokokken auch den Corynebakterien zugeordnet. Das Hämin-Transportsystem HrtA korreliert zwar stark negativ mit dem *Malodor Score* ( $r^2= -0,77$ ), jedoch sind die absoluten *Read*-Abundanzen dieser Annotation sehr gering, sodass sie bei der weiterführenden Bewertung vernachlässigt werden kann.

Zur graphischen Darstellung der korrelierten Funktionen innerhalb des metabolischen Kontextes wurde das KEGG-MAPPER SEARCH&COLOR PATHWAY-*Tool* verwendet. Von den 2359 eingegebenen KO-Nummern konnten 75 nicht gefunden werden. Alle übrigen Funktionen wurden gefunden und insgesamt 357 Stoffwechselwegen zugeordnet und dargestellt. Positiv korrelierte Funktionen wurden grün, negativ korrelierte rot und nicht korrelierte gelb oder schwarz dargestellt. Mit insgesamt 1044 KO-Nummern konnten die meisten Funktionen der Karte „*Metabolic Pathways*“, also dem gesamten Metabolismus zugeordnet werden (s. Anhang). Die einzelnen Stoffwechselwege wurden zunächst anhand der Bezeichnung durchsucht. Stoffwechselwege, die im Zusammenhang mit der Bildung von Geruchskomponenten stehen könnten, wurden entsprechend geladen und näher betrachtet. Im Folgenden werden drei dieser Stoffwechselwege und die taxonomische Zuordnung einzelner Enzymfunktionen näher betrachtet und es wird ein möglicher Zusammenhang zur Körpergeruchsentstehung abgeleitet.

In Abbildung 35 sind die innerhalb des Stoffwechselweges des Leucin-, Isoleucin- und Valin-Abbaus zugeordneten Enzymfunktionen dargestellt. Insgesamt konnten 27 unterschiedliche Enzymfunktionen aus den Metatranskriptomik-Datensätzen zugeordnet werden. Davon korrelieren zwei Enzymfunktionen positiv, vier negativ und 21 nicht mit dem Körpergeruch. Die ersten drei Schritte des Abbaus der Aminosäuren L-Leucin, L-Valin und L-Isoleucin verlaufen analog und führen zur Entstehung von CoA-Konjugaten verzweigtkettiger Fettsäuren und schließlich zur Bildung freier verzweigtkettiger Fettsäuren. Der erste Schritt, die Transaminierung wird dabei von der BCAA-Aminotransferase (EC 2.6.1.42) ausgeführt. Dieses Enzym wurde bereits in Abschnitt 2.4.6 beschrieben und korreliert nicht deutlich mit dem Körpergeruch ( $r^2= -0,25$ ). Abweichend von der Verstoffwechslung der anderen beiden

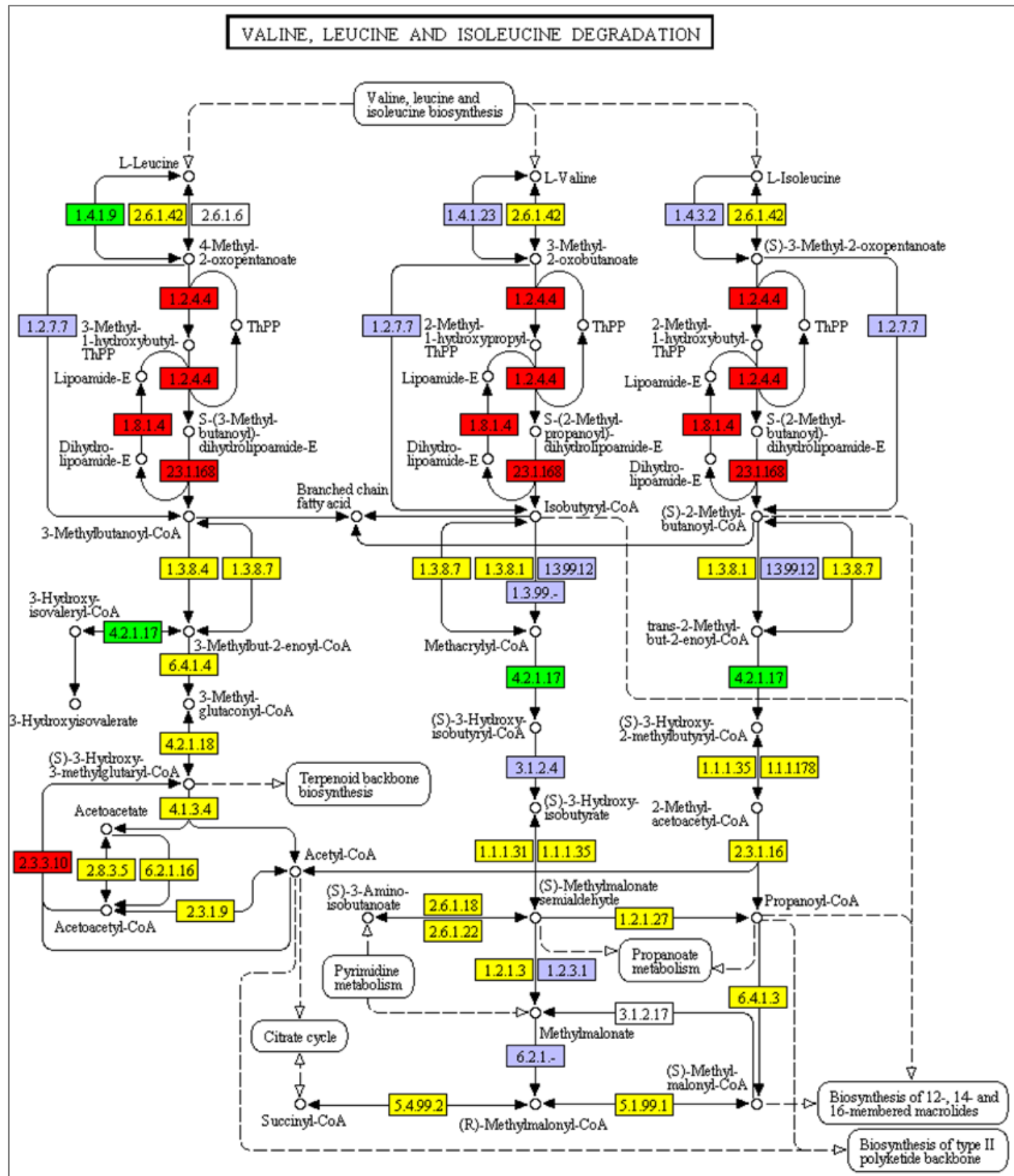


Abbildung 35: Darstellung positiv (grün), negativ (rot) und nicht mit dem Körpergeruch korrelierter (gelb) Enzymfunktionen im Stoffwechselweg des Leucin-, Isoleucin- und Valin-Abbaus auf Grundlage der KO-Ontologie mithilfe des KEGG-MAPPER SEARCH&COLOR PATHWAY-Tools. EC 1.4.1.9: Leucin-Dehydrogenase; EC 1.2.4.4: Verzweigt-kettige- $\alpha$ -Ketosäuren-Dehydrogenase; EC 1.8.1.4: Dihydrolipoyl-Dehydrogenase; EC 2.3.1.168: Dihydrolipoyllysinrest-(2-Methylpropanoyl)-Transferase; EC 4.2.1.17: Enoyl-CoA-Hydratase; EC 2.3.3.10: Acetoacetyl-CoA-Transacetylase.

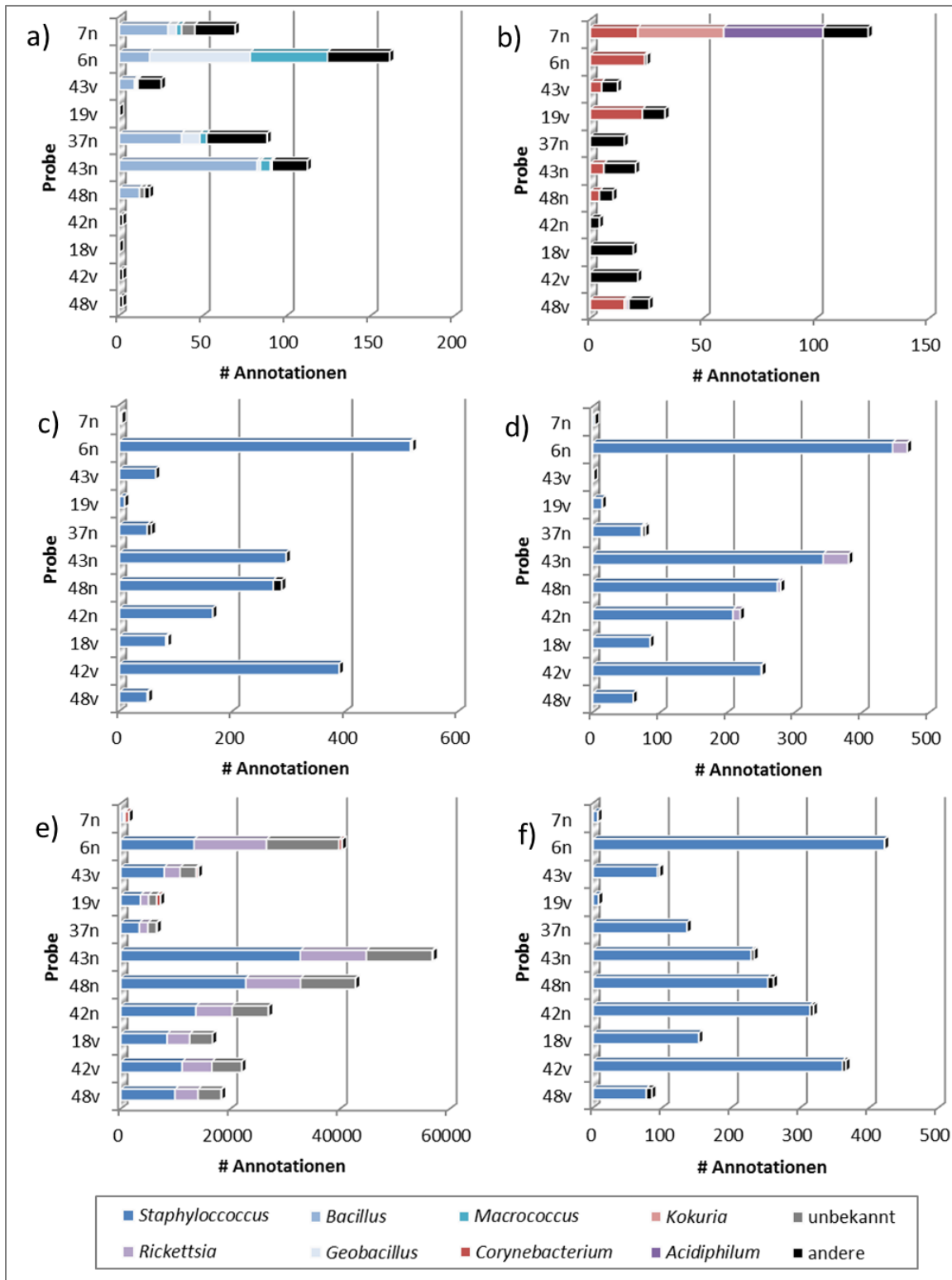


Abbildung 36: Taxonomische Profile relevanter Funktionen aus dem Stoffwechselweg des Leucin-, Isoleucin- und Valin-Abbaus auf Grundlage der KO-Ontologie. a) EC 1.4.1.9 (Leucin-Dehydrogenase); b) EC 4.2.1.17 (Enoyl-CoA-Hydratase); c) EC 1.2.4.4 (Verzweigt-kettige- $\alpha$ -Ketosäuren-Dehydrogenase); d) EC 2.3.1.168 (Dihydrolipoyllysine-(2-Methylpropanoyl)-Transferase); e) EC 1.8.1.4 (Dihydrolipoyl-Dehydrogenase); f) EC 2.3.3.10 (Acetoacetyl-CoA-Transacetylase).

Aminosäuren wird für die Transaminierung von Leucin zusätzlich die Leucin-Dehydrogenase (EC 1.4.1.9) zugeordnet, welche L-Leucin zu 4-Methyl-2-Oxopentanoat umwandelt. Dieses Enzym korreliert leicht positiv mit dem *Malodor Score* ( $r^2 = 0,57$ ) und wird auf taxonomischer Ebene den Genera *Bacillus*, *Geobacillus*, *Macrococcus* sowie anderen zugeordnet (s. Abbildung 36 a). Diese Organismen treten weder in der Liteartur noch in der rRNA-basierten taxonomischen Beschreibung dieser Metatranskriptomik-Datensätze als dominante Organismen auf. Daher kann vermutet werden, dass es sich bei dem taxonomischen Ursprung der entsprechenden Sequenzen um Staphylokokken handelt. Die Substanz 4-Methyl-2-Oxopentanoat wurde ebenfalls von Natsch et al. als Spaltungsprodukt der N<sup>α</sup>-Acylglutamin-Aminoacylase (N-AGA) aus frischen geruchslosen Schweißproben beschrieben (Natsch et al., 2006). Zum Geruch dieser Säure ist lediglich bekannt, dass sie stinkt (Sigma Aldrich Co. LL, 2013). Eine Spezifikation des Geruches und eine eventuelle Assoziation zu Körpergeruch ist nicht bekannt. Auch das analoge Zwischenprodukt aus dem Isoleucin-Abbau 3-Methyl-2-Oxopentanoat wurde als spaltbares Glutaminkonjugat in frischem Schweiß von Natsch et al. beschrieben (Natsch et al., 2006).

Die nächsten beiden enzymatischen Schritte des Abbaus werden durch die Verzweigt-kettige- $\alpha$ -Ketosäuren-Dehydrogenase (EC 1.2.4.4) katalysiert, wobei es sich bei der Dihydrolipoyllysinrest-(2-Methylpropanoyl)-Transferase (EC 2.3.1.168) um eine Untereinheit des Verzweigt-kettige- $\alpha$ -Ketosäuren-Dehydrogenase-Komplexes handelt und die Dihydrolipoyl-Dehydrogenase (EC 1.8.1.4) Dihydrolipoamid zu Lipoamid umwandelt. Diese Enzymfunktionen korrelieren leicht negativ mit dem Körpergeruch ( $r^2 \leq -0,5$ ) und werden auf taxonomischer Ebene überwiegend den Staphylokokken und seltener dem Genus *Rickettsia*, einem *Alpha-Proteobacterium*, zugeordnet (s. Abbildung 36 c-e). Die Umwandlung der methylverzweigten 2-Ketosäuren führt in allen drei Fällen zu CoA-Konjugaten verzweigter Fettsäuren und schließlich zur Bildung freier verzweigter Fettsäuren wie 3-Methylbutansäure (Isovaleriansäure), Isobutansäure und (S)-2-Methylbutansäure. Isovaleriansäure wird in der Literatur als eine Geruchskomponente des Schweißes beschrieben und die Arbeitsgruppe um James et al. postuliert eine zentrale Rolle der Bildung dieser Substanz durch Staphylokokken bei der Körpergeruchsentstehung (James et al., 2013). Von Isobutansäure ist bekannt, dass diese stinkt, wobei jedoch auch hier kein sensorischer Zusammenhang zu Körpergeruch bekannt ist (Sigma Aldrich Co. LL, 2016). Das (S)-Enantiomer der 2-Methylbutansäure wird als fruchtig und süß beschrieben, im Gegensatz zu seinem käsigen nach Schweiß riechenden (R)-Enantiomer. Die Acetoacetyl-CoA-Transacetylase (EC 2.3.3.10) wandelt Acetoacetyl-CoA und Acetyl-CoA zu (S)-3-Hydroxy-3-Methylglutaryl-CoA um, welches in die Terpenoid-Biosynthese einfließt. Die Enzymfunktion korreliert negativ mit Körpergeruch ( $r^2 \leq -0,6$ ) und wird auf taxonomischer Ebene den Staphylokokken zugeordnet (s. Abbildung 36 f).

Besonders auffällig in diesem Stoffwechselweg ist die Enoyl-CoA-Hydratase (EC 4.2.1.17), welche auch am Fettsäure-Abbau beteiligt ist (vgl. Abschnitt 2.4.6). Diese wird auf taxonomischer Ebene durch die Genera *Corynebacterium*, *Kokuria* (ein Actinobakterium), *Acidiphilum* (ein Alpha-Proteobakterium) und andere beschrieben (s. Abbildung 36 b). Da die letzteren beiden Genera weder in der Literatur noch in der rRNA-basierten Taxonomie dieser Datensätze als dominant beschrieben sind, handelt es sich bei dem taxonomischen Ursprung der zugrunde liegenden Sequenzen vermutlich um unbekannte und/oder unkultivierte Organismen. Weiterhin steht das taxonomische Profil dieser Enzymfunktion (s. Abbildung 36 b) nicht im Einklang mit den in Abbildung 32 c) dargestellten Zuordnungen auf Basis der Subsystems-Ontologie. Die Enzymfunktion korreliert stark positiv mit dem Körpergeruch ( $r^2 \leq 0,8$ ). Das Enzym wandelt im Leucin-, Isoleucin- und Valin-Abbauweg 3-Methylbut-2-enoil-CoA und 3-Hydroxy-isovaleryl-CoA ineinander um. Letztere Substanz kann dann durch Abspaltung von CoA 3-Hydroxy-Isovaleriansäure (3-Hydroxy-3-methylbutanonsäure) bilden, welche unangenehm riecht (Sigma Aldrich Co. LL, 2012). Da sowohl die taxonomische Zuordnung dieser Enzymfunktion als auch die *Read*-Abundanzen, bzw. die Anzahl der Annotationen stark von den übrigen für diesen Stoffwechselweg beschriebenen Funktionen abweicht, kann angenommen werden, dass die Enoyl-CoA-Hydratase in den Metatranskriptomik-Datensätzen möglicherweise an einem anderen Stoffwechselweg als dem Leucin-, Isoleucin- und Valin-Abbau beteiligt ist.

Bei dem zweiten in diesem Zusammenhang betrachteten Stoffwechselweg handelt es sich um den Glutathion-Stoffwechsel. Von insgesamt 42 unterschiedlichen Enzymfunktionen korrelieren zwei positiv und zwei negativ mit dem Körpergeruch. 18 Enzymfunktionen werden dargestellt, welche nicht oder nur schwach korrelieren. Besonders interessant ist dabei der Glutathion-Konjugat-Abbauweg, welcher auch von Starkenmann et al. als Ursprung von Cysteinyl-S-Konjugaten geruchsbildender Sulfanylalkanole postuliert wurde (Starkenmann et al., 2005). Dieser besteht aus drei enzymatischen Schritten und beginnt mit der Übertragung eines unspezifischen Restes auf das Glutathion und der Bildung eines Glutathionyl-S-Konjugates durch die Glutathion-Transferase (EC 2.5.1.8) (s. Abbildung 37). Dieser Mechanismus dient dem Schutz der Zelle vor toxischen Fremdstoffen und oxidativem Stress. In den hier vorliegenden Metatranskriptomik-Datensätzen korreliert die Enzymfunktion leicht positiv mit dem *Malodor Score* ( $r^2 = 0,62$ ). Das taxonomische Profil der Glutathion-Transferase wird durch die Genera *Staphylococcus*, *Ralstonia* und andere beschrieben (s. Abbildung 38 a). Dabei werden die entsprechenden *Reads* der Datensätze 7n, 43v, 19v, 37n und 42v ausschließlich anderen Genera zugeordnet. Die maximale Anzahl der jeweiligen *Read*-Abundanzen ist mit maximal 15 (Datensatz 48n) gering.

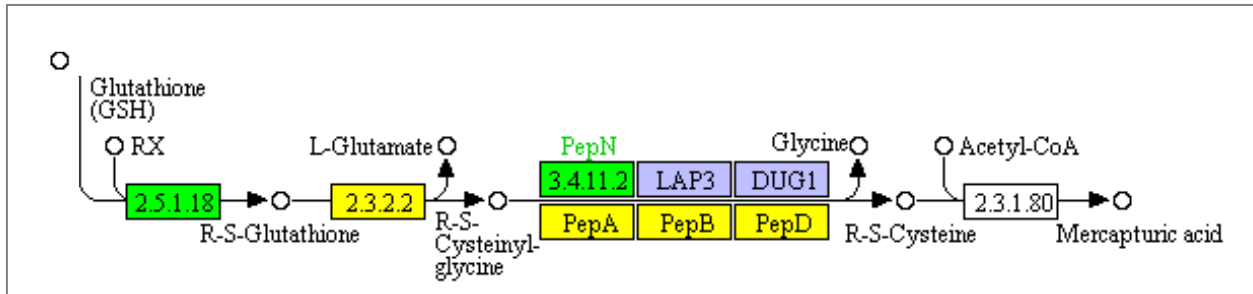


Abbildung 37: Darstellung positiv (grün), negativ (rot) und nicht korrelierter (gelb) Enzymfunktionen im Glutathion-Stoffwechsel auf Grundlage der KO-Ontologie mithilfe des KEGG-MAPPER SEARCH&COLOR PATHWAY – Tools. EC 2.5.1.8: Glutathion-Transferase; EC 2.3.2.2:  $\gamma$ -Glutamyltransferase; EC 3.4.11.1: Aminopeptidase PepA; EC 3.4.11.2: PepN-Aminopeptidase.

Im darauffolgenden Abbauweg wird zunächst der Glutamatrest des Glutathions durch die  $\gamma$ -Glutamyltransferase (EC 2.3.2.2) abgespalten. Diese Enzymfunktion korreliert nicht mit dem Körpergeruch ( $r^2 = -0,28$ ) und ist auf taxonomischer Ebene von den Staphylokokken dominiert (s. Abbildung 38 b). Dabei werden mit mindestens 21 (7n) und maximal 1089 (6n) deutlich mehr Annotationen als für die vorangehende Reaktion verzeichnet. Auffallend ist, dass Datensätze, die im vorangegangenen Schritt durch Annotationen zu anderen Genera dominiert wurden, für diese enzymatische Reaktion relativ wenige Annotationen aufweisen (vgl. Abbildung 38 a). Das auf diese Weise entstandene Cysteinylglycyl-Konjugat wird in einer weiteren Reaktion durch Abspaltung von Glycin zu R-S-Cystein abgebaut. Dieser Abbauweg wurde von Natsch et al. für Konjugate geruchsbildender Sulfanylalkanole wie 3-Methyl-3-Sulfanylhexan-1-ol durch die TpdA-Dipeptidase in *Corynebacterium Ax20* demonstriert (Emter und Natsch, 2008). In den hier vorliegenden Metatranskriptomik-Daten, welche auf Basis der KO-Ontologie ausgewertet wurden, werden diesem Abbauschritt insgesamt vier verschiedene funktionale Annotationen zugeordnet (s. Abbildung 37). Die Peptidasen PepB und PepD weisen jedoch insgesamt lediglich eine bzw. vier *Read*-Abundanzen in dem Datensatz der Probe 42n auf und werden daher nicht weiter betrachtet. Die PepA-Aminopeptidase korreliert nicht ( $r^2 = 0,15$ ) mit dem *Malodor Score* und wird auf taxonomischer Ebene durch die Genera *Staphylococcus*, *Corynebacterium*, *Streptobacillus*, *Treponema* und andere beschrieben, wobei in den meisten Datensätzen Staphylokokken und im Falle der Datensätze 7n, 43v und 19v Corynebakterien dominieren. Die Genera *Streptobacillus* und *Treponema* wurden bisher nicht als Mitglieder der Achselhöhlen-Mikrobiota beschrieben. Daher handelt es sich bei den entsprechenden Annotationen wahrscheinlich um fehlerhafte Annotation von Sequenzen mit anderem Ursprung. Die PepN-Aminopeptidase (EC 3.4.11.2) katalysiert ebenfalls die Spaltung eines Cysteinylglycyl-Konjugates in R-S-Cystein und Glycin. Diese Enzymfunktion korreliert in den hier beschriebenen Datensätzen positiv ( $r^2 = 0,69$ ) mit dem *Malodor Score* und wird auf taxonomischer Ebene den



Genera *Corynebacterium*, *Propionibacterium*, *Saccharopolyspora* und der Gruppe „andere“ zugeordnet. Dabei sind Annotationen zu Corynebakterien in den meisten Datensätzen dominant. Lediglich im Datensatz der Probe 37 dominiert *Propionibacterium*. Der nächste Abbauschritt des Cystein-Konjugates ist in der Repräsentation des Stoffwechselweges auf Grundlage der KO-Ontologie nicht abgebildet. Aufgrund der unterschiedlichen taxonomischen Zuordnungen der einzelnen Annotationen und der z. T. starken Unterschiede in den *Read*-Abundanzen ist davon auszugehen, dass der beschriebene Stoffwechselweg nicht im selben Organismus abläuft.

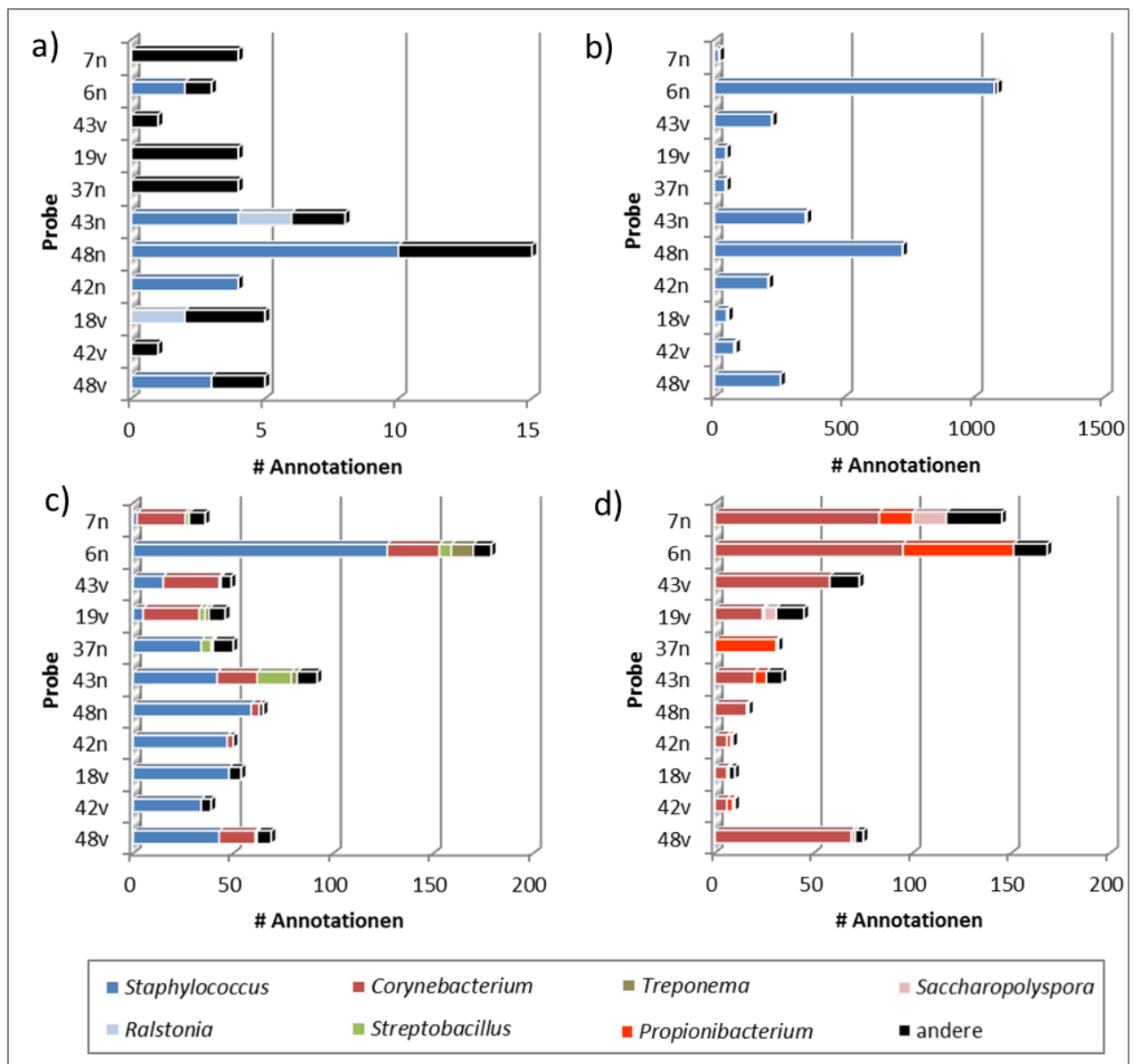


Abbildung 38: Taxonomische Profile relevanter Funktionen aus dem Glutathion-Metabolismus auf Grundlage der KO-Ontologie. a) Glutathion-Transferase (EC 2.5.1.8); b)  $\gamma$ -Glutamyltransferase (EC 2.3.2.2); c) PepA-Amino-peptidase (EC 3.4.11.1); d) PepN-Amino-peptidase (EC 3.4.11.2).

