

**Aus dem Department für Diagnostische Labormedizin der  
Universität Tübingen**

**Institut für Medizinische Mikrobiologie und Hygiene**

**Untersuchung des SNP-Abstands von nosokomialen  
*Pseudomonas aeruginosa*- Stämmen zur Beurteilung  
des Nutzens eines SNP-Grenzwertes zur  
Unterscheidung zwischen Transmission und Mutation**

**Inaugural-Dissertation**

**Zur Erlangung des Doktorgrades  
der Medizin**

**Der medizinischen Fakultät  
Der Eberhard Karls Universität  
zu Tübingen**

**vorgelegt von**

**Madeleine Susan Desch**

**2021**

Dekan: Professor Dr. B. Pichler

1. Berichterstatter: Professor Dr. M. Willmann

2. Berichterstatter: Privatdozentin Dr. I. Brecht

Tag der Disputation: 07.09.2021

# Inhaltsverzeichnis

<b>ABKÜRZUNGSVERZEICHNIS.....</b>	<b>IV</b>
<b>1 EINLEITUNG.....</b>	<b>1</b>
1.1 Bedeutung von <i>Pseudomonas aeruginosa</i> als nosokomialer Krankheitserreger.....	1
1.1.1 Pathogenität von <i>P. aeruginosa</i> .....	1
1.1.2 Risikokollektiv für <i>P. aeruginosa</i> und Klinik .....	3
1.2 Genetik und Phylogenie von <i>P. aeruginosa</i> .....	5
1.2.1 Pangenom, Core-Genom und akzessorisches Genom .....	5
1.2.2 Mutationsrate und SNP-Abstand.....	7
1.3 Next-Generation Sequencing (NGS) .....	8
1.4 Ausbruchanalysen .....	10
1.5 Zielsetzung und Hypothesen.....	11
<b>2 MATERIAL UND METHODEN .....</b>	<b>14</b>
2.1 Studienaufbau und erhobene Daten .....	14
2.1.1 Herkunft der Daten.....	14
2.1.2 Einschlusskriterien für Patienten und Stämme (ID-Vergabe) .....	15
2.2 Mikrobiologische Diagnostik.....	16
2.2.1 Kultur, Identifikation und Antibiotikaresistenztestung .....	16
2.2.2 Next-Generation Sequencing (NGS).....	17
2.3 Phylogenetische Analyse .....	18
2.3.1 Einteilung in Patientencluster .....	18
2.4 Berücksichtigung des Zeitfaktors.....	20
2.5 Statistische Analyse.....	20
<b>3 ERGEBNISSE.....</b>	<b>22</b>
3.1 Eingeschlossene Patienten und Stämme.....	22
3.1.1 Ausgangslage und Ausschlussverfahren .....	22

3.1.2	Demographische Daten der zugrunde liegenden Studienpopulation der Patientencluster .....	26
3.1.3	Clustergröße der Patientencluster .....	29
3.1.4	Probenmaterialien der Patientencluster .....	30
3.1.5	Tabellarische Übersicht über die demographischen Daten, die Clustergröße und die Probenmaterialien der Patientencluster .....	32
<b>3.2</b>	<b>Darstellung des SNP-Abstands in den Patientenclustern .....</b>	<b>33</b>
3.2.1	SNP-Abstand ohne Berücksichtigung des Zeitfaktors .....	33
3.2.2	SNP-Abstand mit Berücksichtigung des Zeitfaktors .....	36
<b>3.3</b>	<b>Unterschiede des SNP-Abstands zwischen gleichem und ungleichem Probenmaterial.....</b>	<b>39</b>
3.3.1	SNP-Abstand ohne Berücksichtigung des Zeitfaktors .....	39
3.3.2	SNP-Abstand mit Berücksichtigung des Zeitfaktors .....	51
<b>3.4</b>	<b>Indexstamm-Verbindungen .....</b>	<b>63</b>
3.4.1	SNP-Abstand ohne Berücksichtigung des Zeitfaktors .....	64
3.4.2	SNP-Abstand mit Berücksichtigung des Zeitfaktors .....	64
3.4.3	Tabellarische Übersicht der deskriptiven Daten und des Mann-Whitney-U-Tests der Indexstamm-Verbindungen .....	65
<b>3.5</b>	<b>Korrelationsanalysen .....</b>	<b>66</b>
3.5.1	Korrelation des SNP-Abstands mit der Clustergröße .....	66
3.5.2	Korrelation des SNP-Abstands mit dem maximalen zeitlichen Abstand .....	68
3.5.3	Korrelation des SNP-Abstands mit der Anzahl der Probenmaterialien	70
<b>4</b>	<b>DISKUSSION .....</b>	<b>73</b>
<b>5</b>	<b>ZUSAMMENFASSUNG .....</b>	<b>81</b>
<b>6</b>	<b>TABELLENVERZEICHNIS .....</b>	<b>83</b>
<b>7</b>	<b>ABBILDUNGSVERZEICHNIS .....</b>	<b>84</b>
<b>8</b>	<b>LITERATURVERZEICHNIS .....</b>	<b>87</b>
<b>9</b>	<b>ERKLÄRUNG ZUM EIGENANTEIL .....</b>	<b>95</b>