Einen weiteren im Zusammenhang mit potentiell geruchsbildenden Molekülen interessanten Stoffwechselweg stellt der Phenylalanin-Stoffwechsel dar. In Abbildung 40 ist der Ausschnitt der Umwandlung von Phenylacetaldehyd zu Phenylacetylglutamin dargestellt. Dabei korreliert das Enzym Phenylacetaldehyd-Dehydrogenase (EC 1.2.1.39), welches Phenylacetaldehyd zu Phenylacetat umwandelt, positiv mit dem Malodor Score ( $r^2 = 0,56$ ). Phenylacetat sowie das verwandte Molekül 4-Hydroxy-Phenylacetat wurden ebenfalls von Natsch et al. als Spaltungsprodukt des Enzyms N-AGA in frischem Schweiß identifiziert (Natsch et al., 2006). Das taxonomische Profil des Enzyms Phenylacetaldehyd-Dehydrogenase wird ausschließlich durch Corynebakterien und andere Genera beschrieben (s. Abbildung 40 a). Dabei weisen die Datensätze der Proben 7n, 43v, 19v und 43n, welche auf rRNA-Ebene ebenfalls durch corynebakterielle Annotationen dominiert werden, als einzige Datensätze Sequenzen mit Homologie zu Corynebakterien auf. Die Gesamtzahl der Annotationen ist mit

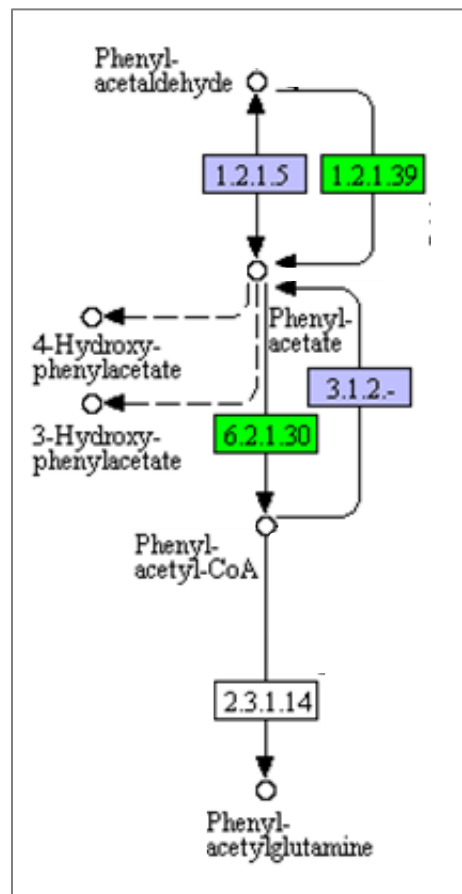


Abbildung 39: Darstellung der positiv (grün) korrelierten Enzymfunktionen Phenylacetaldehyd-Dehydrogenase (EC 1.2.1.39) und Phenylacetat-Ligase (EC 6.2.1.30) aus dem Phenylalanin-Stoffwechsel auf Grundlage der KO-Ontologie mithilfe des KEGG-MAPPER SEARCH&COLOR PATHWAY –Tools.

maximal 26 (7n) im Vergleich zu den anderen in diesem Kapitel betrachteten Funktionen eher gering. In den Datensätzen der Proben 6n, 37n, 48n, 42n, 42v und 48v, werden keine Sequenzhomologien zur Phenylacetat-Ligase beschrieben.

Der nächste enzymatische Schritt führt zur Bildung von Phenyl-Acetyl-CoA und wird durch das Enzym Phenylacetat-Ligase (EC 6.2.1.30) katalysiert. Diese Funktion korreliert positiv ( $r^2 = 0,65$ ) mit dem Körpergeruch. Auf taxonomischer Ebene wird diese Funktion Mitgliedern der Gattungen *Finegoldia* (7n), *Bordetella* (19v) sowie *Ralstonia* (37n und 18v) zugeordnet. Aufgrund der insgesamt sehr geringen Read-Abundanzen sowie der Tatsache, dass die Genera *Bordetella* und *Ralstonia* aus der Literatur nicht im Zusammenhang mit der Achselhöhlen-Mikrobiota beschrieben sind, handelt es sich um Sequenzen aus unbekanntem Organismen, die vermutlich ohne biologische Bedeutung sind. Der nächste enzymatische Schritt ist die Übertragung von Phenylacetat auf ein Glutamin. Dieser Schritt wird durch die Glutamin N-Phenylacetyltransferase katalysiert (EC 2.3.1.14) und wird in keinem der hier beschriebenen Metatranskriptomik-Datensätze beschrieben.

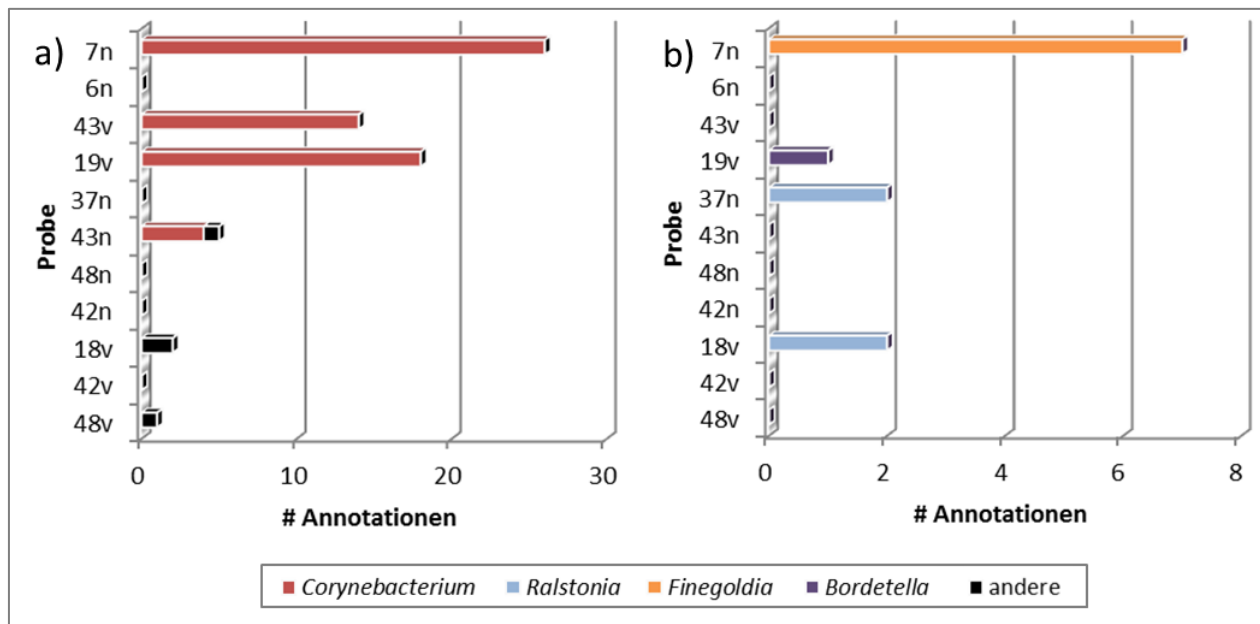


Abbildung 40: Taxonomische Profile der Annotationen zur Phenylacetaldehyd-Dehydrogenase (EC 1.2.1.39) und Phenylacetat-Ligase (EC 6.2.1.30) aus dem KO-Annotationssystem auf Genus-Ebene. Dargestellt sind Zuordnungen zu den Gattungen *Corynebacterium* (dunkelrot), *Ralstonia* (hellblau), *Finegoldia* (orange), *Bordetella* (lila) und zu der Gruppe „andere“ (schwarz).

## V. Diskussion

Die Beschreibung des menschlichen Körpers als Supraorganismus, bestehend aus somatischen Zellen und den Zellen der Mikrobiota sowie die mengenmäßige Überlegenheit der Mikrobiota im Hinblick auf die Zellzahl und auf die genetische Information, führt unweigerlich zur Anerkennung der Rolle der Mikrobiota bei physiologischen Vorgängen im Körper und der Entstehung von Krankheiten. Dies wiederum führt zur Notwendigkeit, die menschliche Mikrobiota auf taxonomischer und auf funktionaler Ebene zu beschreiben. Kultivierungsbasierte Ansätze sind jedoch nicht in der Lage, alle in einem Lebensraum vorkommenden Bakterienspezies abzubilden, da etwa 90 % - 99 % der Organismen nicht ohne Weiteres kultivierbar sind. Die Hochdurchsatzsequenzierung stellt die passende Methode zur Untersuchung der menschlichen Mikrobiota dar, da sie die Beschreibung einer großen Anzahl von z. T. unbekanntem Sequenzen ermöglicht. Die Metatranskriptomik ermöglicht die Beschreibung der Zusammensetzung und Funktionen einer Lebensgemeinschaft auf RNA-Ebene zu einem bestimmten Zeitpunkt. Das Probenmaterial, das von der menschlichen Mikrobiota gewonnen werden kann, ist jedoch in der Regel stark limitiert, sodass bisher nur zwei Studien zur Metatranskriptomik der menschlichen Mikrobiota veröffentlicht wurden (Gosalbes et al., 2011; Twin et al., 2013). Die wahrscheinlich reichhaltigste Quelle für Probenmaterial der Mikrobiota im menschlichen Körper ist der Darm, dessen Lebensgemeinschaften auch als erste in der Literatur beschrieben wurden (Gosalbes et al., 2011). Aber auch die vaginale Mikrobiota liefert im Zusammenhang mit einer bakteriellen Vaginose genügend Probenmaterial zur Durchführung von Metatranskriptom-Analysen (Twin et al., 2013). In der vorliegenden Arbeit wurden erstmals zwei bisher nicht beschriebene Habitate der menschlichen Mikrobiota auf Metatranskriptom-Ebene analysiert: Eine Probe aus dem Außenohr eines Patienten mit Gehörgangsentzündung sowie Proben aus der Achselhöhle männlicher Probanden im Zusammenhang mit der Anwendung eines Deodorants und Wasser. Im Folgenden werden zum einen die entwickelten und angewandten Techniken und ihre Eignung für die Analyse der menschlichen Mikrobiota diskutiert. Zum anderen werden die gewonnenen Erkenntnisse inhaltlich diskutiert und Bezüge zwischen Taxonomie und Funktion sowie die Auswirkungen der aktiven Mikrobiota auf den Menschen im Hinblick auf Gehörgangsentzündung und Körpergeruchsentstehung beschrieben.

### 1 Methode zur Beschreibung der menschlichen Mikrobiota

Eine Methode zur Beschreibung der Mikrobiota der Achselhöhle wurde bereits 2013 postuliert (Fredrich et al., 2013). Diese umfasste im Wesentlichen sechs Arbeitsschritte: Probennahme,

RNA-Extraktion, cDNA-Synthese und Amplifikation, Herstellung einer Sequenzierbibliothek, Sequenzierung mittels NGS-Techniken und Datenanalyse auf taxonomischer und funktionaler Ebene (Abbildung 41). Diese Arbeitsschritte wurden für die Proben der Lebensgemeinschaften des Außenohrs und der Achselhöhle umgesetzt und werden im Folgenden näher beschrieben.

## **1.1 Etablierung einer Labormethode zur Beschreibung der mikrobiellen Lebensgemeinschaften des Ohrs und der Achselhöhle**

Im Rahmen von Kooperationen wurden zwei unterschiedliche Probentypen der Lebensgemeinschaften des Ohrs bzw. der Achselhöhle zur Verfügung gestellt. Bei der Ohrprobe handelte es sich um einen Tupfer mit wenigen  $\mu\text{l}$  Probenmaterial, welches bis zum Zeitpunkt der RNA-Extraktion unverdünnt vorlag. Bei den Proben aus der Achselhöhle handelte es sich um 1 ml Flüssigkeit mit einer mittleren Zelldichte im Bereich von  $10^5$  KBE/ml. Bei dem Lösungsmittel handelte es sich um eine physiologische Kochsalzlösung. Jeder der durchgeführten und beschriebenen Arbeitsschritte verzerrt die messbare Aktivität und Zusammensetzung der untersuchten Lebensgemeinschaften gegenüber dem Zustand unmittelbar vor der Probennahme. So wirkt sich die Verwendung eines künstlichen Lösungsmittels bei den Proben der Achselhöhle aufgrund der starken Veränderung der Umgebungsbedingungen vermutlich auch auf die Aktivität der Lebensgemeinschaften und somit auf die Menge und Art darstellbarer RNA-Sequenzen aus. Weiterhin könnte die lange Transferzeit der Probe bei Raumtemperatur im Falle der Ohrprobe ebenfalls das gemessene transkriptionelle Profil verändert haben. Dennoch werden im Rahmen dieser Arbeit schließlich erstmals Metatranskriptomik-Datensätze generiert, die zum einen weitestgehend mit den in der Fachliteratur beschriebenen Erkenntnissen übereinstimmen, und zum anderen – im Falle der Achselhöhlen-Proben – Unterschiede in den transkriptionellen Profilen unterschiedlicher Proben abbilden. Die möglicherweise durch die einzelnen Arbeitsschritte eingeführten Verzerrungen können daher als geringfügig im Vergleich zu den gelieferten Informationen betrachtet werden.

Bei der Labormethode zur Metatranskriptom-Analyse der Proben stellte die RNA-Isolierung den kritischen Arbeitsschritt dar. Die Limitierung des Probenmaterials und damit einhergehend der RNA-Ausbeute führte auch gleichzeitig zur Limitierung der analysierbaren Proben. Nur wenn genügend RNA isoliert werden konnte, konnte die jeweilige Probe weiterführend bearbeitet werden. Daher konnten von insgesamt 48 Proben aus der Achselhöhle lediglich 14 analysiert werden. Die Mindestmenge an benötigter RNA lag etwa bei 500 pg. Die Ausbeute an Gesamt-RNA nach Säulenextraktion lag bei 35 ng für das Ohr, bzw. bei durchschnittlich 2,3 ng für die Achselhöhle. Da die Ohrprobe im Zusammenhang mit einer Gehörgangsentzündung genommen

wurde und bei mikrobiellen Entzündungen erhöhte Keimzahlen vorliegen, ist der Unterschied in der RNA-Ausbeute um einen Faktor  $> 10$  gegenüber der Achselhöhle nachvollziehbar. Auch die Anwesenheit eukaryotischer Pathogene wie *Candida* spp. und die Immunreaktion des Wirtes erhöhen vermutlich die Menge an isolierter RNA im Vergleich zu den Proben aus der Achselhöhle. Auch die Beschaffenheit der Probe hat Einfluss auf die Extraktionsgeschwindigkeit und somit auf die Menge und Integrität der RNA, da RNAsen sich nur sehr schwer inaktivieren lassen. Im Gegensatz zu einem Probenröhrchen mit Flüssigkeit, das längere Zeit zum Auftauen benötigt, kann Probenmaterial von einem Tupfer direkt extrahiert werden. Die Säulenextraktion stellt im Allgemeinen eine schnelle, zuverlässige und effiziente Möglichkeit zur RNA-Gewinnung dar, wobei jedoch sehr kleine RNAs ( $< 100$  nt) nicht extrahiert werden. Die Qualität der extrahierten RNA variierte für die Proben aus der Achselhöhle zwischen einer sehr schlechten, nicht determinierbaren Qualität für die Probe 43n und einer guten Qualität mit einer RIN von 8 für die Probe 7n. Da es sich bei der aus der Ohrprobe extrahierten Gesamt-RNA um ein Gemisch aus prokaryotischer und eukaryotischer RNA handelt, stellt die RIN von 7,7 vermutlich lediglich einen ungenauen Wert dar. Das entsprechende Elektropherogramm weist jedoch auf eine gute RNA-Integrität hin (s. Abbildung 2). Aufgrund der sehr geringen Mengen an extrahierter Gesamt-RNA wurde auf eine Abreicherung ribosomaler RNA verzichtet. Dies führte dazu, dass rRNA-Sequenzen gemäß ihrer natürlich vorkommenden zahlenmäßigen Überlegenheit auch in den Sequenzdatensätzen gegenüber nichtribosomalen RNAs dominant waren. Dies stellte jedoch bei der Auswertung keinen Nachteil dar, da die Annotationen der rRNA-Sequenzen für die taxonomische Beschreibung der aktiven Lebensgemeinschaften genutzt werden konnten und die taxonomische Zusammensetzung zuverlässiger beschreiben konnten als funktionale RNAs.

Um eine Sequenzierung der Extrakte durchzuführen, erfolgte zunächst die cDNA-Synthese, welche mit sogenannten *Random- und Poly-T-Primern* geprimt wurde. Diese Methode ermöglichte die unselektive Erfassung aller in den Extrakten befindlichen RNAs, was wiederum insbesondere im Falle der Ohrprobe die Beschreibung einer großen Bandbreite mikrobieller Taxa und auch menschlicher Sequenzen ermöglichte. Der Nachteil dieser Methode liegt jedoch darin, dass es sich um einen ungerichteten Ansatz handelt, bei dem, im Gegensatz zur 16S-rRNA-Amplikon Sequenzierung, weder der Startpunkt noch der Endpunkt der Amplifikate bestimmt ist. Solche Ansätze können zwar die Beschreibung vieler diverser rRNAs und mRNA gewährleisten, erschweren jedoch die Interpretation der Daten (Morales und Holben, 2011). Die cDNA-Amplifikation erfolgte isothermal und die amplifizierte doppelsträngige cDNA hatte eine sehr breite Längenverteilung im Bereich zwischen 100 bp und 1000 bp.




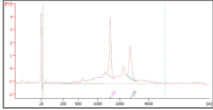
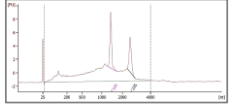

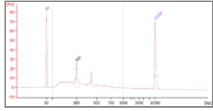
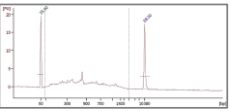
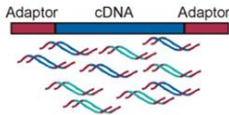
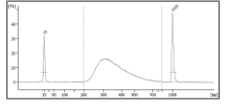




| Arbeitsschritt                                  | Schema  | Ohr  | Achselhöhle   |
|---|---|--|---|
| 1. Probennahme der humanen Mikrobiota           | <p>Probe</p>                   | <p>Ohr-Probe eines 77-jährigen Patienten mit Otitis externa</p>  <p>Dr. Harald Brauer,<br/>Hals-, Nasen-, Ohrenarzt,<br/>Bielefeld</p> | <p>Proben aus den Achselhöhlen gesunder Probanden</p> <p>(~ 1 ml Flüssigkeit)</p> <p>Industrieller Kooperationspartner</p>              |
| 2. Extraktion der Gesamt-RNA                    | <p>RNA</p>                     |  <p>Säulenbasierte RNA-Isolierung</p>  |  <p>Säulenbasierte RNA-Isolierung</p>                |
| 3. cDNA-Synthese und Amplifikation              | <p>cDNA</p>                    |  <p>Random-Priming und isothermale Amplifikation</p>   |  <p>Random-Priming und isothermale Amplifikation</p> |
| 4. Herstellung von cDNA-Sequenzier-Bibliotheken | <p>Sequenzierbibliothek</p>  | <p>Adapter-Ligation</p> <p>vertis Biotechnologie AG</p>  |  <p>Adapter-Ligation</p>                           |
| 5. Sequenzierung mittels NGS-Techniken          | <p>Sequenz-Reads</p>         | <p>~ 144 x 10<sup>6</sup> Reads</p> <p>vertis Biotechnologie AG</p>  | <p>~ 6 x 10<sup>6</sup> Read-Paare/Probe</p> <p>Sequencing Core Facility,<br/>CeBiTec Bielefeld</p>                                     |
| 6. Phylogenetik und funktionale Analyse         | <p>Datenanalyse</p>          | <p>METRANS</p>   |    |

Abbildung 41: Postulierte Methode zur Metatranskriptom-Analyse kleiner menschlicher Proben und die Umsetzung für die Probe des Außenohres und die Proben aus der Achselhöhle in der vorliegenden Arbeit; (Fredrich et al., 2013) modifiziert.

Anschließend wurde die cDNA mittels Ultraschall fragmentiert und es wurde mittels Adapterligation eine cDNA-Sequenzierbibliothek synthetisiert. Für die Ohrprobe wurde dieser Arbeitsschritt durch einen externen Dienstleister, die vertis Biotechnologie AG, durchgeführt. Für die Achselhöhlen-Proben erfolgte dieser Schritt am CeBiTec. Trotz der unterschiedlichen Mengen und Integritäten der RNA wurden für die cDNA-Sequenzierbibliotheken der Achselhöhlenproben sehr einheitliche Längenprofile von 200 bp bis 700 bp erzielt. Das Maximum der Längenverteilung lag bei 300 bp.

Aufgrund der beschriebenen Längenverteilung und der technischen Möglichkeiten wurde für die Achselhöhlen-Proben eine *Paired-End*-Sequenzierung auf der Illumina MiSeq-Plattform gewählt. Die Leseweite betrug 250 nt in jede Richtung. Somit wurden Sequenzen mit einer Länge > 500 bp nicht vollständig erfasst und in der späteren Analyse verworfen. Sequenzen mit einer Länge  $\leq$  250 bp wurden doppelt sequenziert. Insgesamt wurden im Mittel 6 Millionen *Reads* pro Probe erzeugt. Bei einer mittleren Keimzahl im Bereich von  $10^5$  entspricht dies etwa 10 *Reads* pro Keim. Aufgrund des kollektiven Charakters der Methode reicht diese Sequenziertiefe sowohl für eine allgemeine Beschreibung der Aktivität der Lebensgemeinschaften als auch für die Beschreibung spezifischer Funktionen im Zusammenhang mit Körpergeruchsentstehung aus. Die Sequenzierung der Ohrprobe erfolgte ebenfalls durch die vertis Biotechnologie AG auf der Illumina HiSeq-Plattform und lieferte etwa 144 Millionen *Reads* mit einer Leseweite von lediglich 100 bp. In einer proteinkodierenden RNA deckt diese lediglich einen Bereich von maximal 33 Aminosäuren ab. Die Qualität der Sequenzierdaten unterscheidet sich ebenfalls unter den zwei Probenotypen. Die mittlere Qualität der Sequenzdaten aus dem Ohr liegt lediglich an den ersten 75 Basenpositionen bei einem *Phred-Score* von  $\geq 30$ , was einer Fehlerwahrscheinlichkeit von  $\leq 0,1$  % entspricht. Die hinteren 25 Basen wiesen einen *Phred-Score* im Bereich zwischen 25 und 30 auf. Die Qualitätskurven der Sequenzdatensätze der Achselhöhle weisen ein auf den ersten Blick ungewöhnliches Profil auf, welches sich jedoch durch die doppelte Sequenzierung kurzer Sequenzen ( $\leq 250$  bp) erklären lässt. Aufgrund dieser Besonderheit liegt die mittlere Qualität der Sequenzen in diesem Bereich bei einem *Phred-Score* zwischen 50 und 60 (Fehlerwahrscheinlichkeit zwischen 0,001 % und 0,0001 %). Im hinteren Bereich der Sequenzlängen (251 bp – 500 bp) liegt die Qualität im Bereich zwischen 37 und 30. Ein mittlerer *Phred-Score* von 30 wird für die Sequenzdatensätze der Achselhöhlenproben in der Regel nicht unterschritten. Diese Sequenzdaten sind denen der Ohrprobe sowohl im Hinblick auf die Länge als auch im Hinblick auf die Qualität überlegen.



**Fazit:**

- In den im Rahmen dieser Arbeit durchgeführten Experimenten wurde eine Labormethode entwickelt, welche erstmals die Beschreibung der Lebensgemeinschaften des Außenohrs und der Achselhöhle auf Metatranskriptom-Ebene ermöglicht.
- Aufgrund der sehr geringen Mengen an Probenmaterial stellte die RNA-Isolierung den kritischen Arbeitsschritt dar.

## 1.2 Bioinformatik-Analyse von Metatranskriptomik-Datensätzen der menschlichen Mikrobiota

Die Erzeugung immer größerer Datenmengen mittels NGS-Sequenzieretechniken erfordert die Entwicklung immer weiter voranschreitender bioinformatischer Analysesysteme zur Gewährleistung der Auswertbarkeit der generierten Daten. Im Vergleich zur Sequenzierung einzelner Transkriptome bekannter Organismen fehlt in der Metatranskriptomik die Möglichkeit, die Daten auf ein bereits vorhandenes Genom zu mappen. Die funktionale und taxonomische Zuordnung einzelner Sequenzen ist somit nicht immer eindeutig. Auch im Vergleich zu stark gerichteten Ansätzen wie der 16S-rDNA- oder 16S-rRNA-Sequenzierung, welche wie die Metagenomik und Metatranskriptomik auf Datenbank-Abfragen beruhen, fehlt bei der hier verwendeten Methode die Festlegung definierter Start- und Endpunkte der Sequenzen. Dies kann dazu führen, dass der Informationsgehalt einzelner Sequenzzuordnungen qualitativ stark variiert, wenn beispielsweise stark konservierte oder stark variable RNA-Regionen von den Sequenzdaten abgebildet werden.

Eine Analysepipeline zur Auswertung *random* geprimter und amplifizierter Metatranskriptome muss im Wesentlichen die folgenden Probleme lösen, um die inhaltliche Auswertung von Metatranskriptomik-Datensätzen zu ermöglichen: Qualitätskontrolle, Datenreduktion, funktionale und taxonomische Zuordnung, Normalisierung und Darstellung. In der vorliegenden Arbeit wurden die beiden Analysepipelines METRANS und MG-RAST verwendet. Während die Datensätze aus den Achselhöhlenproben ausschließlich mit MG-RAST analysiert wurden, bietet der Datensatz aus dem Ohr die Möglichkeit eines direkten Vergleiches der beiden Analysepipelines. Im Folgenden werden die einzelnen Analyseschritte beschrieben und sowohl aus der Perspektive der Anwendbarkeit als auch auf inhaltlicher Ebene bewertet.

Die Qualitätskontrolle dient der Entfernung einzelner Sequenzen mit geringer Qualität, bzw. dem Eliminieren von Sequenzabschnitten mit geringer Qualität. In MG-RAST werden Sequenzen getrimmt, sobald mindestens 5 Basen mit einem Qualitätswert (*Phred-Score*) < 15 vorliegen. Dies führt zu einer zahlenmäßigen Erhaltung der Sequenzdaten, jedoch auch zur Verkürzung der auswertbaren Sequenzen. Im Falle der Ohrprobe werden dadurch lediglich 2,4 % der Sequenzen entfernt. In METRANS wird ein *Sliding Window Approach* realisiert, welcher im Normalfall ebenfalls zu einer Kürzung der Sequenzen führt. Bei der Auswertung der Ohrprobe hat sich dieser jedoch als nicht praktikabel erwiesen. Als Konsequenz wurde ein Fenster gewählt, welches die gesamte *Read-Länge* abdeckt, was schließlich analog zur Situation in MG-RAST in einem *Quality-Filtering* mündete, bei dem etwa 55 % aller in der Ohrprobe verfügbaren Sequenzen entfernt wurden. Der Vorteil dieser Vorgehensweise liegt grundsätzlich in dem Erhalt der Sequenzlänge und insbesondere bei der Wahl eines hohen Qualitätsschwellenwertes in einer guten Qualität der weiterführend analysierten Sequenzen.

Die Datenreduktion dient der Erzielung einer handhabbaren Auswertungsdauer und Sequenzanzahl durch Zusammenfassung von aus der Amplifikation hervorgegangenen Sequenzduplikaten. In METRANS wird der DNACLUST-Algorithmus genutzt, um Sequenzen mit mindestens 98 % Sequenzidentität zusammenzufassen. Ein ähnliches Vorgehen wird auch in der RNA-Pipeline von MG-RAST angewendet, wobei die minimale Sequenzähnlichkeit bei 97 % liegt. Diese Methode fasst sehr effizient sowohl aus der Amplifikation hervorgegangene als auch natürlich vorkommende Sequenzen mit großer Ähnlichkeit zusammen. Im Falle von METRANS führt dieser Schritt zu einer weiteren Datenreduktion um einen Faktor 13. Im Falle von MG-RAST liegt die Datenreduktion lediglich bei einem Faktor von 2,7, wobei der vorangegangene Schritt der *rRNA-Identification* in dieser Betrachtung nicht berücksichtigt ist. Ein Nachteil dieser Methode liegt darin, dass alle nachfolgenden Analyseschritte nicht mit den nativen Sequenzen, sondern lediglich mit der Konsensus-Sequenz durchgeführt werden, was eine potentielle Fehlerquelle darstellt. In der Protein-*Pipeline* von MG-RAST werden hingegen alle Sequenzen, die in ihren ersten 20 Basen mit einer anderen Sequenz im Datensatz übereinstimmen, entfernt. Auf diese Weise werden lediglich Sequenzartefakte, die durch die Amplifikation hervorgegangen sind, entfernt. Dies führt zu einer Datenreduktion um einen Faktor 14. Zusätzlich wird in einem nachfolgenden Analyseschritt noch einmal geclustert. Somit wird in beiden Pipelines im Falle der Ohrprobe eine sehr effiziente Datenreduktion realisiert.

Die funktionale und taxonomische Zuordnung der Sequenzen erfolgt in beiden Pipelines im Wesentlichen über eine Datenbankabfrage. In METRANS erfolgt die Identifizierung nicht proteinkodierender RNAs durch eine Abfrage in den *SILVA*-Datenbanken *LSU*, *SSU* und *rfam*. Die

funktionale Abfrage wird in den Datenbanken *KEGG*, *SwissProt* und *COG/EGGNOGs* durchgeführt. MG-RAST verwendet hingegen die nicht-redundanten Datenbanken *M5RNA* und *M5NR*. Auf diese Weise können gleichzeitig Sequenzhomologien der Datenbanken *SILVA*, *Greengenes* und *RDP* bzw. *GenBank*, *SEED*, *IMG*, *UniProt*, *KEGG* und *eggNOGs* abgefragt werden. Die Verwendung der nichtredundanten Datenbanken in MG-RAST stellt einen großen Vorteil im Vergleich zur Verwendung einzelner Datenbanken dar. Auf diese Weise wird der Informationsgehalt der erzielten Ergebnisse bei etwa gleichbleibender Analysezeit wesentlich vergrößert. Auch bietet MG-RAST die Möglichkeit, nicht-bakterielle Taxa auf funktionaler und taxonomischer Ebene darzustellen. Weiterhin ist eine spezifischere Darstellung von Funktionen und der funktionalen Taxonomie möglich, indem Subdatensätze in die *Workbench* geladen werden und zwischen den einzelnen Ontologien gewechselt werden kann. Bei der Verwendung der MG-RAST-Plattform über das *Online-Interface* hat der *User* jedoch keinen Einfluss auf die Aktualität der zugrundeliegenden Datenbanken und die Versionierung ist nicht ohne Weiteres nachvollziehbar. Im lokal betriebenen METRANS hingegen können die jeweils aktuellsten Datenbank-Versionen für die Analyse verwendet werden.

Um eine quantitative Vergleichbarkeit zwischen einzelnen Annotationen innerhalb eines Datensatzes oder zwischen einzelnen Datensätzen zu gewährleisten, bedarf es einer Normalisierung der Sequenzdaten. Im Rahmen einzelner Analysemöglichkeiten in MG-RAST, wie beispielsweise der *Heatmap*-Funktion, werden Möglichkeiten zur Normalisierung geboten. Grafiken normalisierter Datensätze können sehr einfach auf Knopfdruck erstellt werden. Allerdings ist für den Benutzer nicht immer ohne Weiteres nachvollziehbar, auf welche Weise die Normalisierung erfolgt ist und ob die gewählte Methode für den vorliegenden Datensatz geeignet ist. In METRANS erfolgt die Normalisierung von Sequenzabundanzanzen mithilfe des RPKM-Wertes, welcher sowohl die Größe des Datensatzes als auch die Genlänge der identifizierten Referenzsequenz berücksichtigt. Diese Methode ist in höchstem Maße speziell für die Analyse von Transkriptomik- und Metatranskriptomik-Datensätzen geeignet und stellt einen entscheidenden Vorteil der METRANS-Pipeline gegenüber MG-RAST dar.

Die Darstellung der Daten kann in beiden Pipelines sowohl in Tabellenform als auch in graphischer Form erfolgen. Dabei bietet MG-RAST deutlich mehr Darstellungsarten als METRANS und ermöglicht auch die Darstellung zuvor ausgewählter Subdatensätze. Jedoch sind nicht alle möglichen Darstellungen für die vorliegenden Datensätze und die betrachteten Eigenschaften sinnvoll. Daher wurde bei der Verwendung beider *Pipelines* meistens die Datenausgabe in Tabellenform gewählt und anschließend eine eigene graphische Darstellung umgesetzt. Ein weiterer bedeutender Aspekt bei der Darstellung und Interpretation der Datensätze ist der

Umgang mit Zuordnungen zu „falschen“ Datenbankeinträgen. Dabei handelt es sich meist um Zuordnungen, die biologisch nicht sinnvoll und/oder aus der Literatur nicht bekannt sind und auf einem fehlerhaften Eintrag innerhalb einer Datenbank beruhen. Dies ist beispielsweise bei den taxonomischen Zuordnungen zum Genus *Coptotermes* oder SP-beta-ähnliche Viren der Fall. MG-RAST bietet für dieses Problem keine Lösung an und überlässt die Interpretation der Daten vollständig dem User. METRANS hingegen besitzt eine Hinweisfunktion, die den User auf ungewöhnlich hohe *Read*-Abundanzen zu einer Zuordnung aufmerksam macht. Somit lassen sich falsch zugeordnete rRNA-Sequenzen in funktionalen Darstellungen besser erkennen.

Bei der Datenanalyse in den *Pipelines* können unabhängig vom analysierten Datensatz häufig bestimmte Muster beobachtet werden. So können beispielsweise Annotation zu den Staphylokokken mit deutlich höheren Sequenzidentitäten beschrieben werden als zu Corynebakterien. Die Ursache dieses Phänomens liegt wahrscheinlich in der deutlich erschwerteren Kultivierbarkeit von Corynebakterien im Vergleich zu Staphylokokken. Daher sind Corynebakterien-Sequenzen seltener in Datenbanken vertreten und es können keine Daten aus kultivierungsunabhängigen Studien verlässlich zugeordnet werden. Die Klassifizierungen auf Genus-Ebene mithilfe der rRNA-basierten Taxonomie sind größtenteils biologisch sinnvoll und stehen mit den Erkenntnissen aus der Literatur im Einklang. Auf Spezies-Ebene wird die Zuordnung jedoch in MG-RAST deutlich unzuverlässiger, da sich die wenigsten Sequenzen für die Zuordnung auf dieser taxonomischen Ebene eignen. METRANS verzichtet bei der rRNA-basierenden Taxonomie gänzlich auf die Spezies-Ebene, was formal korrekt ist. Die funktionale Taxonomie ist in beiden *Pipelines* deutlich diverser als die rRNA-basierte und enthält häufig Zuordnungen, die nicht im Einklang mit der Literatur stehen. Die Ursache liegt vermutlich darin, dass funktionale Sequenzen nicht die Biomarker-Eigenschaften der rRNA-Sequenzen besitzen und somit falsche Zuordnungen liefern können.

Bei der funktionalen Analyse stellt die Auswahl der Ontologie und der Betrachtungsebene eine wesentliche Leistung dar, die stets vom User erbracht werden muss. Häufig werden jedoch einzelne *Reads* auf höherer Ebene mehreren Stoffwechselwegen zugeordnet, da die jeweiligen Funktionen an mehreren Stoffwechselwegen beteiligt sind. Eine biologisch sinnvolle Zuordnung lässt sich dabei nicht immer treffen, wobei die Betrachtung der funktionalen Taxonomie aller beteiligten Funktionen eines Stoffwechselweges bereits einen Hinweis geben kann. Auch sind einzelne Taxa und Funktionen in den einzelnen funktionalen Datenbanken nicht gleich repräsentiert, was besonders in der funktionalen Charakterisierung in METRANS veranschaulicht wird. Es kann daher zu unterschiedlichen Zuordnungen einzelner Sequenzen in Abhängigkeit von der gewählten Ontologie kommen.

Insgesamt zeigt sich insbesondere aufgrund der Normalisierung mittels RPKM-Wert, dass die METRANS-Pipeline deutlich spezifischer auf Metatranskriptomik-Datensätze zugeschnitten ist als MG-RAST. MG-RAST bietet hingegen mit den nicht-redundanten Datenbanken *M5RNA* und *M5NR* und der *Workbench*-Funktion die Möglichkeit, die Datensätze aus den Perspektiven unterschiedlicher Ontologien zu betrachten, was sich besonders für detailliertere Analysen eignet. Weiterhin bietet MG-RAST eine bessere Handhabbarkeit für den User, weshalb diese Pipeline schließlich zur vergleichenden Analyse der Datensätze aus der Achselhöhle gewählt wurde. Für beide Pipelines gilt jedoch, dass sie lediglich die kollektive Darstellung der Sequenzdaten leisten können. Zur detaillierten Analyse einzelner Funktionen, wie beispielsweise der N<sup>α</sup>-Acylglutamin-Aminoacylase (N-AGA) oder der Cystathionin-β-Lyase, müssen die *Pipelines* verlassen werden und einzelne Sequenzen betrachtet werden. In diesem Fall dient MG-RAST jedoch zur Identifizierung und Auswahl von infrage kommenden Sequenzen.

**Fazit:**

- Für die Analyse von großen Metatranskriptomik-Datensätzen müssen Bioinformatik-Pipelines die folgenden Aufgaben bewältigen: Qualitätskontrolle, Datenreduktion, funktionale und taxonomische Zuordnung, Normalisierung und Darstellung. Alle Aufgaben werden in den beiden verwendeten *Analyse-Pipeline* METRANS und MG-RAST teils auf unterschiedliche Weise gelöst.
- Während METRANS insbesondere durch die Anwendung des RPKM-Wertes besser auf die Analyse von Metatranskriptomik-Datensätzen ausgelegt ist, bietet MG-RAST mit den nichtredundanten Datenbanken *M5NR* und *M5RNA* und mit der *Workbench*-Funktion die Möglichkeit, Daten differenzierter zu betrachten. Auch ermöglicht MG-RAST eine gute Handhabbarkeit für den User.
- Bei der Interpretation der Daten stellen die Kultivierbarkeit bestimmter Organismengruppen, die Eignung einzelner Sequenzabschnitte zur Zuordnung, die Auswahl geeigneter Ontologien und der Umgang mit fehlerhaften Zuordnungen wesentliche Aspekte dar.
- Beide Analyse-Pipelines bieten lediglich die kollektive Betrachtung von Datensätzen und müssen verlassen werden, sobald einzelne Funktionen detailliert auf Einzelsequenzebene analysiert werden sollen.

## 2 Metatranskriptomik der Mikrobiota des menschlichen Ohrs im Zusammenhang mit Gehörgangsentzündung

In dieser Arbeit wurde eine Probe eines 77-jährigen Otitis-externa-Patienten auf Metatranskriptom-Ebene untersucht. Diese Arbeit stellt die erste publizierte Metatranskriptom-Analyse des menschlichen Außenohrs im Zusammenhang mit Otitis externa dar. Im Folgenden wird die taxonomische Zusammensetzung und das funktionale Profil der aktiven Lebensgemeinschaft dieser Probe beschrieben. Weiterhin werden Funktionen, die im Zusammenhang mit Gehörgangsentzündung stehen, analysiert.

### 2.1 Taxonomische Betrachtung der aktiven Lebensgemeinschaft

Die taxonomische Darstellung der aktiven Mikrobiota der vorliegenden Probe belegt eindeutig die polymikrobielle Natur der Gehörgangsentzündung. Unabhängig von der verwendeten *Pipeline* und der gewählten Betrachtungsebene (rRNA-basiert oder funktional) werden bereits auf der Ebene des Phylums bakterielle und (je nach Methode) eukaryotische Organismen abgebildet, die bereits aus der Literatur im Zusammenhang mit Otitis externa bekannt sind. Auf Genus-Ebene zeigt sich in allen Fällen ein stark von Staphylokokken dominiertes Profil, von denen die Mehrheit auf Spezies-Ebene als *S. aureus* und z. T. *S. epidermidis* klassifiziert wird. Neben der Spezies *Pseudomonas aeruginosa*, welche in diesem Datensatz nicht erfasst wird, stellt *S. aureus* den häufigsten Verursacher der Gehörgangsentzündung dar (Lee et al., 2013; Rosenfeld et al., 2014; Schaefer und Baugh, 2012). Jedoch wurde auch *S. epidermidis* sowie weitere *Staphylococcus*-Spezies, wie *S. haemolyticus* im Zusammenhang mit Gehörgangsentzündung isoliert. Diese Spezies werden in dem vorliegenden Datensatz ebenfalls beschrieben, jedoch mit einer relativ geringen Abundanz, sodass anzunehmen ist, dass *S. aureus* das Hauptpathogen aus der Gruppe der Staphylokokken darstellt.

Eine weitere, in der vorliegenden Probe sehr abundante Gruppe von Pathogenen, welche ebenfalls im Zusammenhang mit Otitis externa bekannt sind, stellt der eukaryotische Genus *Candida* dar. Die Spezies *C. parapsilosis* und *C. albicans* werden im Einklang mit den Beschreibungen aus der Literatur ebenfalls im vorliegenden Datensatz dargestellt (Ahmad et al., 2007; Martin et al., 2005; McDowall, 1974; Pontes et al., 2009). Das zeitgleiche Vorkommen und die transkriptionelle Aktivität von *S. aureus* und *Candida* spp. weist auf eine mögliche Sekundärinfektion des Patienten durch *Candida* hin. Die Koisolation von *S. aureus* und *Candida*-Spezies ist häufig in der Literatur beschrieben und kommt besonders häufig im Zusammenhang mit Prothesen und Implantaten auf biomedizinischen Materialien vor. Bei dem Alter des Patienten wäre denkbar, dass auch das Tragen eines Hörgerätes die Infektion begünstigt. Zu

weiteren Grunderkrankungen des Patienten liegen jedoch keine Informationen vor. *Aspergillus*-Spezies, welche ebenfalls im Zusammenhang mit Otitis externa bekannt sind, werden hier nicht gefunden. Weitere Pilzspezies, die in sehr geringer Abundanz in der vorliegenden Probe zugeordnet werden, jedoch nicht aus der Literatur im Zusammenhang mit Gehörgangsentzündung bekannt sind, stellen *Lodderomyces elongisporus*, *Candida tropicalis*, *Meyerozyma guilliermondii*, *Scheffersomyces stipitis* und *Debaryomyces hansenii* dar. Ob es sich bei diesen Zuordnungen um medizinisch relevante Pathogene oder um falsch positive Zuordnungen der beiden dominanten *Candida*-Spezies handelt, ist nicht bekannt.

Darüber hinaus werden insbesondere in der funktionalen Taxonomie auch menschliche Sequenzen beschrieben. Weitere Zuordnungen aus der Gruppe der Wirbeltiere mit sehr geringen relativen Abundanzen stellen *Mus musculus*, *Macaca mulatta* und *Rattus norvegicus* dar. Aufgrund der Art und des Kontextes der Probe und der Probennahme ist es sehr wahrscheinlich, dass es sich hierbei um fehlerhafte Zuordnungen menschlicher RNA-Sequenzen handelt. Die Tatsache, dass die Zuordnungen zum Menschen vor allem auf funktionaler Ebene gemacht werden, legt nahe, dass es sich bei den entsprechenden RNA-Sequenzen um alternative Spleißvarianten funktionaler Gene gehandelt haben könnte, welche im entsprechenden Filterschritt in MG-RAST nicht entfernt wurden. Daher stellen diese Zuordnungen wahrscheinlich lediglich einen Teil aller Zuordnungen dar, die ohne den Filterschritt möglich gewesen wären. Dennoch ermöglichen diese Annotationen die Darstellung der transkriptionellen Aktivität des Wirtsorganismus innerhalb der aktiven Lebensgemeinschaft der Gehörgangsentzündung.

Innerhalb der aktiven Lebensgemeinschaft finden sich darüber hinaus Zuordnungen zu Bakterienarten, wie beispielsweise *Propionibacterium acnes*, *Fingoldia magna*, *Corynebacterium amycolatum* und weitere *Corynebacterium*- und *Staphylococcus*-Spezies, die auch als Mitglieder der gesunden Haut-Mikrobiota bekannt sind. Diese Organismen werden jedoch lediglich durch einen geringen Anteil der Sequenzdaten beschrieben, was auf eine Verschiebung der Mikrobiota hin zu den beiden pathogenen Spezies im Zusammenhang mit der Erkrankung hindeutet. Darüber hinaus wird die coryneforme Spezies *Turicella otitidis* mit einer relativ geringen Abundanz beschrieben. Dieser Organismus ist sehr spezifisch für den äußeren Gehörgang. Er stellt sowohl ein residentes Mitglied des gesunden Gehörganges dar, wurde jedoch auch bereits im Zusammenhang mit Otitis externa beschrieben (von Graevenitz und Funke, 2014). Wie genau die Wechselwirkung zwischen diesem Organismus und dem Wirtsorganismus in dieser Probe ist, lässt sich jedoch nicht beschreiben.

Insgesamt steht das für die vorliegende Probe beschriebene taxonomische Profil der aktiven Mikrobiota im Einklang mit den in der Fachliteratur publizierten Erkenntnissen und zeigt, dass sich die entwickelte Methode für die Beschreibung der aktiven Mikrobiota des Menschen gut eignet.

## 2.2 Funktionale Betrachtung der Mikrobiota des menschlichen Ohrs

Zur funktionalen Analyse der aktiven Mikrobiota der Gehörgangsentzündung wurde der vorliegende Metatranskriptomik-Datensatz zum einen allgemein beschrieben und zum anderen wurden spezifische Funktionen im Zusammenhang mit der Pathogenität der beteiligten Organismen analysiert. Die allgemeine Analyse erfolgte auf Basis der COG-Ontologie und zeigte, dass neun funktionale Gruppen, welche überwiegend dem Primärmetabolismus zugeordnet werden, mit insgesamt 80 % den Großteil des Datensatzes beschreiben. Lediglich die funktionale Klasse „Allgemeine Funktionsvorhersage“ stellte innerhalb dieser Gruppe eine unspezifische Beschreibung dar. Diese Analyse zeigte, dass sich die angewandte Methode zur Darstellung aller Funktionen einer Lebensgemeinschaft zu einem bestimmten Zeitpunkt eignet. Jedoch können auf diese Weise keine spezifischen Aussagen zur Pathogen-Wirt-Beziehung gemacht und ein Zusammenhang zu Otitis externa hergestellt werden. Daher wurden im nächsten Schritt Pathogenitätsfaktoren von *S. aureus* und *Candida*-Spezies sowie spezifische Abwehrmechanismen des Menschen analysiert.

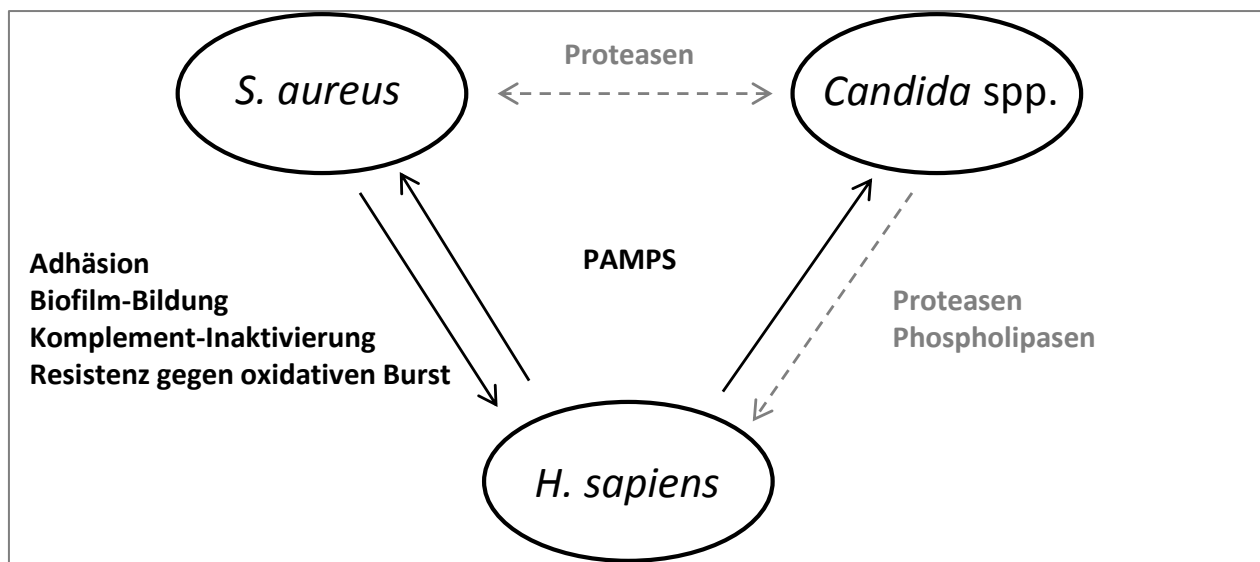


Abbildung 42: Schema der Pathogen-Wirt-Beziehungen der Ohrprobe im Zusammenhang mit Gehörgangsentzündung. Schwarze Pfeile: spezifische Faktoren; graue gestrichelte Pfeile: unspezifische Faktoren. Abkürzung PAMPS: Pathogen-associated molecular patterns.



Insgesamt wurden 29 unterschiedliche Pathogenitätsfaktoren von *S. aureus* beschrieben, wobei sechs Faktoren besonders dominant waren und im Zusammenhang mit der Umgehung der Immunantwort (ClfA und ClfB), Adhäsion und Biofilm-Bildung (Ebbs, Fibronectin-binding protein A, Serin-aspartate repeat protein C) und Resistenz gegen oxidativen Burst (CatA) stehen (s. Abbildung 42). Diese Funktionen zeigen sehr spezifische Pathogenitätsmechanismen von *S. aureus*, die allgemein für Infektionen mit diesem Keim bekannt sind und nun erstmals im Zusammenhang mit der Gehörgangsentzündung beschrieben werden. Für die Pathogene der *Candida*-Spezies sind mit Proteasen und Phospholipasen nur sehr allgemeine Pathogenitätsfaktoren bekannt, welche sich gegen den Wirtsorganismus richten. Diese Funktionen können für die vorliegende Probe in sehr hoher Abundanz für den Genus *Candida* beschrieben werden. Allerdings ist nicht eindeutig belegbar, ob und in welcher Weise alle identifizierten Funktionen an der Pathogenität beteiligt sind. Im Hinblick auf die Kooperation zwischen *S. aureus* und *Candida* spp. können Proteasen beschrieben werden, die im Allgemeinen die Adhäsion beider Organismen an Schleimhäuten und die Degradation der F<sub>c</sub>-Region von IgG ermöglichen (s. Abbildung 42). Auch diese Funktionen stellen sehr unspezifische Faktoren der Kooperation zwischen den beiden Organismen dar. Der spezifische Faktor Als3p, welcher eine Rolle bei der Ausbildung eines gemeinsamen Biofilms spielt, konnte nicht im Datensatz erfasst werden. Eine Erklärung hierfür wäre die Möglichkeit, dass es im Falle der Gehörgangsentzündung nicht zu einer Kooperation zwischen den beiden Organismen kommt, oder dass die Kooperation erst zu einem späteren Zeitpunkt als dem der Probennahme etabliert wird. Vergleichsproben zu einem früheren und späteren Zeitpunkt liegen jedoch nicht vor, sodass komparative Aussagen für diesen Datensatz nicht möglich sind. Die Sequenzdaten, die zum Menschen annotiert und nicht während der Analyse herausgefiltert wurden, zeigen mit *Toll-like* Rezeptoren und Lectin-Rezeptoren spezifische Funktionen, die der Erkennung von Pathogenen dienen und den Kreis innerhalb der Pathogen-Wirt-Interaktionen schließen (s. Abbildung 42).

Insgesamt kann in der vorliegenden Arbeit erstmals ein Metatranskriptomik-Datensatz gezeigt werden, der die Lebensgemeinschaft der Gehörgangsentzündung auf taxonomischer und funktionaler Ebene beschreibt. Dabei werden sowohl *S. aureus* und *Candida* spp. als abundante Pathogene beschrieben, als auch spezifische Pathogen-Wirt-Interaktionen auf Grundlage von spezifischen und unspezifischen Pathogenitätsfaktoren und Rezeptoren dargestellt.

**Fazit:**

- Die taxonomische Analyse des Metatranskriptomik-Datensatzes des Außenohrs zeigt ein polymikrobielles Profil, welches von *S. aureus* und *Candida* spp. dominiert wird und im Einklang mit den Beschreibungen aus der Fachliteratur steht.
- Die funktionale Analyse zeigt die Anwesenheit von Transkripten von *S. aureus*, welche im Zusammenhang mit Infektion, vor allem der Umgehung der Immunantwort, Adhäsion und Biofilm-Bildung und oxidativem Stress stehen. Unspezifische Pathogenitätsfaktoren konnten für *Candida*-Spezies beschrieben werden. Menschliche Komponenten der Immunantwort wurden ebenfalls beschrieben. Darüber hinaus wurden mögliche Interaktionen zwischen dem Wirt und den Pathogenen dargestellt.

### **3 Metatranskriptomik der menschlichen Achselhöhlen-Mikrobiota im Zusammenhang mit Körpergeruchsentstehung**

#### **3.1 Taxonomische Betrachtung der Achselhöhlen-Mikrobiota**

Die taxonomischen Profile der aktiven Lebensgemeinschaften der untersuchten Proben wurden zum einen rRNA-basiert und zum anderen auf der Ebene der funktionalen Taxonomie auf allen taxonomischen Ebenen dargestellt. Im allgemeinen wird die aktive Mikrobiota von Bakterien dominiert, was die Eignung der gewählten Methode zur Probennahme bestätigt. Auf rRNA-Ebene kann nahezu die gesamte aktive Lebensgemeinschaft durch die sechs Genera *Staphylococcus*, *Corynebacterium*, *Propionibacterium*, *Fingoldia*, *Anaerococcus* und *Peptoniphilus* beschrieben werden. Dies steht im Einklang mit DNA- und kultivierungsbasierten Daten aus der Literatur und zeigt, dass die gesamte Mikrobiota der Achselhöhle transkriptionell aktiv ist (Costello et al., 2009; Grice et al., 2009; Leyden et al., 1981; Taylor et al., 2003). Bei Betrachtung der Profile können drei Typen unterschieden werden: Lebensgemeinschaften, deren Aktivität durch die Staphylokokken (Typ I) oder Corynebakterien (Typ II) dominiert wird oder Lebensgemeinschaften, in denen sowohl Staphylokokken als auch Corynebakterien dominieren (Typ III). Insgesamt wird der Typ I jedoch am häufigsten beobachtet. Diese

Ergebnisse bestätigen ebenfalls Daten aus der Literatur, insbesondere die Daten von Taylor et al. (Taylor et al., 2003). Die Beobachtung von Troccaz et al. aus 16S-rDNA-Amplikon-Studien, dass Corynebakterien und Staphylokokken sich gegenseitig ausschließen kann in der Hauptkomponentenanalyse bestätigt werden. Auch eine Assoziation von Corynebakterien mit starkem Körpergeruch wird insbesondere durch das von Corynebakterien dominierte taxonomische Profil des Probanden 7 nahegelegt, welcher den stärksten innerhalb der Studie beobachteten Körpergeruch besitzt. Aber auch andere Proben von Lebensgemeinschaften mit hohem Corynebakterien-Anteil, wie beispielsweise die Proben 6n und 19v, stehen mit relativ starkem Körpergeruch in Zusammenhang. Eine ähnliche Korrelation zwischen Körpergeruch und Corynebakterien-Anteil analog zu den Daten von Taylor et al. lässt sich hier allerdings nicht darstellen. Zum einen ist der Stichprobenumfang mit nur 14 Proben sehr gering. Da einige dieser Proben von den selben Probanden genommen wurden, ist die Unabhängigkeit der Stichproben nicht gegeben. Darüber hinaus lassen sich Proben nach unterschiedlichen Anwendungen nur bedingt vergleichen. So führt die Deo-Anwendung häufig zu einer Reduktion der Keimzahl um einen Faktor 10-100, während die Wasser-Anwendung durch die Veränderung des Milieus der Hautoberfläche gegebenenfalls bestimmte Spezies oder Genera selektiert. Jedoch konnten in dieser Arbeit Korrelationen einzelner Spezies zum Körpergeruch beschrieben werden. Dabei korrelierten Corynebakterien in der Regel positiv und Staphylokokken in der Regel negativ, was erneut die Assoziation von Corynebakterien mit Körpergeruch bestätigt. Insgesamt wird in der vorliegenden Arbeit jedoch lediglich die Intensität, nicht jedoch die Qualität des Körpergeruches betrachtet. Zwar liegen experimentelle Daten auch zur Qualität vor, diese sind jedoch aufgrund des Studiendesigns nicht aussagekräftig. Es ist somit möglich, dass Corynebakterien ebenfalls mit einer anderen Qualität an Körpergeruch assoziiert sind als Staphylokokken. Dieses Phänomen wurde bereits von Leyden et al. beschrieben (Leyden et al., 1981).

Das gewählte Studiendesign und die Kombination an untersuchten Proben ermöglichen es hingegen, inter- und intraindividuelle Ähnlichkeiten zwischen einzelnen Probanden und in Abhängigkeit von der Zeit und Probennahmebedingung zu beschreiben. Insgesamt ist die intraindividuelle Ähnlichkeit der beiden Probanden 13 und 18, deren aktive Lebensgemeinschaften durch Staphylokokken dominiert sind, gegenüber denen der anderen Probanden stärker, da Corynebakterien-dominierte Lebensgemeinschaften eine größere intraindividuelle Variabilität aufweisen. Dies deutet darauf hin, dass Staphylokokken-dominierte Lebensgemeinschaften stabiler sind, möglicherweise, weil sie Veränderungen von Umweltbedingungen besser überdauern können als Corynebakterien. Dies steht im Einklang mit den *In-vitro*-Daten von Kwaszewska et al. die eine stärkere Resistenz von Staphylokokken im

Gegensatz zu den Corynebakterien zeigten (Kwaszewska et al., 2014). Im Falle der Proben des Probanden 19 ist eine zeitliche Veränderung der aktiven Lebensgemeinschaft von Typ II (Corynebakterien-dominiert) über Typ III (gemischt) hin zu Typ I (Staphylokokken-dominiert) zu beobachten. Dies zeigt, wie sehr sich die Zusammensetzung und Aktivität der mikrobiellen Lebensgemeinschaft eines Menschen ändern kann.

Auf Spezies-Ebene weisen Staphylokokken-dominierte Lebensgemeinschaften deutlich geringere Diversitäten und Äquitäten auf. Somit kann analog zu Beschreibungen aus der Literatur zu anderen Körperstellen angenommen werden, dass Staphylokokken den Lebensraum besetzen und die Ansiedlung anderer Spezies somit unterbinden (Cogen et al., 2008). Bei den dominanten *Staphylococcus*-Spezies handelt es sich um *S. epidermidis*, *S. aureus*, *S. hominis* und *S. haemolyticus*. Die dominanten Corynebakterien-Spezies sind *C. tuberculostearicum*, *C. amycolatum*, *C. striatum* (in einigen Proben) sowie *C. lipophiloflavum*. Diese Staphylokokken- und Corynebakterien-Spezies sind auch aus der Literatur bekannt. Insbesondere *C. striatum* und *C. tuberculostearicum* wurden mit Körpergeruch in Verbindung gebracht (Natsch et al., 2003; Troccaz et al., 2015). Insgesamt weisen die Sequenzen mit Homologie zu Staphylokokken eine deutlich höhere Sequenzidentität zu den Referenzsequenzen auf (i. d. R. 95 - 100 %) als Corynebakterien (häufig < 90 %). Dies hängt vermutlich mit der allgemein erschwerten Kultivierbarkeit von Corynebakterien zusammen. Es ist somit anzunehmen, dass die entsprechenden Sequenzen auf Spezies-Ebene falsch zugeordnet werden, da sie aus unkultivierten, bisher nicht beschriebenen Organismen stammen. Weiterhin werden *Propionibacterium*, *Fingoldia* und *Peptoniphilus*-Spezies beschrieben, die jedoch nicht nennenswert mit dem Körpergeruch korrelieren. Interessant sind hingegen die *Anaerococcus*-Spezies. Während *A. lactolyticus* und *A. prevotii* stark negativ mit Körpergeruch korrelieren, verhalten sich andere Spezies in der Korrelation unauffällig. Dies hebt die Bedeutung und Notwendigkeit der Klassifizierung der Mikrobiota auf Spezies-Ebene hervor. Insgesamt sind die Sequenzähnlichkeiten der diesen Gram-positiven anaeroben Kokken zugeordneten Sequenzen mit den jeweiligen Referenzsequenzen ebenfalls recht gering im Vergleich zu der Situation bei den Staphylokokken. Auch dies ist vermutlich der erschwerten Kultivierbarkeit und der geringen Repräsentation dieser Spezies in den Sequenzdatenbanken geschuldet.

Zusätzlich zu der Datenanalyse in MG-RAST wurden die taxonomischen Profile auf Grundlage von 16S-SSU-rRNA-Sequenzen und der *GenBank*-Datenbank als Referenz dargestellt. Bei der *GenBank*-Datenbank handelt es sich um die größte und aktuellste Datenbank, die auch Referenzsequenzen aus aktuellen Mikrobiom-Studien enthält. Diese Art der Analyse war in MG-RAST nicht möglich. In den auf diese Weise generierten Profilen zeigt sich, dass der Großteil der

Sequenzen aus bisher unkultivierten Organismen stammt. Ein weiterer Anteil der Sequenzen konnte gar nicht zugeordnet werden, vermutlich weil diese Sequenzen aus gänzlich unbekanntem Organismen stammen. Somit stammte mit einem Anteil von 74 % - 88 % der überwiegende Großteil der Daten aus weitestgehend unbekanntem Organismen. Lediglich der übrige Anteil der Sequenzen konnte auf Genus-Ebene zugeordnet werden und entsprach hinsichtlich seiner Zusammensetzung in etwa den in MG-RAST dargestellten Profilen. Dieses Ergebnis unterstreicht noch einmal die Bedeutung kultivierungsunabhängiger Techniken zur Beschreibung der menschlichen Mikrobiota und liefert gleichzeitig eine Erklärung für die z. T. unpezifischen Zuordnungen von Sequenzen auf funktionaler Ebene und auf der Ebene der funktionalen Taxonomie. Denn wenn ein Organismus nicht kultivierbar ist und lediglich auf Grundlage von DNA-Amplifikons beschrieben ist, befinden sich höchstwahrscheinlich keine funktionalen Informationen in den jeweiligen Referenzdatenbanken. Eine Beschreibung der funktionalen RNAs ist somit häufig nur durch relativ geringe Sequenzähnlichkeiten zu bekannten Funktionen aus kultivierten Organismen möglich.

**Fazit:**

- Die taxonomischen Profile der Lebensgemeinschaften der in dieser Arbeit untersuchten Proben der Achselhöhle stehen im Einklang mit kultivierungsbasierten und kultivierungsunabhängigen Daten aus der Literatur und deuten darauf hin, dass alle Mitglieder der Mikrobiota transkriptionell aktiv sind.
- Positive Korrelationen zwischen Corynebakterien-Spezies und negative Korrelationen zwischen Staphylokokken-Spezies und Körpergeruch können beschrieben werden.
- Staphylokokken-dominierte Lebensgemeinschaften scheinen stabiler und weniger divers zu sein, während Corynebakterien-dominierte Lebensgemeinschaften in Abhängigkeit von der Zeit und der Deo- bzw. Wasser-Behandlung stärker variieren.
- Auf Spezies-Ebene können aus der Literatur bekannte Organismen beschrieben werden. Bei den meisten Vertretern, insbesondere den Corynebakterien und Gram-positiven anaeroben Kokken handelt es sich jedoch größtenteils um bislang unbekannte und unkultivierte Organismen. Der Großteil der Sequenzdaten wird weitestgehend unbekanntem und/oder unkultiviertem Organismen zugeordnet.

### 3.2 Funktionale Betrachtung der Achselhöhlen-Mikrobiota

Die funktionale Analyse der Datensätze wurde zunächst allgemein auf Grundlage der COG-Ontologie durchgeführt. Dabei waren die Profile der Datensätze der Proben 13n und 7n besonders auffallend, da sie stark von den übrigen Profilen abwichen und lediglich sehr wenige Sequenzen zugeordnet werden konnten (Reduktion um einen Faktor von etwa 100 bzw. 500 im Vergleich zum Mittelwert). Die Ursache dieses Phänomens konnte nicht geklärt werden. Die RNA-Integrität und taxonomische Zusammensetzung auf Genus-Ebene scheinen jedoch nicht zu den Fehlerquellen zu gehören. Bei den dominanten Funktionen der übrigen Datensätze handelte es sich um Funktionsgruppen des Primärstoffwechsels mit den COG-Klassen Energiegewinnung und –umwandlung sowie Translation, ribosomale Struktur und Biogenese, während Funktionsgruppen, die im Zusammenhang mit eukaryotischen Systemen stehen, nur in sehr geringer Abundanz beschrieben wurden. Es konnten jedoch keine deutlichen Unterschiede (mit Ausnahme des Phänomens der Datensätze 7n und 13 n) zwischen den funktionalen Profilen von Proben unterschiedlicher Probanden, unterschiedlicher Probennahmetage oder in Abhängigkeit von der Behandlung mit Deo und Wasser beschrieben werden. Die Unterschiede auf funktionaler Ebene fallen somit deutlich geringer aus als auf taxonomischer Ebene. Diese Befunde stehen im Einklang mit der im Rahmen der Untersuchungen des HMP gemachten Beobachtungen, dass funktionale Unterschiede auf interindividueller und temporaler Ebene, aber auch bei Betrachtung unterschiedlicher Habitate äußerst gering sind. Daher erscheint auch für die Achselhöhle die vom HMP-Konsortium gewählte Erklärung sinnvoll, dass funktionale Unterschiede womöglich in Varianzen von Funktionen mit geringen Abundanzen begründet sind (Human Microbiome Project Consortium, 2012).

Zur Untersuchung von Unterschieden in den funktionalen Datensätzen wurde zusätzlich analog zu den taxonomischen Profilen eine Hauptkomponentenanalyse durchgeführt. Jedoch lieferte auch diese Analyse auf Grundlage einer Datentransformation kein deutliches *Clustering* von Datensätzen im Zusammenhang mit der Identität der Probanden, dem Probennahmezeitpunkt oder der Behandlung mit Deo oder Wasser. Jedoch ist im Falle der Datensätze der Probanden 18 und 19 eine Verschiebung der Proben im Zusammenhang mit Deo- bzw. Wasser-Anwendung zu beobachten. Die Verschiebung bei Deo-Anwendung steht dabei im Zusammenhang mit den funktionalen Gruppen Kohlenhydrate, Chaperone und allgemeine Funktionsvorhersage. Die Verschiebung bei Wasser-Anwendung steht mit den funktionalen Gruppen anorganische Ionen und unbekannte Funktionen. Dieses Phänomen ist auf taxonomischer Ebene in dieser Form nicht zu beobachten und legt somit eine Veränderung der Funktionen der Achselhöhlen-Mikrobiota auf RNA-Ebene nach Wasser-Behandlung nahe. Eine mögliche Interpretation dieses

Effektes wäre die schlagartige Verdünnung der in der Regel feuchten Achselhöhlen-Oberfläche und die damit einhergehende Herabsetzung der Tonizität und der verfügbaren Mikronährstoffe für die Mitglieder der Achselhöhlen-Mikrobiota. Die Tatsache, dass der Deo-Effekt für die Probanden 18 und 19 auf dieser funktionalen Betrachtungsebene weniger stark ausfällt als der Wassereffekt, könnte mit der allgemeinen antimikrobiellen Wirkung eines Deodorants zusammenhängen, die weniger selektiv auf die Mitglieder der Achselhöhlen-Mikrobiota einwirkt. Unterschiedliche Transkriptmengen von Genen, die Chaperone kodieren, stellen dabei womöglich eine Antwort auf den durch das Deo verursachten Stressreiz dar.

Aufgrund der begrenzten Aussagekraft der allgemeinen funktionalen Analyse, wurde zusätzlich eine spezifische Analyse auf Grundlage der in der Literatur beschriebenen Funktionen von *C. striatum* durchgeführt, welche im Zusammenhang mit Körpergeruchsentstehung stehen. Es wurden keine Sequenzen mit 100%iger Ähnlichkeit zu einer der bekannten Funktionen N<sup>α</sup>-Acylglutamin-Aminoacylase (N-AGA), Cystathionin-β-Lyase oder *Thiol Precursor Dipeptidase* (TpdA) identifiziert. In einigen Proben wurden zwar Sequenzähnlichkeiten beschrieben, diese liegen jedoch in der Regel unterhalb von 90 %. Im Zusammenhang mit der Beobachtung, dass auch auf taxonomischer Ebene Zuordnungen zu *C. striatum* eher selten sind, ist anzunehmen, dass die beschriebenen Funktionen einen anderen taxonomischen Ursprung haben. Der aus der Literatur bekannte Organismus scheint also im Falle der vorliegenden Proben keine bedeutende Rolle bei der Körpergeruchsentstehung zu spielen. Daher wurden im nächsten Schritt die bekannten Funktionen und ihre taxonomische Zuordnung näher untersucht.

Zur Beschreibung von Aminoacylase-Transkripten wurde zunächst die funktionale Taxonomie der übergeordneten unspezifischen Annotation COG1473 analysiert. Die taxonomischen Profile lieferten im Mittel 191 Annotationen pro Datensatz, welche auf Genus-Ebene hauptsächlich den Staphylokokken und Corynebakterien zugeordnet wurden, wobei die Profile im Wesentlichen Ähnlichkeiten zu den taxonomischen Profilen auf rRNA-Ebene aufwiesen und entsprechend entweder von Staphylokokken, Corynebakterien oder von beiden Genera dominiert wurden. Ein leicht negativer Korrelationskoeffizient von -0,37 zum *Malodor Score* legt nahe, dass die gewählte COG-Kategorie zu unspezifisch ist um Aminoacylase-Funktionen innerhalb der Datensätze zu beschreiben.

Im nächsten Analyseschritt wurde daher gezielt nach Annotationen zur Enzymfunktion Aminoacylase gesucht und die drei Annotationen identifiziert. Die Annotation „mögliche Aminoacylase“ wird dabei ausschließlich der Spezies *Anaerococcus lactolyticus* zugeordnet, was einen Zusammenhang zwischen der Arbeit von Fujii et al. darstellt, die eine Aminoacylase-Aktivität in einem Achselhöhlenisolat des Genus *Anaerococcus* beschreibt (Fujii et al., 2014).

Weiterhin weist die Gruppe der Sequenzen mit Annotation zur Aminoacylase unter anderem Homologien zu Sequenzen aus *A. lactolyticus* und *A. tetradicus* auf. Eine mögliche Aminoacylase-Aktivität durch *Anaerococcus lactolyticus* oder andere *Anaerococcus*-Spezies kann daher auf Basis der vorliegenden Daten untermauert werden. Darüber hinaus können viele Annotationen zu Aminoacylasen aus *S. epidermidis* beschrieben werden. Aminoacylasen aus Staphylokokken mit Beteiligung an der Körpergeruchsentstehung sind nicht aus der Literatur bekannt. Jedoch liegen diese Sequenzen in verhältnismäßig hoher Abundanz vor und könnten einen möglichen Erklärungsansatz für die Diskrepanz zwischen der hohen Menge an 3M2H und HMHA und der theoretisch sehr geringen Menge an N-AGA in menschlichem Schweiß liefern. Sowohl die Annotationen zur Aminoacylase als auch einige der Annotationen zur funktionalen Klasse COG1473 könnten Transkripte von Aminoacylasen mit Beteiligung an der Körpergeruchsentstehung darstellen. Annotationen zur N<sup>α</sup>-Acylglutamin-Aminoacylase werden ausschließlich der Spezies *C. striatum* zugeordnet und können nur in den Datensätzen der Proben des Probanden 7 und 19 beschrieben werden, welche auch auf rRNA-Ebene einen dominanten Anteil an Annotationen zu Corynebakterien aufweisen. Insgesamt können nur maximal 10 Sequenzen pro Datensatz mit dieser Zuordnung beschrieben werden. Dies ermöglichte die manuelle Untersuchung dieser Daten auf Einzelsequenzebene. In diesem Zusammenhang wurden die beiden Referenzsequenzen aus *Corynebacterium ureicelerivorans* (GenBank-ID WP\_038613795.1) und *Corynebacterium imitans* (GenBank-ID WP\_051904756.1) mit einer mittleren Sequenzidentität von 92,8 % identifiziert. Die Sequenzidentität zu der von Natsch et al. beschriebenen N-AGA liegt im Mittel bei lediglich 73,8 % und in den meisten Fällen unterhalb von 80 %. Jedoch können für die meisten Sequenzen die von Natsch et al. für N-AGA identifizierten Sequenzmotive MHACGHDXH (hier in der Form MHACGHDMH) und RADXDLPXXE (hier RADFDALPITE) beschrieben werden, was auf eine analoge Funktion der Genprodukte dieser Transkripte zu N-AGA schließen lässt. Insgesamt belegen die Daten, dass die corynebakterielle Aminoacylase-Aktivität in den hier untersuchten Proben nicht aus *C. striatum*, sondern von bisher nicht beschriebenen Corynebakterien stammt.

Die zugeordneten Sequenzen weisen eine sehr geringe Sequenzhomologie zu N-AGA auf und alignieren ausschließlich in der N-terminalen Zinkbindedomäne. Homologien zur C-terminalen Domäne können in den vorliegenden Datensätzen nicht beschrieben werden. Die Ursache könnte in einer falschen Zuordnung zu anderen Referenzsequenzen oder aufgrund einer Sequenzidentität unterhalb des Schwellenwertes in einer nicht erfolgten Zuordnung liegen. In beiden Fällen wäre davon auszugehen, dass die in den Proben vorliegenden Transkripte und die daraus abgeleiteten Proteinsequenzen insgesamt eine Sequenzhomologie kleiner 80 % zu N-AGA aufweisen. Somit wäre die Nutzung dieser Sequenzen in einem *In-vitro*-Assay zur Spaltung



von Glutaminkonjugaten nicht mehr durch das Patent von Natsch et al. geschützt und könnte zum Screening von aktiven Deosubstanzen zur Körpergeruchsverminderung genutzt werden (Natsch et al., 2005). Die beschriebenen konservierten Sequenzen liefern einen guten Ansatzpunkt zum Primerdesign zur Amplifikation der vollständigen Sequenzen aus genomischer DNA aus Proben der Achselhöhlen-Mikrobiota und zur Aufreinigung der entsprechenden Enzyme.

Weiterhin wurden in den vorliegenden Metatranskriptomik-Datensätzen Funktionen im Zusammenhang mit der Abspaltung geruchsbildender Sulfanylalkanole analysiert. Die Analyse von Annotationen zur von Starckenmann et al. (Starckenmann et al., 2005) beschriebenen Cystein-S-Konjugat- $\beta$ -Lyase lieferte lediglich zwei Sequenzen in einem der Datensätze, sodass davon auszugehen ist, dass diese Enzymfunktion keine bedeutende Rolle im Metatranskriptom der vorliegenden Proben der Achselhöhlen-Mikrobiota spielt und daher wahrscheinlich auch nicht an der Körpergeruchsentstehung beteiligt ist. Hingegen lieferte die Analyse von Annotationen zur von Natsch et al. beschriebenen Cystathionin- $\beta$ -Lyase im Mittel 47 *Read*-Abundanzen pro Datensatz. Diese Sequenzen wurden auf taxonomischer Ebene überwiegend Staphylokokken (*S. epidermidis*, *S. hominis*, *S. aureus* und *S. capitis*) und Corynebakterien (*C. lipophiloflavum*, *C. efficiens* und *C. striatum*) zugeordnet. Eine weiterführende Analyse der identifizierten Sequenzen in der CDD-Datenbank zeigte, dass die überwiegende Mehrheit dieser Sequenzen den funktionalen Gruppen MetC und MalY zugeordnet wurde. Dabei wurden corynebakterielle Sequenzen ausschließlich der Gruppe MalY und Staphylokokken-Sequenzen ausschließlich der Gruppe MetC zugeordnet, was im Einklang mit der Beschreibung von James et al. steht und diese zunächst bestätigt (James et al., 2013). Insgesamt konnte eine leicht negative Korrelation der  $\beta$ -Lysen mit dem *Malodor Score* beschrieben werden, was darauf hindeutet, dass die Beschreibung dieser Sequenzen auf Grundlage der gewählten Annotation zu unspezifisch ist.

Da es sich bei den funktionalen Cystathionin- $\beta$ -Lysen aus den Publikationen von Natsch et al. (Natsch et al., 2004) und James et al. (James et al., 2013) um Proteine der MalY-Gruppe handelt, wurde diese Gruppe im nächsten Schritt weiterführend analysiert. Auf taxonomischer Ebene wurde auch diese funktionale Gruppe überwiegend *Staphylococcus*- und *Corynebacterium*-Spezies zugeordnet und korrelierte leicht positiv mit dem *Malodor Score* ( $r^2 = 0,34$ ), was darauf hindeutet, dass zumindest einige der identifizierten Mitglieder dieser Gruppe an der Körpergeruchsentstehung beteiligt sein könnten. Unter den identifizierten corynebakteriellen Referenzsequenzen befand sich die von Natsch et al. beschriebene Cystathionin- $\beta$ -Lyase aus *C. striatum* sowie andere Funktionen mit den Annotation hypothetisches Protein, Plastocyanin und

Aminotransferase aus unterschiedlichen Corynebakterien-Spezies (Emter und Natsch, 2008). Bei diesen Referenzsequenzen handelt es sich um eine Gruppe hoch homologer Proteine mit dem katalytischen Lysinrest innerhalb einer stark konservierten Region. Lediglich zu den folgenden drei Referenzsequenzen konnten innerhalb der vorliegenden Metatranskriptomik-Datensätze Sequenzidentitäten von 100 % beschrieben werden: WP\_023020984.1 (*Corynebacterium* (Multispecies) *hypothetisches Protein*), WP\_005324171.1 (*C. pseudogenitalium*, Plastocyanin) und WP\_005326419.1 (*C. tuberculostearicum*, Plastocyanin). Daher ist anzunehmen, dass es sich bei den *in vivo* aktiven Proteinen mit Cystathionin- $\beta$ -Lyase-Aktivität nicht um die von Natsch et al. beschriebene Cystathionin- $\beta$ -Lyase aus *C. striatum* Ax 20 handelt, sondern um eine oder mehrere der drei im Rahmen dieser Analyse beschriebenen Proteine aus unterschiedlichen Corynebakterien-Spezies. Somit konnten in dieser Arbeit erstmals auf Grundlage von Daten aus *In-vivo*-Metatranskriptomik-Datensätzen Sequenzen beschrieben werden, welche an der Körpergeruchsentstehung durch Abspaltung von Sulfanylalkanolen beteiligt sein könnten. Ob tatsächlich eine Cystathionin- $\beta$ -Lyase-Aktivität der Genprodukte der identifizierten Sequenzen vorliegt, müsste in funktionalen Studien, beispielsweise mit heterolog exprimierten und aufgereinigten Proteinen getestet werden. Diese Arbeit liefert somit eine Grundlage für die Beschreibung von Proteinen und ggf. Substratspezifitäten, die bisher im Zusammenhang mit Körpergeruchsentstehung unbekannt waren.

Unter den Referenzsequenzen der MalY-Gruppe aus Staphylokokken wurde überwiegend die funktionale Annotation Aminotransferase aus unterschiedlichen Spezies identifiziert. Bei den Proteinen handelt es sich ebenfalls um eine hoch homologe Gruppe. Dabei konnte für den Großteil (elf) der beschriebenen Referenzsequenzen eine 100%ige Sequenzidentität innerhalb der Metatranskriptomik-Datensätze beschrieben werden. Diese Sequenzen stellen somit mögliche Kandidaten für Cystathionin- $\beta$ -Lyase-Aktivität aus Staphylokokken dar. Auch hier müsste die potentielle Beteiligung der entsprechenden Genprodukte zunächst *in vitro* verifiziert werden. Im positiven Fall können diese Sequenzen eine Verbindung zu den Arbeiten von Starkenmann et al. (Starkenmann et al., 2005) schaffen und evt. neue Substratspezifitäten und Kinetiken für die Freisetzung geruchsbildender Sulfanylalkanole mit Beteiligung an der Körpergeruchsentstehung liefern.

Die Analyse von Sequenzen mit Homologie zur ebenfalls von Natsch et al. (Emter und Natsch, 2008) beschriebenen TpdA-Dipeptidase lieferte Zuordnungen zu corynebakteriellen Referenzsequenzen. Jedoch wurden innerhalb der vorliegenden Datensätze keine Sequenzen mit einer 100 %igen Sequenzidentität beschrieben, sodass anzunehmen ist, dass es sich bei dem taxonomischen Ursprung dieser Sequenzen um bislang unbekannte Organismen, vermutlich

Corynebakterien, handelt. Die Analyse der unspezifischeren Annotation COG2195 lieferte größtenteils Referenzsequenzen aus Staphylokokken und Anaerokokken mit der funktionalen Annotation zur Peptidase T sowie die Annotation Aminoacyl-Histidin-Dipeptidase aus *Fingoldia magna*. Eine Überprüfung, ob die beschriebenen Sequenzen mit der Körpergeruchsentstehung in der Achselhöhle in Zusammenhang stehen, kann nur durch eine funktionale Charakterisierung der entsprechenden Enzyme erreicht werden.

Funktionen im Zusammenhang mit geruchsbildenden Steroiden wurden nicht identifiziert. Vermutlich laufen die entsprechenden Reaktionen nicht *in vivo* ab. Oder sie wurden aufgrund geringer Sequenzhomologie zu bekannten analogen Funktionen nicht identifiziert.

Sequenzen mit Homologie zur  $\beta$ -Glucuronidase und  $\gamma$ -Glutamyltranspeptidase werden verhältnismäßig selten identifiziert und zu Bakteriengattungen annotiert, die nicht im Zusammenhang mit der Achselhöhlen-Mikrobiota und Schweißgeruchsentstehung in der Literatur beschrieben sind. Es werden keine 100%igen Sequenzidentitäten beschrieben. Vermutlich handelt es sich bei dem Ursprung dieser Sequenzen um unkultivierbare Organismen, die bisher nicht beschrieben wurden. Aufgrund der sehr geringen *Read*-Abundanzen mit den entsprechenden Zuordnungen kann angenommen werden, dass die beiden Funktionen keine bedeutende Rolle bei der Körpergeruchsentstehung spielen.

Die Abundanzen und Zuordnungen der *Reads* mit Homologie zu Genen des Fettsäureabbaus sind sehr divers. Es werden zum einen starke Unterschiede in den beschriebenen *Read*-Abundanzen beobachtet. Zum anderen werden die Funktionen unterschiedlichen Bakteriengattungen zugeordnet. Neben Annotationen zu Corynebakterien, Staphylokokken und Anaerokokken werden auch Annotationen zu anderen Bakteriengattungen beschrieben, die nicht im Zusammenhang mit der Achselhöhlen-Mikrobiota bekannt sind. Dies lässt ebenfalls auf eine unbekannt taxonomische Herkunft dieser Sequenzen schließen. Bei Betrachtung der beschriebenen Zuordnungen ist anzunehmen, dass die beschriebenen Enzymfunktionen nicht Teil des Fettsäureabbauweges in einzelnen Organismen sind. Da die am Fettsäureabbau beteiligten Enzyme auch an anderen Stoffwechselwegen beteiligt sind, ist anzunehmen, dass die beschriebenen Sequenzen andere biologische Funktionen erfüllen. Auf Grundlage dieser Beschreibung kann angenommen werden, dass der Abbau langkettiger verzweigter Fettsäuren durch den Stoffwechselweg der  $\beta$ -Oxidation in Corynebakterien keine bedeutende Rolle bei der Körpergeruchsentstehung spielt. Dies bestätigt die letzte Hypothese von James et al. (James et al., 2013).

Die von James et al. aufgestellte Hypothese der Beteiligung von BCAA-Aminotransferasen aus Staphylokokken durch den Abbau verzweigtkettiger Aminosäuren wurde ebenfalls überprüft. Es wurden überwiegend Zuordnungen zu BCAA-Aminotransferasen aus Staphylokokken in den vorliegenden Datensätzen beschrieben. In Datensätzen, die auf rRNA-Ebene auch von corynebakteriellen Sequenzen dominiert wurden, konnten auch Zuordnungen zu BCAA-Aminotransferasen aus Corynebakterien beschrieben werden. Insbesondere im Datensatz 7n des Probanden 7 mit dem stärksten Körpergeruch waren diese besonders dominant. Auch Annotationen zu BCAA-Aminotransferasen aus anderen Organismen konnten beschrieben werden, die Abundanzen sind jedoch im Vergleich zu den Zuordnungen zu Staphylokokken sehr gering. Eine mögliche Beteiligung dieser Enzymfunktionen an der Körpergeruchsentstehung verursacht durch Staphylokokken ist somit denkbar. In der Korrelationsanalyse einzelner Enzymfunktionen mit dem Körpergeruch zeigt sich jedoch eine leicht negative Korrelation ( $r^2 = 0,25$ ). Somit ist zu vermuten, dass die mögliche Beteiligung dieser Enzymfunktion an der Entstehung von Körpergeruch eher indirekter Natur ist.

Die *Read*-Abundanzen mit Sequenzhomologie zu L-Laktat-Dehydrogenase und Acetat-Kinase sind im Vergleich zu denen der zuvor betrachteten Funktionen sehr hoch, und deuten auf eine große Bedeutung alternativer Fermentationswege in der Achselhöhlen-Mikrobiota hin. Die jeweiligen Sequenzen werden überwiegend den Staphylokokken zugeordnet. Weiterhin sind in geringerem Maße Zuordnungen zu Corynebakterien, Anaerokokken und im Falle der Laktat-Dehydrogenase zu Propionibakterien beschrieben. Die Annahme, dass mikroaerophile Staphylokokken und Propionibakterien zur Entstehung von Körpergeruch durch Bildung von Acetat und Propionat beitragen, kann somit untermauert werden. Weiterhin ist die Beteiligung von Corynebakterien und Anaerokokken möglich. Eine direkte Proportionalität der *Read*-Abundanzen zum *Malodor Score* ist jedoch nicht erkennbar. Möglicherweise bilden diese beiden Substanzen gewissermaßen einen Grundgeruch in der Achselhöhle und werden z. T. auch nach der Bildung weiter verstoffwechselt.

Zur Beschreibung von Funktionen, die bisher nicht im Zusammenhang mit Körpergeruchsentstehung beschrieben sind, wurde eine Korrelationsanalyse aller Funktionen in den Metatranskriptomik-Datensätzen mit dem *Malodor Score* durchgeführt und graphisch auf die KEGG-Stoffwechselkarten gemappt. Von den 2359 beschriebenen Funktionen korrelierten jeweils 11 % positiv bzw. negativ mit dem Körpergeruch. Auffallend war, dass die 10 am stärksten positiv korrelierten Funktionen überwiegend den Corynebakterien und die 10 am stärksten negativ korrelierten Funktionen überwiegend den Staphylokokken zugeordnet wurden. Dies steht im Einklang mit den Beschreibungen aus der Literatur und mit den

taxonomischen Beschreibungen aus dieser Arbeit, dass eine Corynebakterien-dominierte Mikrobiota mit starkem Körpergeruch assoziiert ist (Leyden et al., 1981; Taylor et al., 2003; Troccaz et al., 2015). Besonders stark positiv korrelierten dabei Funktionen im Zusammenhang mit Nährstoffaufnahme (Hämoxigenase und Sialidase) sowie die beiden Funktionen aus dem Tryptophan-Stoffwechsel, TrpCF und TrpD. Keiner der beteiligten Metabolite ist im Zusammenhang mit Körpergeruch beschrieben worden, sodass die entsprechenden Enzymfunktionen vermutlich nicht direkt mit der Bildung körpergeruchsverursachender Substanzen durch Corynebakterien in Zusammenhang stehen. Besonders hervorzuheben ist jedoch die positive Korrelation der Enoyl-CoA-Hydratase mit dem *Malodor Score*. Das Enzym ist durch die Hydratisierung der Doppelbindung zwischen dem C2- und C3-Atom in Hydroxysäuren an vielen katabolen Stoffwechselwegen beteiligt. In den Arbeiten von Natsch et al. sind diverse 3-Hydroxysäuren als Spaltungsprodukte des Enzyms N-AGA aus *C. striatum* beschrieben worden (Natsch et al., 2006). Darunter befinden sich die besonders häufig beschriebenen Komponenten 3-Hydroxy-3-Methylhexansäure (HMHA), 3-Hydroxy-4-Methylhexansäure, sowie die entsprechenden Heptan-, Octan-, Nonan- und Decansäuren (Natsch et al., 2006). Darüber hinaus wird die Funktion in der graphischen Darstellung auf den Leucin- Valin- und Isoleucin-Abbauweg gemappt. Bei dem Produkt der entsprechenden Enzymreaktion würde es sich um 3-Hydroxyisovaleriansäure handeln. Jedoch ist aufgrund der taxonomischen Zuordnung der anderen Funktionen zu diesem Stoffwechselweg anzunehmen, dass die Enoyl-CoA-Hydratase an einem anderen Stoffwechselweg mit einem anderen Produkt beteiligt ist. Ob und falls ja, wie genau diese Enzymfunktion mit der Körpergeruchsentstehung im Zusammenhang stehen könnte, bleibt zunächst ungeklärt, da der Ursprung der glutaminkonjugierten Vorläufer in Schweiß selber vermutet wird. Die mit dem Körpergeruch stark negativ korrelierten Funktionen aus Staphylokokken sind diversen Stoffwechselvorgängen zuzuordnen. Dies deutet darauf hin, dass der negative Korrelation der Staphylokokken mit Körpergeruch keine spezifischen molekularen Mechanismen zugrunde liegen, sondern dass die stabile Besiedlung des Lebensraumes Achselhöhle durch Staphylokokken die Besiedlung durch andere körpergeruchsverursachende Bakterien verhindert.

Bei der graphischen Darstellung wurde insbesondere der Abbauweg der verzweigt-kettigen Fettsäuren Valin, Leucin und Isoleucin analysiert, welcher ausgehend von den jeweiligen Aminosäuren in vier Abbauschritten, die jeweils durch die gleichen Enzyme katalysiert werden, zur Bildung verzweigt-kettiger volatiler Fettsäuren durch Staphylokokken führt, deren Beteiligung am Körpergeruch bereits in der Literatur beschrieben und von James et al. postuliert wurde (James et al., 2013). In den hier vorliegenden Datensätzen werden die vier ersten Schritte im Abbauweg tatsächlich den Staphylokokken zugeordnet, was im Einklang mit der Hypothese

steht. Neben der bereits beschriebenen BCAA-Aminotransferase katalysiert die Leucin-Dehydrogenase die Umwandlung von Leucin zu 4-Methyl-2-Oxopentanonsäure und korreliert positiv mit dem *Malodor Score*. 4-Methyl-2-Oxopentanonsäure wurde ebenfalls von Natsch et al. als Spaltungsprodukt des Enzyms N-AGA beschrieben und es ist bekannt, dass diese Substanz stinkt, wobei eine genauere sensorische Charakterisierung in der Literatur nicht verfügbar ist. Die Tatsache, dass die nächsten drei Schritte des Abbauweges negativ mit dem *Malodor Score* korrelieren deutet zum einen auf eine mögliche Beteiligung sowie auf eine mögliche bisher im Zusammenhang mit Körpergeruch unbekannt Quelle dieser Substanz hin. Zum anderen sprechen diese Daten gegen den Ursprung verzweigtkettiger Fettsäuren mit Beteiligung an Körpergeruchsentstehung wie Isovaleriansäure aus dem Leucin-Abbauweg von Staphylokokken. Jedoch ist zu beachten, dass bei der Bestimmung des *Malodor Score* lediglich ein quantitatives Maß des Körpergeruches festgelegt wird. Daten zur Qualität sind zwar auch vorhanden, besitzen jedoch wenig Aussagekraft und Konsistenz, weswegen sie hier nicht weiter betrachtet werden. Grundsätzlich wäre es jedoch auch denkbar, dass Staphylokokken-dominierte Lebensgemeinschaften einen qualitativ anderen und quantitativ eher schwächeren Körpergeruch erzeugen als durch Corynebakterien dominierte Lebensgemeinschaften. Eine ähnliche Beobachtung ist bereits aus der Literatur bekannt (Leyden et al., 1981). Eine weiterführende Überprüfung der Beteiligung der einzelnen Enzyme des Leucin-Abbauweges an der Körpergeruchsentstehung ist daher notwendig.

Weiterhin wurde der Glutathion-Abbauweg detailliert analysiert. Die Tatsache, dass die einzelnen Funktionen z. T. unterschiedlichen Organismen zugeordnet werden, lässt jedoch darauf schließen, dass sie nicht im gleichen Organismus ablaufen. Eine Beteiligung dieses Stoffwechselweges an der Körpergeruchsentstehung ist daher sehr unwahrscheinlich. Einen weiteren interessanten Stoffwechselweg stellt der Phenylalanin-Metabolismus dar. Insbesondere wurde hier die Umwandlung von Phenylacetaldehyd zu Phenylacetat durch die corynebakterielle Phenylacetaldehyd-Dehydrogenase betrachtet. Phenylacetat kann weiterführend zu 4-Hydroxy-Phenylacetat umgewandelt werden. Beide Substanzen wurden von Natsch et al. ebenfalls als Spaltungsprodukte des Enzyms N-AGA beschrieben (Natsch et al., 2006). Ein Zusammenhang zwischen der direkten Bildung dieser Substanzen durch Corynebakterien und Körpergeruch müsste jedoch in weiteren Experimenten überprüft werden.

Insgesamt wurde im Rahmen dieser Arbeit die Achselhöhlen-Mikrobiota auf Transkriptomebene betrachtet. Zuverlässige Aussagen zur Situation auf Proteom- und Metabolomebene sind nicht möglich. Jedoch können die beschriebenen funktionalen Profile Hinweise auf Transkripte mit Beteiligung an der Körpergeruchsentstehung geben.

**Fazit:**

- Auf funktionaler Ebene sind intraindividuelle Ähnlichkeiten deutlich schwächer ausgeprägt als auf taxonomischer. Dominante Funktionen stehen im Zusammenhang mit dem Primärstoffwechsel. Die Wasser-Anwendung löst vermutlich aufgrund der raschen Änderung osmotischer Bedingungen stärkere Veränderungen in den funktionalen Profilen aus als die Deo-Anwendung.
- Es wurden keine 100%igen Sequenzidentitäten zu den von Natsch et al. beschriebenen Enzymfunktionen N<sup>α</sup>-Acylglutamin-Aminoacylase (N-AGA), Cystathionin-β-Lyase oder *Thiol Precursor Dipeptidase* (TpdA) aus *C. striatum* identifiziert. Die jeweiligen Enzymfunktionen stammen daher wahrscheinlich aus unkultivierten Corynebakterien-Spezies.
- Sequenzidentitäten zu Aminoacylasen konnten für die Genera *Corynebacterium*, *Staphylococcus* und *Anaerococcus* beschrieben werden und liefern möglicherweise eine Erklärung für die Diskrepanz zwischen der Menge geruchsbildender Substanzen und theoretisch vorliegenden Enzymmengen mit Aminoacylase-Funktion. Sequenzhomologien zur von Natsch et al. beschriebenen N<sup>α</sup>-Acylglutamin-Aminoacylase liegen in der Regel unterhalb von 80 % und stammen vermutlich aus bisher unbekanntem Corynebakterien.
- Es wurden MalY-Sequenzen aus Corynebakterien und Staphylokokken mit möglicher Cystathionin-β-Lyase-Funktion erstmals beschrieben.
- Funktionen im Zusammenhang mit geruchsbildenden Steroiden, β-Glucuronidasen und γ-Glutamyltranspeptidasen wurden nicht in den Datensätzen identifiziert und spielen wahrscheinlich keine Rolle bei der Körpergeruchsentstehung.
- Die β-Oxidation spielt ebenfalls keine bedeutende Rolle in den funktionalen Profilen der hier untersuchten Proben und ist somit wahrscheinlich nicht an der Körpergeruchsentstehung beteiligt.
- BCAA-Aminotransferasen aus Staphylokokken sowie L-Laktat-Dehydrogenase und Acetat-Kinase wurden in hohen Abundanzen beschrieben und könnten an der Körpergeruchsentstehung beteiligt sein.
- Corynebakterielle Funktionen, insbesondere die Enoyl-CoA-Hydratase, korrelieren positiv und Staphylokokken-Funktionen negativ mit dem Körpergeruch.
- Es konnten Enzymfunktionen mit möglicher Beteiligung an der Körpergeruchsentstehung beschrieben werden, welche bisher nicht bekannt waren.

## VI. Ausblick

Auf der Grundlage der schnellen Entwicklung von Sequenzieretechniken ermöglichen die beiden Methoden der Metagenomik und der Metatranskriptomik die Beschreibung ganzer Lebensgemeinschaften. Die Detektion einzelner Gene bzw. RNAs erfolgt dabei *de novo* und die Genauigkeit der Auflösung liegt auf der Ebene einzelner Basen. Da mit der Metatranskriptomik die aktiven Mitglieder und Funktionen einer Lebensgemeinschaft abgebildet werden, stellt diese Technik bereits einen gezielteren Ansatz dar als die Metagenomik.

In der vorliegenden Arbeit konnte die Anwendung der Metatranskriptomik zur Beschreibung der menschlichen Mikrobiota demonstriert werden. Dabei konnten sowohl die Zusammensetzung der Lebensgemeinschaften und die aktiven Funktionen allgemein, als auch spezifische Funktionen im Zusammenhang mit dem beschriebenen Habitat dargestellt werden. Auf diese Weise können Krankheiten und andere phänotypische Eigenschaften beschrieben und auf molekularer Ebene analysiert werden. Komparative Ansätze bieten dabei die Möglichkeit, die Aktivität im zeitlichen oder örtlichen Verlauf zu analysieren und auf inter- oder intraindividuelle Ebene abzubilden. Auf diese Weise können Korrelationen zwischen einzelnen Organismen oder Funktionen und dem Phänotyp beschrieben werden. Die Darstellung von Kausalitäten ist jedoch Gegenstand weiterführender gezielter Analysen.

Den stärksten limitierenden Faktor für die Metatranskriptom-Analyse menschlicher Lebensgemeinschaften stellen derzeit noch die geringen Probenmengen dar, die sich aus solchen Habitaten gewinnen lassen. Vor dem Hintergrund der enormen Weiterentwicklung der Einzelmolekülsequenzierung sowie molekularer Methoden zur Aufarbeitung der RNA für die Sequenzierung ist jedoch anzunehmen, dass diese Limitierung schon bald weitestgehend überwunden sein wird.

Die Erzeugung immer größerer Datenmengen wird auch in Zukunft die Entwicklung effizienter *Analyse-Pipelines* zur Datenverarbeitung erfordern und vorantreiben. Dabei wird eine stärkere Differenzierung der Analysestrategien in Abhängigkeit von den Datensätzen erforderlich sein. Ein modularer Aufbau von *Analyse-Pipelines*, bei denen der Anwender nicht nur einzelne Parameter einstellen, sondern auch ganze Analyseverläufe in Abhängigkeit vom Datensatz zusammenstellen könnte, würde die notwendige Flexibilität für die Analyse diverser Datensätze gewährleisten. So würde sich eine *Analyse-Pipeline*, in der der Sequenzabgleich mit der *M5NR*-Datenbank aus MG-RAST sowie die Darstellung der Daten mittels RPKM-Wert analog zu METRANS implementiert wäre, die optimale Beschreibung von *Shotgun*-Metatranskriptomik-Datensätzen, wie sie im Rahmen dieser Arbeit generiert wurden, ermöglichen. Gezieltere Ansätze,



---

beispielsweise auf Grundlage ausgewählter Biomarker oder Zielsequenzen, würden mit weniger komplexen *Pipelines* analysiert werden können, was den Vorteil einer Verkürzung der Analysezeit hätte.

Eine weitere Herausforderung stellt die Interpretation der Datensätze dar. Zwar erfolgt die Detektion einzelner Sequenzen *de novo*, ihre taxonomische und funktionale Zuordnung basiert jedoch auf der Beschreibung von Ähnlichkeiten zu bekannten Sequenzen. Diese stammen jedoch größtenteils aus kultivierungsbasierten Datensätzen und sind daher nur bedingt in der Lage, die Beschreibung von Daten nichtkultivierbarer Organismen zu ermöglichen. Da die kultivierungsunabhängige Beschreibung von Organismen und Funktionen jedoch parallel zur Entwicklung der „Meta-omik“-Technologien voranschreitet, werden in Zukunft zunehmend Daten aus unkultivierten Organismen zur Verfügung stehen und die Analyse neuer Datensätze erleichtern. Bei der Beschreibung von Kausalitäten zwischen einzelnen Funktionen der Mikrobiota und dem Phänotyp werden jedoch auch weiterhin funktionale Analysen unumgänglich sein. Auf diese Weise werden funktionale Annotationen spezifischer.

Die Metatranskriptomik kann als Basis zur Identifizierung relevanter Funktionen oder, wie in dieser Arbeit demonstriert, als Methode zur Beschreibung relevanter Funktionen aus unbekanntem Organismen angewandt werden. So ist beispielsweise die Beschreibung bisher unbekannter N<sup>α</sup>-Acylglutamin-Aminoacylasen aus unbekanntem Organismen geeignet, die molekularen Mechanismen der Körpergeruchsentstehung besser zu verstehen und entsprechende Wirksubstanzen dagegen zu entwickeln. Aber auch die Beschreibung von MALY-Sequenzen mit abweichenden funktionalen Annotationen zeigt, dass molekulare Vorgänge der menschlichen Mikrobiota mithilfe der Metatranskriptomik besser verstanden werden können. Die Erforschung molekularer Mechanismen der Pathogenität bei Infektionserkrankungen stellt ein weiteres Anwendungsfeld der Metatranskriptomik dar. Dabei können Interaktionen zwischen dem Wirtsorganismus und den Pathogenen sowie zwischen den Mitgliedern der Lebensgemeinschaft dargelegt werden.

Insgesamt ermöglicht die Metatranskriptomik der menschlichen Mikrobiota umfassende und detaillierte Einblicke in die Zusammensetzung und Funktion diverser Habitats des Supraorganismus des Menschen. Die daraus gewonnenen Erkenntnisse können sowohl in der Grundlagenforschung als auch in für die Entwicklung kosmetischer oder pharmazeutischer Wirksubstanzen genutzt werden.

## VII. Literaturverzeichnis

Ahmad, N., Etheridge, C., Farrington, M. und Baguley, D. M.: Prospective study of the microbiological flora of hearing aid moulds and the efficacy of current cleaning techniques, *J. Laryngol. Otol.*, 121(2), 110–113, doi:10.1017/S0022215106002222, 2007.

Akiba, S., Arai, N., Kusuoku, H., Takagi, Y., Hagura, T., Takeuchi, K. und Fuji, A.: The N-terminal amino acid of apolipoprotein D is putatively covalently bound to 3-hydroxy-3-methyl hexanoic acid, a key odour compound in axillary sweat, *Int. J. Cosmet. Sci.*, 33(3), 283–286, doi:10.1111/j.1468-2494.2010.00636.x, 2011.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. und Lipman, D. J.: Basic local alignment search tool, *J. Mol. Biol.*, 215(3), 403–410, doi:10.1016/S0022-2836(05)80360-2, 1990.

Arsović, N. A., Banko, A. V., Dimitrijević, M. V., Djordjević, V. Z., Milovanović, J. P. und Arsenijević, V. A.: Protease activities of *Candida* spp. isolated from otitis externa: preliminary result, *Acta Chir. Iugosl.*, 56(3), 113–116, 2009.

Austin, C. und Ellis, J.: Microbial pathways leading to steroidal malodour in the axilla, *J. Steroid Biochem. Mol. Biol.*, 87(1), 105–110, 2003.

Baker, B. J. und Banfield, J. F.: Microbial communities in acid mine drainage, *FEMS Microbiol. Ecol.*, 44(2), 139–152, doi:10.1016/S0168-6496(03)00028-X, 2003.

Barzantny, H., Brune, I. und Tauch, A.: Molecular basis of human body odour formation: insights deduced from corynebacterial genome sequences, *Int. J. Cosmet. Sci.*, 34(1), 2–11, doi:10.1111/j.1468-2494.2011.00669.x, 2012a.

Barzantny, H., Schröder, J., Strotmeier, J., Fredrich, E., Brune, I. und Tauch, A.: The transcriptional regulatory network of *Corynebacterium jeikeium* K411 and its interaction with metabolic routes contributing to human body odor formation, *J. Biotechnol.*, 159(3), 235–248, doi:10.1016/j.jbiotec.2012.01.021, 2012b.

Barzantny, H., Guttman, S., Lässig, C., Brune, I. und Tauch, A.: Transcriptional control of lipid metabolism by the MarR-like regulator FamR and the global regulator GlxR in the lipophilic axilla isolate *Corynebacterium jeikeium* K411, *Microb. Biotechnol.*, 6(2), 118–130, doi:10.1111/1751-7915.12004, 2013.

Baumann, T., Bergmann, S., Schmidt-Rose, T., Max, H., Martin, A., Enthaler, B., Terstegen, L., Schweiger, D., Kalbacher, H., Wenck, H., Jedlitschky, G. und Jovanovic, Z.: Glutathione-conjugated sulfanylalkanols are substrates for ABCC11 and  $\gamma$ -glutamyl transferase 1: a potential new pathway for the formation of odorant precursors in the apocrine sweat gland, *Exp. Dermatol.*, 23(4), 247–252, doi:10.1111/exd.12354, 2014.

Beers, S. L. und Abramo, T. J.: Otitis externa review, *Pediatr. Emerg. Care*, 20(4), 250–256, 2004.

- Boeckmann, B., Blatter, M.-C., Famiglietti, L., Hinz, U., Lane, L., Roechert, B. und Bairoch, A.: Protein variety und functional diversity: Swiss-Prot annotation in its biological context, *C. R. Biol.*, 328(10–11), 882–899, doi:10.1016/j.crv.2005.06.001, 2005.
- Bouslimani, A., Porto, C., Rath, C. M., Wang, M., Guo, Y., Gonzalez, A., Berg-Lyon, D., Ackermann, G., Moeller Christensen, G. J., Nakatsuji, T., Zhang, L., Borkowski, A. W., Meehan, M. J., Dorrestein, K., Gallo, R. L., Bandeira, N., Knight, R., Alexandrov, T. und Dorrestein, P. C.: Molecular cartography of the human skin surface in 3D, *Proc. Natl. Acad. Sci. U. S. A.*, 112(17), E2120-2129, doi:10.1073/pnas.1424409112, 2015.
- Brown, S., Santa Maria, J. P. und Walker, S.: Wall Teichoic Acids of Gram-Positive Bacteria, *Annu. Rev. Microbiol.*, 67, doi:10.1146/annurev-micro-092412-155620, 2013.
- Brune, I., Barzantny, H., Klötzl, M., Jones, J., James, G. und Tauch, A.: Identification of McbR as transcription regulator of *aecD* and genes involved in methionine and cysteine biosynthesis in *Corynebacterium jeikeium* K411, *J. Biotechnol.*, 151(1), 22–29, doi:10.1016/j.jbiotec.2010.11.005, 2011.
- Callewaert, C., Kerckhof, F.-M., Granitsiotis, M. S., Van Gele, M., Van de Wiele, T. und Boon, N.: Characterization of *Staphylococcus* and *Corynebacterium* clusters in the human axillary region, *PloS One*, 8(8), e70538, doi:10.1371/journal.pone.0070538, 2013.
- Callewaert, C., Hutapea, P., Van de Wiele, T. und Boon, N.: Deodorants and antiperspirants affect the axillary bacterial community, *Arch. Dermatol. Res.*, 306(8), 701–710, doi:10.1007/s00403-014-1487-1, 2014.
- Cavaggioni, A. und Mucignat-Caretta, C.: Major urinary proteins, alpha(2U)-globulins and aphrodisin, *Biochim. Biophys. Acta*, 1482(1–2), 218–228, 2000.
- Centers for Disease Control and Prevention (CDC): Estimated burden of acute otitis externa--United States, 2003-2007, *MMWR Morb. Mortal. Wkly. Rep.*, 60(19), 605–609, 2011.
- Chavakis, T., Preissner, K. T. und Herrmann, M.: The anti-inflammatory activities of *Staphylococcus aureus*, *Trends Immunol.*, 28(9), 408–418, doi:10.1016/j.it.2007.07.002, 2007.
- Chiller, K., Selkin, B. A. und Murakawa, G. J.: Skin microflora and bacterial infections of the skin, *J. Investig. Dermatol. Symp. Proc. Soc. Investig. Dermatol. Inc Eur. Soc. Dermatol. Res.*, 6(3), 170–174, doi:10.1046/j.0022-202x.2001.00043.x, 2001.
- Christina Ander: Bioinformatic methods for the analysis and comparison of metagenomes and metatranscriptomes, Universität Bielefeld, Bielefeld., 2014.
- Clark, W. B., Brook, I., Bianki, D. und Thompson, D. H.: Microbiology of otitis externa, *Otolaryngol.--Head Neck Surg. Off. J. Am. Acad. Otolaryngol.-Head Neck Surg.*, 116(1), 23–25, 1997.

Clarke, S. R. und Foster, S. J.: Surface adhesins of *Staphylococcus aureus*, *Adv. Microb. Physiol.*, 51, 187–224, doi:10.1016/S0065-2911(06)51004-5, 2006.

Cogen, A. L., Nizet, V. und Gallo, R. L.: Skin microbiota: a source of disease or defence?, *Br. J. Dermatol.*, 158(3), 442–455, doi:10.1111/j.1365-2133.2008.08437.x, 2008.

Cole, J. R., Chai, B., Marsh, T. L., Farris, R. J., Wang, Q., Kulam, S. A., Chandra, S., McGarrell, D. M., Schmidt, T. M., Garrity, G. M., Tiedje, J. M. und Ribosomal Database Project: The Ribosomal Database Project (RDP-II): previewing a new autoaligner that allows regular updates and the new prokaryotic taxonomy, *Nucleic Acids Res.*, 31(1), 442–443, 2003.

Collins, M. D., Hoyles, L., Foster, G., Sjödnén, B. und Falsen, E.: *Corynebacterium capitovis* sp. nov., from a sheep, *Int. J. Syst. Evol. Microbiol.*, 51(Pt 3), 857–860, 2001.

Costello, E. K., Lauber, C. L., Hamady, M., Fierer, N., Gordon, J. I. und Knight, R.: Bacterial community variation in human body habitats across space and time, *Science*, 326(5960), 1694–1697, doi:10.1126/science.1177486, 2009.

Davis, I. J.: Axillary Malodor Neutralization, Patent, US5766578 A, 1998.

Decréau, R. A., Marson, C. M., Smith, K. E. und Behan, J. M.: Production of malodorous steroids from androsta-5,16-dienes and androsta-4,16-dienes by *Corynebacteria* and other human axillary bacteria, *J. Steroid Biochem. Mol. Biol.*, 87(4–5), 327–336, 2003.

DeSantis, T. Z., Hugenholtz, P., Larsen, N., Rojas, M., Brodie, E. L., Keller, K., Huber, T., Dalevi, D., Hu, P. und Andersen, G. L.: Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB, *Appl. Environ. Microbiol.*, 72(7), 5069–5072, doi:10.1128/AEM.03006-05, 2006.

Edouard, S., Bibi, F., Dhamodharan, R., Lagier, J.-C., Azhar, E. I., Robert, C., Caputo, A., Yasir, M., Jiman-Fatani, A. A., Alawi, M., Fournier, P.-E. und Raoult, D.: Non-contiguous finished genome sequence and description of *Corynebacterium jeddahense* sp. nov, *Stand. Genomic Sci.*, 9(3), 987–1002, doi:10.4056/sigs.5561028, 2014.

Egert, M., Schmidt, I., Höhne, H.-M., Lachnit, T., Schmitz, R. A. und Breves, R.: rRNA-based profiling of bacteria in the axilla of healthy males suggests right-left asymmetry in bacterial activity, *FEMS Microbiol. Ecol.*, 77(1), 146–153, doi:10.1111/j.1574-6941.2011.01097.x, 2011.

Egert, M., Höhne, H.-M., Weber, T., Simmering, R., Banowski, B. und Breves, R.: Identification of compounds inhibiting the C-S lyase activity of a cell extract from a *Staphylococcus* sp. isolated from human skin, *Lett. Appl. Microbiol.*, 57(6), 534–539, doi:10.1111/lam.12146, 2013.

Eid, J., Fehr, A., Gray, J., Luong, K., Lyle, J., Otto, G., Peluso, P., Rank, D., Baybayan, P., Bettman, B., Bibillo, A., Bjornson, K., Chaudhuri, B., Christians, F., Cicero, R., Clark, S., Dalal, R., Dewinter, A., Dixon, J., Foquet, M., Gaertner, A., Hardenbol, P., Heiner, C., Hester, K., Holden, D., Kearns, G., Kong, X., Kuse, R., Lacroix, Y., Lin, S., Lundquist, P., Ma, C., Marks, P., Maxham, M., Murphy,

- D., Park, I., Pham, T., Phillips, M., Roy, J., Sebra, R., Shen, G., Sorenson, J., Tomaney, A., Travers, K., Trulson, M., Vieceli, J., Wegener, J., Wu, D., Yang, A., Zaccarin, D., Zhao, P., Zhong, F., Korlach, J. und Turner, S.: Real-time DNA sequencing from single polymerase molecules, *Science*, 323(5910), 133–138, doi:10.1126/science.1162986, 2009.
- Eigen, E. und Froebe, K.: Deodorant compositions comprising inhibitors of odor-producing axillary bacterial exoenzymes, Patent, US5643559 A, 1997.
- Emter, R. und Natsch, A.: The sequential action of a dipeptidase and a beta-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by *Corynebacteria*, *J. Biol. Chem.*, 283(30), 20645–20652, doi:10.1074/jbc.M800730200, 2008.
- Ewing, B., Hillier, L., Wendl, M. C. und Green, P.: Base-calling of automated sequencer traces using phred. I. Accuracy assessment, *Genome Res.*, 8(3), 175–185, 1998.
- Fournier, B. und Philpott, D. J.: Recognition of *Staphylococcus aureus* by the innate immune system, *Clin. Microbiol. Rev.*, 18(3), 521–540, doi:10.1128/CMR.18.3.521-540.2005, 2005.
- Francis, H.: Anatomy of the temporal bone, external ear, and middle ear, in Cummings Otolaryngology: head & neck surgery, Mosby Elsevier, Philadelphia., 2010.
- Frankenberg-Dinkel, N.: Bacterial heme oxygenases, *Antioxid. Redox Signal.*, 6(5), 825–834, doi:10.1089/ars.2004.6.825, 2004.
- Fredrich, E., Barzantny, H., Brune, I. und Tauch, A.: Daily battle against body odor: towards the activity of the axillary microbiota, *Trends Microbiol.*, 21(6), 305–312, doi:10.1016/j.tim.2013.03.002, 2013.
- Fredricks, D. N.: Microbial ecology of human skin in health and disease, *J. Investig. Dermatol. Symp. Proc. Soc. Investig. Dermatol. Inc Eur. Soc. Dermatol. Res.*, 6(3), 167–169, doi:10.1046/j.0022-202x.2001.00039.x, 2001.
- Fujii, T., Shinozaki, J., Kajiura, T., Iwasaki, K. und Fudou, R.: A newly discovered Anaerococcus strain responsible for axillary odor and a new axillary odor inhibitor, pentagalloyl glucose, *FEMS Microbiol. Ecol.*, 89(1), 198–207, doi:10.1111/1574-6941.12347, 2014.
- Funke, G., Efstratiou, A., Kuklinska, D., Hutson, R. A., De Zoysa, A., Engler, K. H. und Collins, M. D.: *Corynebacterium imitans* sp. nov. isolated from patients with suspected diphtheria, *J. Clin. Microbiol.*, 35(8), 1978–1983, 1997.
- Gallagher, M., Wysocki, C. J., Leyden, J. J., Spielman, A. I., Sun, X. und Preti, G.: Analyses of volatile organic compounds from human skin, *Br. J. Dermatol.*, 159(4), 780–791, doi:10.1111/j.1365-2133.2008.08748.x, 2008.

- Gao, Z., Perez-Perez, G. I., Chen, Y. und Blaser, M. J.: Quantitation of major human cutaneous bacterial and fungal populations, *J. Clin. Microbiol.*, 48(10), 3575–3581, doi:10.1128/JCM.00597-10, 2010.
- Ghodsi, M., Liu, B. und Pop, M.: DNACLUST: accurate and efficient clustering of phylogenetic marker genes, *BMC Bioinformatics*, 12, 271, doi:10.1186/1471-2105-12-271, 2011.
- Gomis-Rüth, F. X.: Structure and mechanism of metallo-carboxypeptidases, *Crit. Rev. Biochem. Mol. Biol.*, 43(5), 319–345, doi:10.1080/10409230802376375, 2008.
- Gordon, B. I. und Maibach, H. I.: Eccrine anhidrosis due to glutaraldehyde, formaldehyde, and iontophoresis, *J. Invest. Dermatol.*, 53(6), 436–439, 1969.
- Gosalbes, M. J., Durbán, A., Pignatelli, M., Abellan, J. J., Jiménez-Hernández, N., Pérez-Cobas, A. E., Latorre, A. und Moya, A.: Metatranscriptomic approach to analyze the functional human gut microbiota, *PloS One*, 6(3), e17447, doi:10.1371/journal.pone.0017447, 2011.
- Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. und Lopez, R.: A new bioinformatics analysis tools framework at EMBL-EBI, *Nucleic Acids Res.*, 38(Web Server issue), W695-699, doi:10.1093/nar/gkq313, 2010.
- Gower, D. B., Holland, K. T., Mallet, A. I., Rennie, P. J. und Watkins, W. J.: Comparison of 16-androstene steroid concentrations in sterile apocrine sweat and axillary secretions: interconversions of 16-androstenes by the axillary microflora--a mechanism for axillary odour production in man?, *J. Steroid Biochem. Mol. Biol.*, 48(4), 409–418, 1994.
- von Graevenitz, A. und Funke, G.: *Turicella otitidis* and *Corynebacterium auris*: 20 years on, *Infection*, 42(1), 1–4, doi:10.1007/s15010-013-0488-x, 2014.
- Grice, E. A. und Segre, J. A.: The skin microbiome, *Nat. Rev. Microbiol.*, 9(4), 244–253, doi:10.1038/nrmicro2537, 2011.
- Grice, E. A., Kong, H. H., Conlan, S., Deming, C. B., Davis, J., Young, A. C., NISC Comparative Sequencing Program, Bouffard, G. G., Blakesley, R. W., Murray, P. R., Green, E. D., Turner, M. L. und Segre, J. A.: Topographical and temporal diversity of the human skin microbiome, *Science*, 324(5931), 1190–1192, doi:10.1126/science.1171700, 2009.
- Griffiths-Jones, S., Bateman, A., Marshall, M., Khanna, A. und Eddy, S. R.: Rfam: an RNA family database, *Nucleic Acids Res.*, 31(1), 439–441, 2003.
- Guss, J. und Ruckenstein, M.: Infections of the external ear, in *Cummings Otolaryngology: Head & Neck Surgery*, Mosby Elsevier., 2010.
- Haim, M., Trost, A., Maier, C. J., Achatz, G., Feichtner, S., Hintner, H., Bauer, J. W. und Onder, K.: Cytokeratin 8 interacts with clumping factor B: a new possible virulence factor target, *Microbiol. Read. Engl.*, 156(Pt 12), 3710–3721, doi:10.1099/mic.0.034413-0, 2010.

Hajioff, D. und Mackeith, S.: Otitis externa, *BMJ Clin. Evid.*, 2010, 2010.

Hannon Lab: FASTX-Toolkit, [online] Verfügbar über: [http://hannonlab.cshl.edu/fastx\\_toolkit/index.html](http://hannonlab.cshl.edu/fastx_toolkit/index.html) (Letzter Zugriff 3. Januar 2015), 2009.

Harriott, M. M. und Noverr, M. C.: *Candida albicans* and *Staphylococcus aureus* form polymicrobial biofilms: effects on antimicrobial resistance, *Antimicrob. Agents Chemother.*, 53(9), 3914–3922, doi:10.1128/AAC.00657-09, 2009.

Harvey, C. J., LeBouf, R. F. und Stefaniak, A. B.: Formulation and stability of a novel artificial human sweat under conditions of storage and use, *Toxicol. Vitro Int. J. Publ. Assoc. BIBRA*, 24(6), 1790–1796, doi:10.1016/j.tiv.2010.06.016, 2010.

Hasegawa, Y., Yabuki, M. und Matsukane, M.: Identification of new odoriferous compounds in human axillary sweat, *Chem. Biodivers.*, 1(12), 2042–2050, doi:10.1002/cbdv.200490157, 2004.

Hölzle, E.: Pathophysiology of sweating, *Curr. Probl. Dermatol.*, 30, 10–22, 2002.

Hooper, N. M.: Families of zinc metalloproteases, *FEBS Lett.*, 354(1), 1–6, 1994.

Human Microbiome Project Consortium: Structure, function and diversity of the healthy human microbiome, *Nature*, 486(7402), 207–214, doi:10.1038/nature11234, 2012.

Jacob, S., McClintock, M. K., Zelano, B. und Ober, C.: Paternally inherited HLA alleles are associated with women's choice of male odor, *Nat. Genet.*, 30(2), 175–179, doi:10.1038/ng830, 2002.

James, A. G., Casey, J., Hyliands, D. und Mycock, G.: Fatty acid metabolism by cutaneous bacteria and its role in axillary malodour, *World J. Microbiol. Biotechnol.*, 20(8), 787–793, doi:10.1007/s11274-004-5843-8, 2004a.

James, A. G., Hyliands, D. und Johnston, H.: Generation of volatile fatty acids by axillary bacteria, *Int. J. Cosmet. Sci.*, 26(3), 149–156, doi:10.1111/j.1467-2494.2004.00214.x, 2004b.

James, A. G., Austin, C. J., Cox, D. S., Taylor, D. und Calvert, R.: Microbiological and biochemical origins of human axillary odour, *FEMS Microbiol. Ecol.*, 83(3), 527–540, doi:10.1111/1574-6941.12054, 2013.

Jensen, L. J., Julien, P., Kuhn, M., von Mering, C., Muller, J., Doerks, T. und Bork, P.: eggNOG: automated construction and annotation of orthologous groups of genes, *Nucleic Acids Res.*, 36(Database issue), D250-254, doi:10.1093/nar/gkm796, 2008.

Johnson, J. M. und Proppe, D.: Cardiovascular Adjustments to Heat Stress, in *Handbook of Physiology*, Section 4: Environmental Physiology, Oxford University Press, Oxford., 1996.

- jstjohn: SeqPrep, GitHub [online] Verfügbar über: <https://github.com/jstjohn/SeqPrep> (Letzter Zugriff 2. Januar 2015), 2015.
- Juge, N., Tailford, L. und Owen, C. D.: Sialidases from gut bacteria: a mini-review, *Biochem. Soc. Trans.*, 44(1), 166–175, doi:10.1042/BST20150226, 2016.
- Kämpfer, P., Lidders, N., Warfolomeow, I., Falsen, E. und Busse, H.-J.: *Corynebacterium lubricantis* sp. nov., isolated from a coolant lubricant, *Int. J. Syst. Evol. Microbiol.*, 59(Pt 5), 1112–1115, doi:10.1099/ij.s.0.006379-0, 2009.
- Kanehisa, M.: A database for post-genome analysis, *Trends Genet. TIG*, 13(9), 375–376, 1997.
- Kanlayavattanakul, M. und Lourith, N.: Body malodours and their topical treatment agents, *Int. J. Cosmet. Sci.*, 33(4), 298–311, doi:10.1111/j.1468-2494.2011.00649.x, 2011.
- Karaca, Ç. T., Akçay, S. Ş., Toros, S. Z., Oysu, Ç., Verim, A., Çelebi, Ş. und Aksaray, S.: External auditory canal microbiology and hearing aid use, *Am. J. Otolaryngol.*, 34(4), 278–281, doi:10.1016/j.amjoto.2012.12.002, 2013.
- Katsaveli, K., Vayenas, D., Tsiamis, G. und Bourtzis, K.: Bacterial diversity in Cr(VI) and Cr(III)-contaminated industrial wastewaters, *Extrem. Life Extreme Cond.*, 16(2), 285–296, doi:10.1007/s00792-012-0429-0, 2012.
- Kawai, T. und Akira, S.: Toll-like receptors and their crosstalk with other innate receptors in infection and immunity, *Immunity*, 34(5), 637–650, doi:10.1016/j.immuni.2011.05.006, 2011.
- KEGG: KEGG Mapper, [online] Verfügbar über: <http://www.genome.jp/kegg/mapper.html> (Letzter Zugriff 18. September 2016), 2010.
- Kent, W. J.: BLAT--the BLAST-like alignment tool, *Genome Res.*, 12(4), 656–664, doi:10.1101/gr.229202. Article published online before March 2002, 2002.
- Kim, K. K., Lee, J.-S. und Stevens, D. A.: Microbiology and epidemiology of *Halomonas* species, *Future Microbiol.*, 8(12), 1559–1573, doi:10.2217/fmb.13.108, 2013.
- Klein, J.: Otitis externa, otitis media, and mastoiditis., in *Principles and Practice of Infectious Diseases*, PA: Elsevier Churchill Livingstone, Philadelphia., 2009.
- Kojic, E. M. und Darouiche, R. O.: *Candida* infections of medical devices, *Clin. Microbiol. Rev.*, 17(2), 255–267, 2004.
- Kong, H. H.: Skin microbiome: genomics-based insights into the diversity and role of skin microbes, *Trends Mol. Med.*, 17(6), 320–328, doi:10.1016/j.molmed.2011.01.013, 2011.



- Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers, D. und Knight, R.: Experimental and analytical tools for studying the human microbiome, *Nat. Rev. Genet.*, 13(1), 47–58, doi:10.1038/nrg3129, 2012.
- Kuhn, F. und Natsch, A.: Body odour of monozygotic human twins: a common pattern of odorant carboxylic acids released by a bacterial aminoacylase from axilla secretions contributing to an inherited body odour type, *J. R. Soc. Interface R. Soc.*, 6(33), 377–392, doi:10.1098/rsif.2008.0223, 2009.
- Kwaszewska, A., Sobiś-Glinkowska, M. und Szewczyk, E. M.: Cohabitation--relationships of corynebacteria and staphylococci on human skin, *Folia Microbiol. (Praha)*, 59(6), 495–502, doi:10.1007/s12223-014-0326-2, 2014.
- Labows, J. N., Preti, G., Hoelzle, E., Leyden, J. und Kligman, A.: Steroid analysis of human apocrine secretion, *Steroids*, 34(3), 249–258, 1979.
- Labows, J. N., Reilly, J. T., Leyden, J. J. und Preti, G.: Axillary Odor Determination, Formation and Control., in In: Laden K, editor. *Antiperspirants and Deodorants*, pp. 59–82, Marcel Dekker Inc., New York., 1999.
- Langmead, B., Trapnell, C., Pop, M. und Salzberg, S. L.: Ultrafast and memory-efficient alignment of short DNA sequences to the human genome, *Genome Biol.*, 10(3), R25, doi:10.1186/gb-2009-10-3-r25, 2009.
- Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. und Higgins, D. G.: Clustal W and Clustal X version 2.0, *Bioinformatics*, 23(21), 2947–2948, doi:10.1093/bioinformatics/btm404, 2007.
- Lasken, R. S. und McLean, J. S.: Recent advances in genomic DNA sequencing of microbial species from single cells, *Nat. Rev. Genet.*, 15(9), 577–584, doi:10.1038/nrg3785, 2014.
- Lederberg, J. und McCray, A.: 'Omesweet 'omics: - a genealogical treasury of words., *Scientist*, 15(8), 2001.
- Lee, H., Kim, J. und Nguyen, V.: Ear infections: otitis externa and otitis media, *Prim. Care*, 40(3), 671–686, doi:10.1016/j.pop.2013.05.005, 2013.
- Leyden, J. J., McGinley, K. J., Hölzle, E., Labows, J. N. und Kligman, A. M.: The microbiology of the human axilla and its relationship to axillary odor, *J. Invest. Dermatol.*, 77(5), 413–416, 1981.
- Li, K., Bihan, M. und Methé, B. A.: Analyses of the stability and core taxonomic memberships of the human microbiome, *PLoS One*, 8(5), e63139, doi:10.1371/journal.pone.0063139, 2013.
- Makarova, K. S. und Grishin, N. V.: The Zn-peptidase superfamily: functional convergence after evolutionary divergence, *J. Mol. Biol.*, 292(1), 11–17, doi:10.1006/jmbi.1999.3059, 1999.

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., Geer, R. C., He, J., Gwadz, M., Hurwitz, D. I., Lanczycki, C. J., Lu, F., Marchler, G. H., Song, J. S., Thanki, N., Wang, Z., Yamashita, R. A., Zhang, D., Zheng, C. und Bryant, S. H.: CDD: NCBI's conserved domain database, *Nucleic Acids Res.*, 43(Database issue), D222-226, doi:10.1093/nar/gku1221, 2015.

Mardis, E. R.: The impact of next-generation sequencing technology on genetics, *Trends Genet. TIG*, 24(3), 133–141, doi:10.1016/j.tig.2007.12.007, 2008.

Marples, M. J.: *The Ecology of human skin*, Charles C. Thomas, Bannerstone House, Springfield, Illinois., 1956.

Martin, A., Saathoff, M., Kuhn, F., Max, H., Terstegen, L. und Natsch, A.: A functional ABCC11 allele is essential in the biochemical formation of human axillary odor, *J. Invest. Dermatol.*, 130(2), 529–540, doi:10.1038/jid.2009.254, 2010.

Martin, T. J., Kerschner, J. E. und Flanary, V. A.: Fungal causes of otitis externa and tympanostomy tube otorrhea, *Int. J. Pediatr. Otorhinolaryngol.*, 69(11), 1503–1508, doi:10.1016/j.ijporl.2005.04.012, 2005.

McDowall, G. D.: External otitis: otological problems, *J. Laryngol. Otol.*, 88(1), 1–13, 1974.

McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y. M., Buso, N., Cowley, A. P. und Lopez, R.: Analysis Tool Web Services from the EMBL-EBI, *Nucleic Acids Res.*, 41(Web Server issue), W597-600, doi:10.1093/nar/gkt376, 2013.

Metzker, M. L.: Sequencing technologies - the next generation, *Nat. Rev. Genet.*, 11(1), 31–46, doi:10.1038/nrg2626, 2010.

Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E. M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A., Wilkening, J. und Edwards, R. A.: The metagenomics RAST server – a public resource for the automatic phylogenetic and functional analysis of metagenomes, *BMC Bioinformatics*, 9(1), 386, doi:10.1186/1471-2105-9-386, 2008.

Morales, S. E. und Holben, W. E.: Linking bacterial identities and ecosystem processes: can “omic” analyses be more than the sum of their parts?, *FEMS Microbiol. Ecol.*, 75(1), 2–16, doi:10.1111/j.1574-6941.2010.00938.x, 2011.

Moran, M. A.: Metatranscriptomics: Eavesdropping on Complex Microbial Communities, *Microbe*, 4(7), 329–335, 2009.

Morgan, X. C. und Huttenhower, C.: Chapter 12: Human microbiome analysis, *PLoS Comput. Biol.*, 8(12), e1002808, doi:10.1371/journal.pcbi.1002808, 2012.

Murphy, E. C. und Frick, I.-M.: Gram-positive anaerobic cocci--commensals and opportunistic pathogens, *FEMS Microbiol. Rev.*, 37(4), 520–553, doi:10.1111/1574-6976.12005, 2013.

Nair, N., Biswas, R., Götz, F. und Biswas, L.: Staphylococcus aureus in polymicrobial infections: impact on pathogenesis, *Infect. Immun.*, doi:10.1128/IAI.00059-14, 2014.

Nakakido, M., Aikawa, C., Nakagawa, I. und Tsumoto, K.: The staphylococcal elastin-binding protein regulates zinc-dependent growth/biofilm formation, *J. Biochem. (Tokyo)*, 156(3), 155–162, doi:10.1093/jb/mvu027, 2014.

National Research Council (US) Committee on Metagenomics: Challenges and Functional Applications: The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet, National Academies Press (US), Washington (DC). [online] Verfügbar über: <http://www.ncbi.nlm.nih.gov/books/NBK54006/> (Letzter Zugriff 14. März 2016), 2007.

Natsch, A.: What Makes Us Smell: The Biochemistry of Body Odour and the Design of New Deodorant Ingredients, *Chimia*, 69(7–8), 414–420, doi:10.2533/chimia.2015.414, 2015.

Natsch, A., Gfeller, H., Gygax, P., Schmid, J. und Acuna, G.: A specific bacterial aminoacylase cleaves odorant precursors secreted in the human axilla, *J. Biol. Chem.*, 278(8), 5718–5727, doi:10.1074/jbc.M210142200, 2003.

Natsch, A., Schmid, J. und Flachsmann, F.: Identification of odoriferous sulfanylalkanols in human axilla secretions and their formation through cleavage of cysteine precursors by a C-S lyase isolated from axilla bacteria, *Chem. Biodivers.*, 1(7), 1058–1072, doi:10.1002/cbdv.200490079, 2004.

Natsch, A., Gfeller, H., Gygax, P. und Schmid, J.: Isolation of a bacterial enzyme releasing axillary malodor and its use as a screening target for novel deodorant formulations, *Int. J. Cosmet. Sci.*, 27(2), 115–122, doi:10.1111/j.1467-2494.2004.00255.x, 2005.

Natsch, A., Derrer, S., Flachsmann, F. und Schmid, J.: A broad diversity of volatile carboxylic acids, released by a bacterial aminoacylase from axilla secretions, as candidate molecules for the determination of human-body odor type, *Chem. Biodivers.*, 3(1), 1–20, doi:10.1002/cbdv.200690015, 2006.

Natsch, A., Kuhn, F. und Tiercy, J.-M.: Lack of evidence for HLA-linked patterns of odorous carboxylic acids released from glutamine conjugates secreted in the human axilla, *J. Chem. Ecol.*, 36(8), 837–846, doi:10.1007/s10886-010-9826-y, 2010.

Naumann, M., Lowe, N. J., Kumar, C. R., Hamm, H. und Hyperhidrosis Clinical Investigators Group: Botulinum toxin type a is a safe and effective treatment for axillary hyperhidrosis over 16 months: a prospective study, *Arch. Dermatol.*, 139(6), 731–736, doi:10.1001/archderm.139.6.731, 2003.

Nicholas, K. B. und Nicholas, H. B. J.: GeneDoc: a tool for editing and annotating multiple sequence alignments., 1997.

Nicolaides, N.: Skin lipids: their biochemical uniqueness, *Science*, 186(4158), 19–26, 1974.

- Nikolaki, S. und Tsiamis, G.: Microbial diversity in the era of omic technologies, *BioMed Res. Int.*, 2013, 958719, doi:10.1155/2013/958719, 2013.
- Ninkovic, G., Dullo, V. und Saunders, N. C.: Microbiology of otitis externa in the secondary care in United Kingdom and antimicrobial sensitivity, *Auris. Nasus. Larynx*, 35(4), 480–484, doi:10.1016/j.anl.2007.09.013, 2008.
- Noble, W. C.: Skin microbiology: coming of age, *J. Med. Microbiol.*, 17(1), 1–12, doi:10.1099/00222615-17-1-1, 1984.
- Noël, F., Piérard-Franchimont, C., Piérard, G. E. und Quatresooz, P.: Sweaty skin, background and assessments, *Int. J. Dermatol.*, 51(6), 647–655, doi:10.1111/j.1365-4632.2011.05307.x, 2012.
- Nossa, C. W., Oberdorf, W. E., Yang, L., Aas, J. A., Paster, B. J., Desantis, T. Z., Brodie, E. L., Malamud, D., Poles, M. A. und Pei, Z.: Design of 16S rRNA gene primers for 454 pyrosequencing of the human foregut microbiome, *World J. Gastroenterol.*, 16(33), 4135–4144, 2010.
- O'Donovan, D., Corcoran, G. D., Lucey, B. und Sleator, R. D.: *Campylobacter ureolyticus*, Virulence, 5(4), 498–506, doi:10.4161/viru.28776, 2014.
- Oh, J., Freeman, A. F., NISC Comparative Sequencing Program, Park, M., Sokolic, R., Candotti, F., Holland, S. M., Segre, J. A. und Kong, H. H.: The altered landscape of the human skin microbiome in patients with primary immunodeficiencies, *Genome Res.*, 23(12), 2103–2114, doi:10.1101/gr.159467.113, 2013.
- Oksuz, S., Sahin, I., Yildirim, M., Gulcan, A., Yavuz, T., Kaya, D. und Koc, A. N.: Phospholipase and proteinase activities in different *Candida* species isolated from anatomically distinct sites of healthy adults, *Jpn. J. Infect. Dis.*, 60(5), 280–283, 2007.
- Olsen, G. J., Larsen, N. und Woese, C. R.: The ribosomal RNA database project, *Nucleic Acids Res.*, 19 Suppl, 2017–2021, 1991.
- Pace, N. R., Stahl, D. A., Lane, D. J. und Olsen, G. J.: The Analysis of Natural Microbial Populations by Ribosomal RNA Sequences, in *Advances in Microbial Ecology*, edited by K. C. Marshall, pp. 1–55, Springer US., 1986.
- Padmanabhan, R., Dubourg, G., Lagier, J.-C., Couderc, C., Michelle, C., Raoult, D. und Fournier, P.-E.: Genome sequence and description of *Corynebacterium ihumii* sp. nov, *Stand. Genomic Sci.*, 9(3), 1128–1143, doi:10.4056/sigs.5149006, 2014.
- Pareek, C. S., Smoczynski, R. und Tretyn, A.: Sequencing technologies and genome sequencing, *J. Appl. Genet.*, 52(4), 413–435, doi:10.1007/s13353-011-0057-x, 2011.

Percival, S. L., Emanuel, C., Cutting, K. F. und Williams, D. W.: Microbiology of the skin and the role of biofilms in infection, *Int. Wound J.*, 9(1), 14–32, doi:10.1111/j.1742-481X.2011.00836.x, 2012.

Pérez-Cobas, A. E., Gosalbes, M. J., Friedrichs, A., Knecht, H., Artacho, A., Eismann, K., Otto, W., Rojo, D., Bargiela, R., von Bergen, M., Neulinger, S. C., Däumer, C., Heinsen, F.-A., Latorre, A., Barbas, C., Seifert, J., dos Santos, V. M., Ott, S. J., Ferrer, M. und Moya, A.: Gut microbiota disturbance during antibiotic therapy: a multi-omic approach, *Gut*, 62(11), 1591–1601, doi:10.1136/gutjnl-2012-303184, 2013.

Peterkin, G. A.: Otitis externa, *J. Laryngol. Otol.*, 88(1), 15–21, 1974.

Peters, B. M. und Noverr, M. C.: *Candida albicans*-*Staphylococcus aureus* polymicrobial peritonitis modulates host innate immunity, *Infect. Immun.*, 81(6), 2178–2189, doi:10.1128/IAI.00265-13, 2013.

Petrosino, J. F., Highlander, S., Luna, R. A., Gibbs, R. A. und Versalovic, J.: Metagenomic pyrosequencing and microbial identification, *Clin. Chem.*, 55(5), 856–866, doi:10.1373/clinchem.2008.107565, 2009.

Piérard, G. E., Elsner, P., Marks, R., Masson, P., Paye, M. und EEMCO Group: EEMCO guidance for the efficacy assessment of antiperspirants and deodorants, *Skin Pharmacol. Appl. Skin Physiol.*, 16(5), 324–342, doi:72072, 2003.

Pontes, Z. B. V. da S., Silva, A. D. F., Lima, E. de O., Guerra, M. de H., Oliveira, N. M. C., Carvalho, M. de F. F. P. und Guerra, F. S. Q.: Otomycosis: a retrospective study, *Braz. J. Otorhinolaryngol.*, 75(3), 367–370, 2009.

Powell, S., Szklarczyk, D., Trachana, K., Roth, A., Kuhn, M., Muller, J., Arnold, R., Rattei, T., Letunic, I., Doerks, T., Jensen, L. J., von Mering, C. und Bork, P.: eggNOG v3.0: orthologous groups covering 1133 organisms at 41 different taxonomic ranges, *Nucleic Acids Res.*, 40(Database issue), D284-289, doi:10.1093/nar/gkr1060, 2012.

Probst, A. J., Auerbach, A. K. und Moissl-Eichinger, C.: Archaea on human skin, *PLoS One*, 8(6), e65388, doi:10.1371/journal.pone.0065388, 2013.

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J. und Glöckner, F. O.: SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB, *Nucleic Acids Res.*, 35(21), 7188–7196, doi:10.1093/nar/gkm864, 2007.

Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K. S., Manichanh, C., Nielsen, T., Pons, N., Levenez, F., Yamada, T., Mende, D. R., Li, J., Xu, J., Li, S., Li, D., Cao, J., Wang, B., Liang, H., Zheng, H., Xie, Y., Tap, J., Lepage, P., Bertalan, M., Batto, J.-M., Hansen, T., Le Paslier, D., Linneberg, A., Nielsen, H. B., Pelletier, E., Renault, P., Sicheritz-Ponten, T., Turner, K., Zhu, H., Yu, C., Li, S., Jian, M., Zhou, Y., Li, Y., Zhang, X., Li, S., Qin, N., Yang, H., Wang, J., Brunak, S., Doré, J., Guarner, F., Kristiansen, K., Pedersen, O., Parkhill, J., Weissenbach, J., MetaHIT Consortium, Bork, P., Ehrlich,

S. D. und Wang, J.: A human gut microbial gene catalogue established by metagenomic sequencing, *Nature*, 464(7285), 59–65, doi:10.1038/nature08821, 2010.

Ram, R., Lowe, N. J. und Yamauchi, P. S.: Current and emerging therapeutic modalities for hyperhidrosis, part 1: conservative and noninvasive treatments, *Cutis*, 79(3), 211–217, 2007.

Rho, M., Tang, H. und Ye, Y.: FragGeneScan: predicting genes in short and error-prone reads, *Nucleic Acids Res.*, 38(20), e191, doi:10.1093/nar/gkq747, 2010.

Roberts, S. C., Gosling, L. M., Carter, V. und Petrie, M.: MHC-correlated odour preferences in humans and the use of oral contraceptives, *Proc. Biol. Sci.*, 275(1652), 2715–2722, doi:10.1098/rspb.2008.0825, 2008.

Roland, P. S. und Stroman, D. W.: Microbiology of acute otitis externa, *The Laryngoscope*, 112(7 Pt 1), 1166–1177, doi:10.1097/00005537-200207000-00005, 2002.

Rosenfeld, R. M., Schwartz, S. R., Cannon, C. R., Roland, P. S., Simon, G. R., Kumar, K. A., Huang, W. W., Haskell, H. W. und Robertson, P. J.: Clinical practice guideline: acute otitis externa, *Otolaryngol.–Head Neck Surg. Off. J. Am. Acad. Otolaryngol.–Head Neck Surg.*, 150(1 Suppl), S1–S24, doi:10.1177/0194599813517083, 2014.

Roth, R. R. und James, W. D.: Microbial ecology of the skin, *Annu. Rev. Microbiol.*, 42, 441–464, doi:10.1146/annurev.mi.42.100188.002301, 1988.

Roth, R. R. und James, W. D.: Microbiology of the skin: resident flora, ecology, infection, *J. Am. Acad. Dermatol.*, 20(3), 367–390, 1989.

Rowlands, S., Devalia, H., Smith, C., Hubbard, R. und Dean, A.: Otitis externa in UK general practice: a survey using the UK General Practice Research Database, *Br. J. Gen. Pract. J. R. Coll. Gen. Pract.*, 51(468), 533–538, 2001.

Sato, K. und Dobson, R. L.: Mechanism of the antiperspirant effect of topical glutaraldehyde, *Arch. Dermatol.*, 100(5), 564–569, 1969.

Sato, K., Leidal, R. und Sato, F.: Morphology and development of an apoeccrine sweat gland in human axillae, *Am. J. Physiol.*, 252(1 Pt 2), R166-180, 1987.

Sato, K., Kang, W. H., Saga, K. und Sato, K. T.: Biology of sweat glands and their disorders. I. Normal sweat gland function, *J. Am. Acad. Dermatol.*, 20(4), 537–563, 1989.

Sayers, E. W., Barrett, T., Benson, D. A., Bolton, E., Bryant, S. H., Canese, K., Chetvernin, V., Church, D. M., Dicuccio, M., Federhen, S., Feolo, M., Geer, L. Y., Helmberg, W., Kapustin, Y., Landsman, D., Lipman, D. J., Lu, Z., Madden, T. L., Madej, T., Maglott, D. R., Marchler-Bauer, A., Miller, V., Mizrachi, I., Ostell, J., Panchenko, A., Pruitt, K. D., Schuler, G. D., Sequeira, E., Sherry, S. T., Shumway, M., Sirotkin, K., Slotta, D., Souvorov, A., Starchenko, G., Tatusova, T. A., Wagner, L., Wang, Y., John Wilbur, W., Yaschenko, E. und Ye, J.: Database resources of the National

Center for Biotechnology Information, *Nucleic Acids Res.*, 38(Database issue), D5-16, doi:10.1093/nar/gkp967, 2010.

Schaefer, P. und Baugh, R. F.: Acute otitis externa: an update, *Am. Fam. Physician*, 86(11), 1055–1061, 2012.

Shelley, W. B., Hurley, H. J. und Nichols, A. C.: Axillary odor; experimental study of the role of bacteria, apocrine sweat, and deodorants, *AMA Arch. Dermatol. Syphilol.*, 68(4), 430–446, 1953.

Shirtliff, M. E., Peters, B. M. und Jabra-Rizk, M. A.: Cross-kingdom interactions: *Candida albicans* and bacteria, *FEMS Microbiol. Lett.*, 299(1), 1–8, doi:10.1111/j.1574-6968.2009.01668.x, 2009.

Sigma Aldrich Co. LL: Sicherheitsdatenblatt beta-Hydroxyisovaleriansäure, [online] Verfügbar über:

<http://www.sigmaaldrich.com/MSDS/MSDS/DisplayMSDSPage.do?country=DE&language=de&productNumber=55453&brand=ALDRICH&PageToGoToURL=http%3A%2F%2Fwww.sigmaaldrich.com%2Fcatalog%2Fproduct%2Faldrich%2F55453%3Flang%3Dde> (Letzter Zugriff 11. Juni 2016), 2012.

Sigma Aldrich Co. LL: Sicherheitsdatenblatt 4-Methyl-2-oxo-valeriansäure, [online] Verfügbar über:

<http://www.sigmaaldrich.com/MSDS/MSDS/DisplayMSDSPage.do?country=DE&language=de&productNumber=68255&brand=ALDRICH&PageToGoToURL=http%3A%2F%2Fwww.sigmaaldrich.com%2Fcatalog%2Fproduct%2Faldrich%2F68255%3Flang%3Dde> (Letzter Zugriff 10. Juni 2016), 2013.

Sigma Aldrich Co. LL: Sicherheitsdatenblatt Isobuttersäure, [online] Verfügbar über:

<http://www.sigmaaldrich.com/MSDS/MSDS/DisplayMSDSPage.do?country=DE&language=de&productNumber=11754&brand=SIGMA&PageToGoToURL=http%3A%2F%2Fwww.sigmaaldrich.com%2Fcatalog%2Fproduct%2Fsigma%2Fi1754%3Flang%3Dde> (Letzter Zugriff 10. Juni 2016), 2016.

Singh, R. und Ray, P.: Quorum sensing-mediated regulation of staphylococcal virulence and antibiotic resistance, *Future Microbiol.*, 9(5), 669–681, doi:10.2217/fmb.14.31, 2014.

Slingluff, C. L., Petroni, G. R., Molhoek, K. R., Brautigan, D. L., Chianese-Bullock, K. A., Shada, A. L., Smolkin, M. E., Olson, W. C., Gaucher, A., Chase, C. M., Grosh, W. W., Weiss, G. R., Wagenseller, A. G., Olszanski, A. J., Martin, L., Shea, S. M., Erdag, G., Ram, P., Gershenwald, J. E. und Weber, M. J.: Clinical activity and safety of combination therapy with temsirolimus and bevacizumab for advanced melanoma: a phase II trial (CTEP 7190/Mel47), *Clin. Cancer Res. Off. J. Am. Assoc. Cancer Res.*, 19(13), 3611–3620, doi:10.1158/1078-0432.CCR-12-3919, 2013.

Speziale, P., Pietrocola, G., Rindi, S., Provenzano, M., Provenza, G., Di Poto, A., Visai, L. und Arciola, C. R.: Structural and functional role of *Staphylococcus aureus* surface components

recognizing adhesive matrix molecules of the host, *Future Microbiol.*, 4(10), 1337–1352, doi:10.2217/fmb.09.102, 2009.

Speziale, P., Pietrocola, G., Foster, T. J. und Geoghegan, J. A.: Protein-based biofilm matrices in *Staphylococci*, *Front. Cell. Infect. Microbiol.*, 4, 171, doi:10.3389/fcimb.2014.00171, 2014.

Spielman, A. I., Zeng, X. N., Leyden, J. J. und Preti, G.: Proteinaceous precursors of human axillary odor: isolation of two novel odor-binding proteins, *Experientia*, 51(1), 40–47, 1995.

Starkenmann, C., Niclass, Y., Troccaz, M. und Clark, A. J.: Identification of the precursor of (S)-3-methyl-3-sulfanylohexan-1-ol, the sulfury malodour of human axilla sweat, *Chem. Biodivers.*, 2(6), 705–716, doi:10.1002/cbdv.200590048, 2005.

Starkenmann, C., Mayenzet, F., Brauchli, R. und Troccaz, M.: 5 $\alpha$ -Androst-16-en-3 $\alpha$ -ol  $\beta$ -D-glucuronide, precursor of 5 $\alpha$ -androst-16-en-3 $\alpha$ -ol in human sweat, *Chem. Biodivers.*, 10(12), 2197–2208, doi:10.1002/cbdv.201300286, 2013.

Stashak, A.-B. und Brewer, J. D.: Management of hyperhidrosis, *Clin. Cosmet. Investig. Dermatol.*, 7, 285–299, doi:10.2147/CCID.S53119, 2014.

Stevens, D. A., Hamilton, J. R., Johnson, N., Kim, K. K. und Lee, J.-S.: *Halomonas*, a newly recognized human pathogen causing infections and contamination in a dialysis center: three new species, *Medicine (Baltimore)*, 88(4), 244–249, doi:10.1097/MD.0b013e3181aede29, 2009.

Stevens, D. A., Kim, K. K., Johnson, N., Lee, J.-S. und Hamilton, J. R.: *Halomonas johnsoniae*: review of a medically underappreciated genus of growing human importance, *Am. J. Med. Sci.*, 345(5), 335–338, doi:10.1097/MAJ.0b013e31825600de, 2013.

Stewart, J. C. M.: Tomatoes cause under-arm odour, *Med. Hypotheses*, 82(5), 518–521, doi:10.1016/j.mehy.2014.02.001, 2014.

Stroman, D. W., Roland, P. S., Dohar, J. und Burt, W.: Microbiology of normal external auditory canal, *The Laryngoscope*, 111(11 Pt 1), 2054–2059, doi:10.1097/00005537-200111000-00035, 2001.

Styger, G., Prior, B. und Bauer, F. F.: Wine flavor and aroma, *J. Ind. Microbiol. Biotechnol.*, 38(9), 1145–1159, doi:10.1007/s10295-011-1018-4, 2011.

Swiegers, J. H. und Pretorius, I. S.: Modulation of volatile sulfur compounds by wine yeast, *Appl. Microbiol. Biotechnol.*, 74(5), 954–960, doi:10.1007/s00253-006-0828-1, 2007.

Tauch, A., Kaiser, O., Hain, T., Goesmann, A., Weisshaar, B., Albersmeier, A., Bekel, T., Bischoff, N., Brune, I., Chakraborty, T., Kalinowski, J., Meyer, F., Rupp, O., Schneiker, S., Viehoveer, P. und Pühler, A.: Complete genome sequence and analysis of the multiresistant nosocomial pathogen *Corynebacterium jeikeium* K411, a lipid-requiring bacterium of the human skin flora, *J. Bacteriol.*, 187(13), 4671–4682, doi:10.1128/JB.187.13.4671-4682.2005, 2005.



Taylor, D., Daulby, A., Grimshaw, S., James, G., Mercer, J. und Vaziri, S.: Characterization of the microflora of the human axilla, *Int. J. Cosmet. Sci.*, 25(3), 137–145, doi:10.1046/j.1467-2494.2003.00181.x, 2003.

Thermo Scientific: NanoDrop Products - Spectrophotometers and Fluorospectrometers - [www.nanodrop.com](http://www.nanodrop.com), Nanodrop Prod. [online] Verfügbar über: <http://www.nanodrop.com/> (Letzter Zugriff 3. Januar 2015), 2013.

Troccaz, M., Starckenmann, C., Niclass, Y., van de Waal, M. und Clark, A. J.: 3-Methyl-3-sulfanylhexan-1-ol as a major descriptor for the human axilla-sweat odour profile, *Chem. Biodivers.*, 1(7), 1022–1035, doi:10.1002/cbdv.200490077, 2004.

Troccaz, M., Benattia, F., Borchard, G. und Clark, A. J.: Properties of recombinant *Staphylococcus haemolyticus* cystathionine beta-lyase (metC) and its potential role in the generation of volatile thiols in axillary malodor, *Chem. Biodivers.*, 5(11), 2372–2385, doi:10.1002/cbdv.200890202, 2008.

Troccaz, M., Borchard, G., Vuilleumier, C., Raviot-Derrien, S., Niclass, Y., Beccucci, S. und Starckenmann, C.: Gender-specific differences between the concentrations of nonvolatile (R)/(S)-3-methyl-3-sulfanylhexan-1-ol and (R)/(S)-3-hydroxy-3-methyl-hexanoic acid odor precursors in axillary secretions, *Chem. Senses*, 34(3), 203–210, doi:10.1093/chemse/bjn076, 2009.

Troccaz, M., Gaïa, N., Beccucci, S., Schrenzel, J., Cayeux, I., Starckenmann, C. und Lazarevic, V.: Mapping axillary microbiota responsible for body odours using a culture-independent approach, *Microbiome*, 3(1), 3, doi:10.1186/s40168-014-0064-3, 2015.

Turnbaugh, P. J., Ley, R. E., Hamady, M., Fraser-Liggett, C. M., Knight, R. und Gordon, J. I.: The human microbiome project, *Nature*, 449(7164), 804–810, doi:10.1038/nature06244, 2007.

Twin, J., Bradshaw, C. S., Garland, S. M., Fairley, C. K., Fethers, K. und Tabrizi, S. N.: The potential of metatranscriptomics for identifying screening targets for bacterial vaginosis, *PloS One*, 8(9), e76892, doi:10.1371/journal.pone.0076892, 2013.

Urban, J., Fergus, D. J., Savage, A. M., Ehlers, M., Menninger, H. L., Dunn, R. R. und Horvath, J. E.: The effect of habitual and experimental antiperspirant and deodorant product use on the armpit microbiome, *PeerJ*, 4, e1605, doi:10.7717/peerj.1605, 2016.

Wang, Q., Garrity, G. M., Tiedje, J. M. und Cole, J. R.: Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy, *Appl. Environ. Microbiol.*, 73(16), 5261–5267, doi:10.1128/AEM.00062-07, 2007.

Wedekind, C., Seebeck, T., Bettens, F. und Paepke, A. J.: MHC-dependent mate preferences in humans, *Proc. Biol. Sci.*, 260(1359), 245–249, doi:10.1098/rspb.1995.0087, 1995.

Wilke, K., Martin, A., Terstegen, L. und Biel, S. S.: A short history of sweat gland biology, *Int. J. Cosmet. Sci.*, 29(3), 169–179, doi:10.1111/j.1467-2494.2007.00387.x, 2007.

- Witt, E., Frank, R. und Hengstenberg, W.: 6-Phospho-beta-galactosidases of gram-positive and 6-phospho-beta-glucosidase B of gram-negative bacteria: comparison of structure and function by kinetic and immunological methods and mutagenesis of the lacG gene of *Staphylococcus aureus*, *Protein Eng.*, 6(8), 913–920, 1993.
- Yamaguchi, M., Yamazaki, K., Beauchamp, G. K., Bard, J., Thomas, L. und Boyse, E. A.: Distinctive urinary odors governed by the major histocompatibility locus of the mouse, *Proc. Natl. Acad. Sci. U. S. A.*, 78(9), 5817–5820, 1981.
- Yanagawa, R.: Causative agents of bovine pyelonephritis: *Corynebacterium renale*, *C. pilosum* and *C. cystitidis*, *Prog. Vet. Microbiol. Immunol.*, 2, 158–174, 1986.
- Yassin, A. F.: *Corynebacterium ureicelerivorans* sp. nov., a lipophilic bacterium isolated from blood culture, *Int. J. Syst. Evol. Microbiol.*, 57(Pt 6), 1200–1203, doi:10.1099/ijs.0.64832-0, 2007.
- Yoshiura, K., Kinoshita, A., Ishida, T., Ninokata, A., Ishikawa, T., Kaname, T., Bannai, M., Tokunaga, K., Sonoda, S., Komaki, R., Ihara, M., Saenko, V. A., Alipov, G. K., Sekine, I., Komatsu, K., Takahashi, H., Nakashima, M., Sosonkina, N., Mapendano, C. K., Ghadami, M., Nomura, M., Liang, D.-S., Miwa, N., Kim, D.-K., Garidkhuu, A., Natsume, N., Ohta, T., Tomita, H., Kaneko, A., Kikuchi, M., Russomando, G., Hirayama, K., Ishibashi, M., Takahashi, A., Saitou, N., Murray, J. C., Saito, S., Nakamura, Y. und Niikawa, N.: A SNP in the ABCC11 gene is the determinant of human earwax type, *Nat. Genet.*, 38(3), 324–330, doi:10.1038/ng1733, 2006.
- Yu, Y., Lee, C., Kim, J. und Hwang, S.: Group-specific primer and probe sets to detect methanogenic communities using quantitative real-time polymerase chain reaction, *Biotechnol. Bioeng.*, 89(6), 670–679, doi:10.1002/bit.20347, 2005.
- Zdych, E., Peist, R., Reidl, J. und Boos, W.: MalY of *Escherichia coli* is an enzyme with the activity of a beta C-S lyase (cystathionase), *J. Bacteriol.*, 177(17), 5035–5039, 1995.
- Zecconi, A. und Scali, F.: *Staphylococcus aureus* virulence factors in evasion from innate immune defenses in human and animal diseases, *Immunol. Lett.*, 150(1–2), 12–22, doi:10.1016/j.imlet.2013.01.004, 2013.
- Zeng, C., Spielman, A. I., Vowels, B. R., Leyden, J. J., Biemann, K. und Preti, G.: A human axillary odorant is carried by apolipoprotein D, *Proc. Natl. Acad. Sci. U. S. A.*, 93(13), 6626–6630, 1996.
- Zeng, X. N., Leyden, J. J., Lawley, H. J., Sawano, K., Nohara, I. und Preti, G.: Analysis of characteristic odors from human male axillae, *J. Chem. Ecol.*, 17(7), 1469–1492, doi:10.1007/BF00983777, 1991.

## VIII. Anhang

## 1 Referenzsequenzen zu COG1168 (MalY) aus Staphylokokken

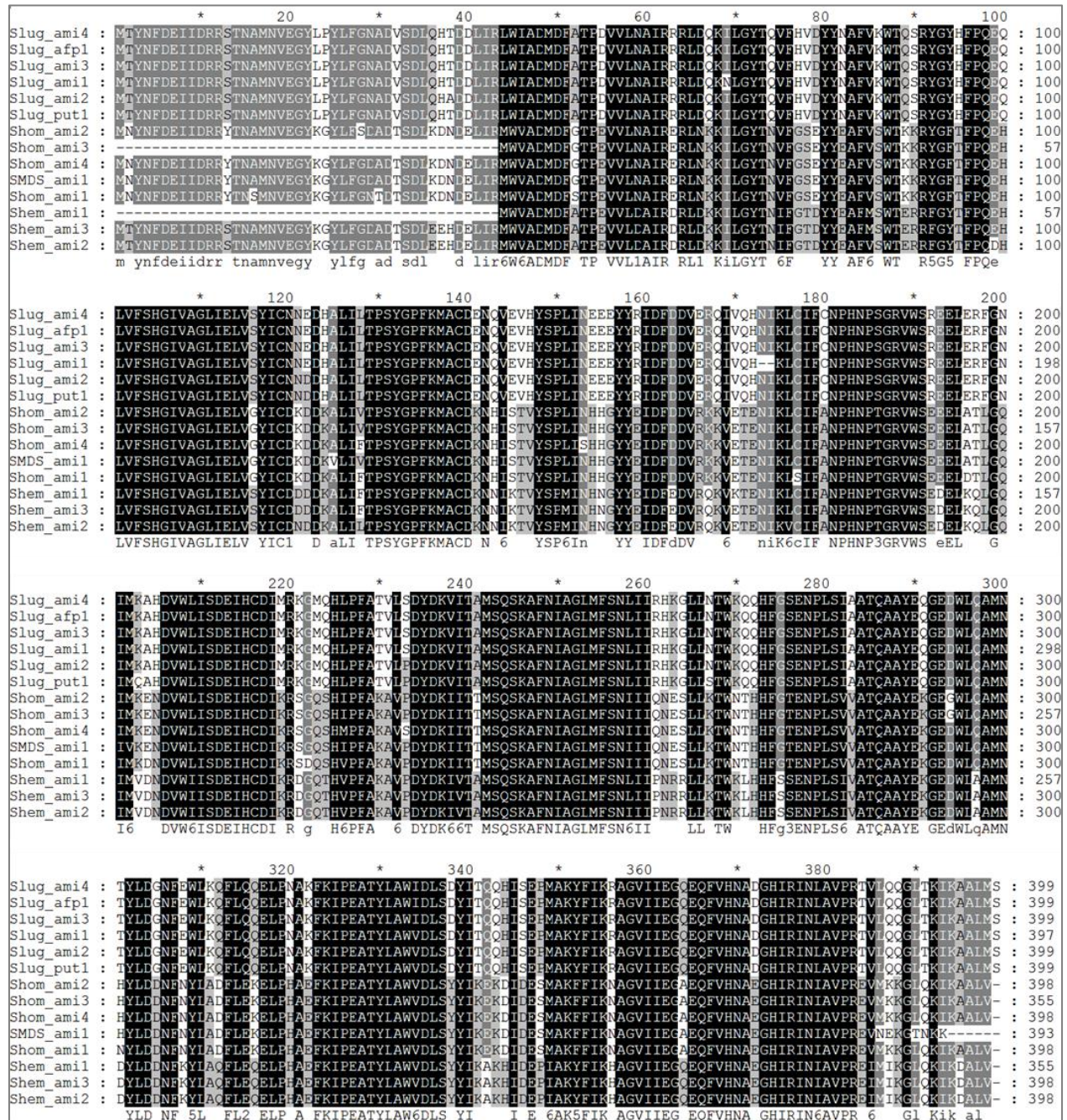


Abbildung 43: Multiple Sequence Alignment der mittels BLASTX identifizierten Referenzsequenzen aus der GenBank-Datenbank zu cDNA-Sequenzen mit Homologie zu COG1168 (MalY). Dargestellt sind lediglich Referenzsequenzen von *Staphylococcus*-Spezies. Slug: *Staphylococcus lugdunensis*, Shom: *Staphylococcus hominis*, Shem: *Staphylococcus haemolyticus*, SMDS: *Staphylococcus* sp. MDS7B; ami: Aminotransferase, put: putative C-S-Lyase, AFP: Aminotransferase.



## 2 Referenzsequenzen zur TpdA-Dipeptidase

```

*          20          *          40          *          60          *          80
CGD7_M20 : ---MTAQNGSQNYTPQRDRIFNLLSKLVSENSVHSTFEALADCHEPAACAVTANALGDLGLDVTTRYPTVDDADTIIVAKREP : 76
Cure_M20 : ---MPEQT-AQHFSPPRDRIFKLLSELVAFNSVHSTFEALADCHEPAACAVTVPALQALGLDVTTRYPTVDDADTIIVGKREP : 75
CJCB_M20 : ---MTTNDAQSYQPQRDRIFALLSLIVSENSHSTFEALADCHEPAACAVTVAALKEGGLDVTTRHATVDDADTIIVATKAP : 76
Ccap_M20 : ---MNTTLF-----TPQRDRIFALLSALVSENSHSTFEALADCHEPAACAVTVALEELGLDVTTRHATVDDADTIIVGKREP : 72
Cstr_tpd : MSNDKAATSTNFNLTTPNREIFQQLSELVAFNSVHSTFEALADCHEPAACAVTVALEELGLDVTTRHATVDDADTIIVGKREP : 80
Ctus_M20 : -----MSTFTPNREIFNLLSALVSENSHSTVEEADCHAAAADVAVALAERGLVEERHATVAKALAIIGRRHV : 70
Ctub_M20 : -----MTALNESIRNDREIFRQLSELVAFNSVHDEEGLEDCTKGAQAVKSALEEAGVSVETITIVAGSTAILGERRG : 74
Cacc_M20 : -----MTNIIDSVRSDRDTIFRQLSELVAFNSVHDEEGLEDACCGASQAVKSALEEAGVSVETITIVAGSTAILGERRG : 74
Ccal_hyp : MSQFSNVQEIQRHLEGQRDQLEFQQLKELVSENSVHSTFEALIDDYEGATNVRVATLTAALGIVTWEYFAEPGGTTNFVATKAG : 80
              R  IF  Ls 66  fnS  Hs  p  l  Dq  A  W  aL  G  V  t  D  i6  4

*          100         *          120         *          140         *          160
CGD7_M20 : KNGAPTVELLYSHYDVVPANNPDAWTNSPFFELTER---DC--RWYGRGAADCKGNLAMHLEALRLLEENGGTDLGLKVVVE : 151
Cure_M20 : INGAPTVELLYSHYDVVPANNPEAWTNSPFFELTER---EG--RWYGRGAADCKGNLAMHLEALRLLEENGGTDLGLKVVVE : 150
CJCB_M20 : ANGAPTVELLYSHYDVVPANNPEAWTNSPFFELTER---EG--RWYGRGAADCKGNLAMHLEALRLLEENGGTDLGLKVVVE : 151
Ccap_M20 : LAGAPTVELLYSHYDVVPAGDFAAWTNSPFFELTER---DC--RWYGRGAADCKGNLAMHLEALRLVREIENGGTDLGLKVVVE : 147
Cstr_tpd : VGDAPTVELLYSHYDVVPAGQNAFAVTNDELFLDER---EG--RWYGRGAADCKGNVAMHLEALRVMVQENGGTDLGLKVVVE : 155
Ctus_M20 : SDDAPTVELLYSHYDVVPAGEFAAWTNSPFFELTER---DC--RWYGRGAADCKGNVAMHLEALRLIDCAGFCPQANLIVLVE : 145
Ctub_M20 : DEGAPTVELLYSHYDVVPAGREAWESDPFFELTEREADGSTRWYGRGAADCKGNVAMHLEALRVDNNGGTQVNLRYLIE : 154
Cacc_M20 : EEGAPTVELLYSHYDVVPAGRAAWESDPFFELTERALAEAGTRWYGRGAADCKGNVAMHLEALRALAANGGTQVNLRYLIE : 154
Ccal_hyp : SPDAKVMVLYSHFDVVPAGPLDQVADSNEPFFELTERDAGEG--RWYGRGAADCKGNLAMHLEALRLVQSLGDTNINLTYVVF : 159
              Apt66LYsH5DVVPA  aW  Pf  lTer  G  RWYgRGAADCKGN6aMHL  ALR  6  Ggt  L  66E

*          180         *          200         *          220         *          240
CGD7_M20 : GSEELGGDGLGKLIIEKPEVFESEDAIIVADSGNVAACVPPTLTLRLGGAQVQVTVVETLIDNQVHSGSYGGCAPDAFAHALV : 231
Cure_M20 : GSEELGGDGLGKLIIEKPEVFESEDAIIVADSGNVAACVPPTLTLRLGGAQVQVTVVETLANQVHSGSYGGCAPDAFAHALM : 230
CJCB_M20 : GSEELGGDGLGKLIIEKPELFAADAILIADSGNVAACVPTVTLRLGGAQVQVTVVETLRGPHSGSYGGCAPDAFAHALV : 231
Ccap_M20 : GSEELGGDGLSLLIEQQPSLFSADVILIDVTGNVAVGCPPTLTLRLGGAQVQVTVVETLEGAVHSGSYGGCAPDAFAHALV : 227
Cstr_tpd : GSEELGGDGLGKLIIDANPELFTADVILICGGNVAVGIPPTLTLRLGGAQVQVTVVETLEGPVHSGSYGGCAPDAFAHALI : 235
Ctus_M20 : GSEELGG-GGLEKLIIEKPELEADVILIDVTGNVAVGPTLTLTLRLGGAQVQVTVVETLRGAVHSGSYGGCAPDPVAHALI : 224
Ctub_M20 : GSEERGG-EGLSGLIKDRPELFAADAILIADAGNAVAGCPPTLTLRLGGAQVQVTVVETLISAVHSGSYGGCAPDAFAKALM : 233
Cacc_M20 : GSEERGG-EGLSGLIKDRPELFAADAILIADAGNAVAGCPPTLTLRLGGAQVQVTVVETLISAVHSGSYGGCAPDAFAKALM : 233
Ccal_hyp : GSEELGG-GALSALIEKPELEADVILIDVTGNVAVGPTLTLTLRLGGAQVQVTVVETLTKGAVHSGSYGGCAPDAFAHALV : 238
              GSEE  GG  gL  LI  Pe6F  ad  I16ad  GN  A  G  PT6TT  LRG6aQ6  v  V  TL  HSG  5GG  APDA  AL6

*          260         *          280         *          300         *          320
CGD7_M20 : VIANSLFDDEHGHTIIDGV---DITAKWDEADRETFRTDITAVLGGVOLLGTVDD---EP--ADLVVAREFAVTMIGFTST : 303
Cure_M20 : VIANSLFDDEHGHTIIDGV---DITAKWDEADRETFRTDITAVLGGVOLLGTVDD---EP--ADLVVAREFAVTMIGFTST : 302
CJCB_M20 : VIANSLFDDEHGHTIIDGV---DITAKWDEADRETFRTDITAVLGGVOLLGTVDD---EP--ADLVVAREFAVTMIGFTST : 303
Ccap_M20 : RIANSLFDDEHGHTIIDGV---FALDKWEGEPYPRQFRDAGVLLGGVOLLGTVDD---EP--ALMVARFAVTMIGFTST : 299
Cstr_tpd : RLIDSLFDDEHGHTIIDGV---DITAKWDEADRETFRTDITAVLGGVOLLGTVDD---EP--ALMVARFAVTMIGFTSV : 307
Ctus_M20 : RLIDSLRFDASGHTIIDGIGDALITATWICQVYLAFAAFADAGVLLGGVOLLGGVSGTAGSPGVALMVARFAVTMIGFTST : 304
Ctub_M20 : RAIDSLSDDEHGHTIIDGV---DITAEWEGQVYFAFAFADAGVLLGGVOLLGGVIMGEDIYFAG--ATPVANMVARFAVTMIGFTST : 309
Cacc_M20 : RLIDSLFDDEHGHTIIDGV---DITSAHWQEQVYFAFAFADAGVLLGGVOLLGGVIMGSKPEG--ATPVANMVARFAVTMIGFTST : 309
Ccal_hyp : RLIDSLRFDDEHGHTIIDGV---DSSAVWEGAFYDPTERTDITAVLGGVOLLGGVIMGNGNP-----SSMLWSREFAVTMIGFTSI : 309
              31  De  GrTtIdG6  d  a  W  G  Y  e  FR  DA  6L  Gv  66G  d  a  66W  RPa636  GFTSt

*          340         *          360         *          380         *          400
CGD7_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITAKQNEQVHKSHPTPWGAHVVEVSGINEPFAATDSAGAAVALGECCLKDAYG : 383
Cure_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITAKQNEEHKIKAHTPWGAHVVEVSSVNEPFAATNPKAGAVATLGCCLKDAYG : 382
CJCB_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITADKQVEHKAHTPWGAHVVEVSGVNEPFAATDEEAFAVATLGCCLKDAYG : 383
Ccap_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITADKQVDRIRDIILAHTPWGAHVVDIAGVKNPASAEREAATLGCCLKDAYG : 379
Cstr_tpd : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITAKQVEHKAHTPWGAHVVEVSGVNEPFAATDEEAFAVATLGCCLKDAYG : 387
Ctus_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITAKQVHSHPTPWGAHVVEVDFINHGHAATDRSRAIALLGCCLKDAYG : 384
Ctub_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITADKQVHRAHTPWGAHVVEVLEAVAGSSTDEKFAAALLGCCLKDAYG : 389
Cacc_M20 : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITADKQVHRAHTPWGAHVVEVLEAVAGSSTDEKFAAALLGCCLKDAYG : 389
Ccal_hyp : PVAAEAVNAVNAFAAQENLNRVFAQCSAAEITADKQVHRAHTPWGAHVVEVLEAVAGSSTDEKFAAALLGCCLKDAYG : 389
              PV  eA  Na6  A  Aq  NLRVFAg  aa  a  6  h6  h  PWGAKV  v  6  N  F  31p  pA  lg  CL  aY

*          420         *          440         *          460         *          480
CGD7_M20 : A-----DKLSVVGSGSIPLITLQNTPEFAEIALYGVVEEMCGIHGVDSEVDETEIERIAVAEAAFLQRFYK : 451
Cure_M20 : A-----DELAVVSGSGSIPLITLQNTPEFAEIALYGVVEEQCGIHGVDSEVDETEIERIAVAEAAFLQRFYK : 450
CJCB_M20 : A-----DELAVVSGSGSIPLITLQNTPEFAEIALYGVVEELCGIHGVDSEVDETEIERIAVAEAAFLQRFYK : 451
Ccap_M20 : T-----DSVSHVSGSGSIPLITLQNTPEFAEIALYGVVEERACIHGVDSEVDETEIERIAVAEAAFLQRFYK : 447
Cstr_tpd : A-----EHLTVVGTGSGSIPLITLQNTPEFAEIALYGVVAIDIAANIHGVDESVDTEIEHVAIAEAFVFLTYGK : 455
Ctus_M20 : EGAGSDPTVTTPTDLDVTVGSGSIPLITLQNTPEFAEIALYGVVDENLACIHSVDESVEEYETIEHVAIAEAFVFLTYGK : 464
Ctub_M20 : -----TDTAAQMGSGSIPLITLQNTPEFAEIALYGVVEEKCTIHSANESVDTEIEHVAIAEAAFLQRFYK : 455
Cacc_M20 : -----KDTAAQMGSGSIPLITLQNTPEFAEIALYGVVEEKCTIHSANESVDTEIEHVAIAEAAFLQRFYK : 455
Ccal_hyp : N-----KETVTGSGSGSIPLITLQNTPEFAEIALYGVVEELTVIHSANESVDEHEIRVATFAEAFVFLTYGK : 457
              G  GGSIPL  Lt  Lq  P  Ae  A65Gveep  IH  1ESV  P  Eie  6a  AEA  FL  r

```

Abbildung 44: Multiple Sequence Alignment der mittels BLASTX identifizierten Referenzsequenzen aus der GenBank-Datenbank zu cDNA-Sequenzen mit Homologie zur TpdA-Dipeptidase. CGD7: *Corynebacterium* sp. GD7 (*Corynebacterium ihumii*), Cure: *Corynebacterium ureicelerivorans*, CJCB: *Corynebacterium* sp. JCB (*Corynebacterium jeddahense*), Ccap: *Corynebacterium capitovis*, Cstr: *Corynebacterium striatum*, Ctus: *Corynebacterium tuscaniense*, Ctub: *Corynebacterium tuberculostearicum*, Cacc: *Corynebacterium accolens*, Ccal: *Corynebacterium callunae*; M20: Peptidase M20, tpd: TpdA-Dipeptidase, hyp: hypothetisches Protein.

### 3 Funktionen mit Korrelation zum Körpergeruch

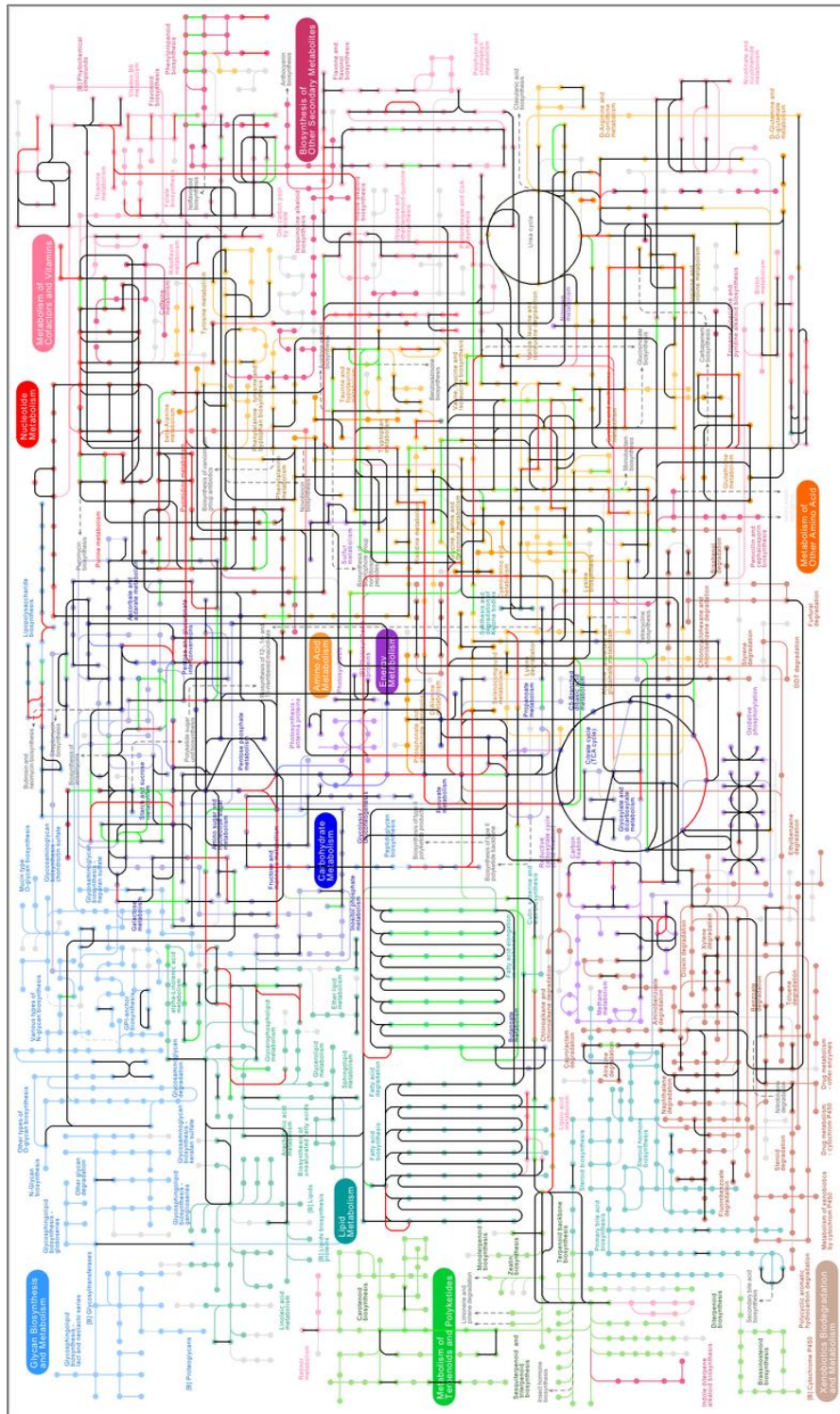


Abbildung 45: Funktionale Zuordnungen innerhalb der Metatranskriptom-Datensätze der Achselhöhlen-Proben mit Korrelation zum Körpergeruch auf Grundlage des KO-Annotationsystems. Die Darstellung basiert auf der mittels KEGG-MAPPER erstellten Stoffwechselkarte „Metabolic Pathways“. Grün: positive Korrelation ( $r^2 \geq 0,5$ ), schwarz: geringfügige Korrelation ( $0,5 \geq r^2 \geq -0,5$ ), rot: negative Korrelation ( $r^2 \leq -0,5$ ).

## Danksagung

Zuerst möchte ich mich bei meinem Ehemann Tobias bedanken für die großartige Unterstützung während dieser nicht immer einfachen Zeit.

Ich bedanke mich bei Herrn apl. Prof. Dr. Andreas Tauch für die Möglichkeit, diese Dissertation anzufertigen und für die Betreuung.

Dem CLIB-Graduiertencluster danke ich für die Bereitstellung des Stipendiums. Insbesondere danke ich den Mitgliedern für eine interessante Zeit.

Ich bedanke mich bei Herrn Dr. Brauer für die Bereitstellung der Ohr-Probe sowie bei Herrn Dr. Rainer Simmering und Herrn Stefan Evers für die Bereitstellung der Achselhöhlen-Proben.

Weiterhin bedanke ich mich ganz herzlich bei Annika Winkler und Andreas Albersmeier für die Durchführung der Sequenzierungen am CeBiTec und für hilfreiche praktische Tipps im Laboralltag.

Ein ganz herzlicher Dank gilt Daniel Wibberg für die tatkräftige Unterstützung bei der Sequenzdaten-Analyse. Die fachlichen Gespräche mit ihm und mit Felix Eickmeier waren von großer Bedeutung für die Auswertung und Interpretation der Daten.

Ein ebenso großer und herzlicher Dank geht an Christina Ander für die Entwicklung und Bereitstellung der *METRANS-Pipeline* unter der Betreuung von Herrn Prof. Dr. Jens Stoye. Die Zusammenarbeit war sehr angenehm und spannend und lieferte entscheidende Beiträge für die Sequenzdatenanalyse.

Bei meinen lieben Büro-Kolleginnen Katharina Pfeifer-Sancar, Kathrin Petri, Almut Mentz und Johanna Milse bedanke ich mich für einen schönen und interessanten Alltag, die vielen Fachgespräche und die schöne Zeit außerhalb der Uni. Das Gleiche gilt für Armin Neshat. Ich hoffe, wir bleiben noch sehr lange in Kontakt.

Ganz herzlich bedanke ich mich bei den „Corynes“, bei den Kollegen aus anderen Arbeitsgruppen und bei Magdalena Kröber für die schöne Zeit am CeBiTec und auch außerhalb.

Der größte Dank gilt meiner lieben Mutter dafür, dass sie mich zum Studium ermutigt und mich währenddessen stets unterstützt hat.

## Erklärung

Die Anfertigung dieser Dissertation mit dem Titel „Metatranskriptom-Analysen der mikrobiellen Lebensgemeinschaften des menschlichen Ohres und der Achselhöhle“ erfolgte unter Betreuung von apl. Prof. Dr. Andreas Tauch in der Arbeitsgruppe „Medizinische Mikrobiologie und Genomik“ am Centrum für Biotechnologie (CeBiTec) der Universität Bielefeld.

Hiermit erkläre ich, dass ich die vorliegende Dissertation selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel verwendet habe. Die Stellen der Dissertation, die anderen Quellen im Wortlaut oder dem Sinn nach entnommen wurden, sind durch Angabe der Herkunft kenntlich gemacht. Dies gilt auch für Zeichnungen, Skizzen, bildliche Darstellungen sowie für Quellen aus dem Internet.

Weiterhin erkläre ich, dass die vorliegende Dissertation weder vollständig noch in Auszügen einer anderen Fakultät mit dem Ziel vorgelegt worden ist, einen akademischen Titel zu erwerben. Ich bewerbe mich hiermit erstmalig um den Doktorgrad der Naturwissenschaften der Universität Bielefeld.

Bielefeld, den

---

Eugenie Fredrich-Vahle