**10 DANKSAGUNG..... 96**

## Abkürzungsverzeichnis

Abb.	Abbildung
blastp	basic local alignment search tool protein
bp	Basenpaare (engl. Base Pairs)
bzw.	beziehungsweise
ca.	circa
DNA	Desoxyribonukleinsäure (engl. Deoxyribonucleic Acid)
engl.	englisch
etc.	et cetera
GC-Gehalt	Guanin- und Cytosin-Gehalt
ID	Identifikationsnummer
IQR	Interquartilsrange
Kap.	Kapitel
95%KI	95%-Konfidenzintervall
MALDI-TOF	Matrix-assisted Laser Desorption Ionization Time -off Light
Mbp	Megabasenpaare (engl. Mega Base Pairs)
MDR	multiresistent (engl. Multi Drug-resistant)
N	Anzahl
NGS	Next-Generation Sequencing
p	Signifikanzniveau
<i>P. aeruginosa</i>	<i>Pseudomonas aeruginosa</i>
PCR	Polymerase-Kettenreaktion (engl. Polymerase Chain Reaction)
R	resistent (Antibiogramm)
S	sensibel (Antibiogramm)
SBH	Sequencing-by-Hybridization

SBS	Sequencing-by-Synthesis
SD	Standard Deviation
SNP	Einzelnukleotid-Polymorphismus (engl. Single Nucleotide Polymorphism)
Tab.	Tabelle
U	Mann-Whitney-U-Statistik
VAP	Beatmungsgerät-assoziierte Pneumonie (engl. Ventilator-associated Pneumonia)
vs.	versus
XDR	extrem Antibiotika-resistent (engl. Extensive Drug-resistant)
Z	Z-Statistik
z.B.	zum Beispiel
$\rho$	Spearman-Korrelationskoeffizient

# 1 Einleitung

## 1.1 Bedeutung von *Pseudomonas aeruginosa* als nosokomialer Krankheitserreger

Trotz der medizinischen Fortschritte in Prävention, Diagnostik und Therapie stellen nosokomiale Infektionen mit infolgedessen ansteigenden Morbiditäts- und Mortalitätsraten das Gesundheitssystem weltweit vor große Herausforderungen. [1], [2]

*Pseudomonas aeruginosa* (*P. aeruginosa*), einer der häufigsten Verursacher von nosokomialen Infektionen unter den gramnegativen Bakterien, bildet dabei mit jährlich mehr als zwei Millionen infizierten Patienten und 90.000 Toten einen signifikanten Anteil und verursacht als vierthäufigster Krankheitserreger in europäischen Krankenhäusern etwa 10 % aller nosokomialen Infektionen. [2]–[6] Sowohl seine Fähigkeit ökologische Nischen zu besiedeln, als auch die schwerwiegende Behandlung von *P. aeruginosa*-Infektionen mit damit einhergehenden längeren Krankenhausaufenthalten und einer höheren Komplikations- bzw. Mortalitätsrate sowie die ansteigenden Resistenzraten erweisen sich dabei als eine ernst zu nehmende medizinische und ökonomische Belastung für das Gesundheitssystem aufgrund der Notwendigkeit von weitreichenderen Diagnostik- und Therapiemaßnahmen. [1], [4], [7]

### 1.1.1 Pathogenität von *P. aeruginosa*

Bei *P. aeruginosa* handelt es sich um ein bis zu 3 µm großes, gramnegatives Bakterium, dem aufgrund seines gehäufteten Vorkommens im Zusammenhang mit Infektionen im Menschen die größte Aufmerksamkeit unter der Gattung der Pseudomonaden zuteilwird. [8], [9]

Seine Fähigkeit sich an neue bzw. sich verändernde Umweltbedingungen anzupassen, spielt für die Besiedlung einer Vielzahl von Lebensräumen und -repertorien eine wichtige Rolle und macht sein ubiquitäres Auftreten in menschlichen, tierischen und pflanzlichen Wirten sowie in der Umwelt möglich. [10]–[13] Dieser genetischen Flexibilität und Anpassungsfähigkeit liegen das



über 6 Megabasenpaaren (Mbp, engl. Mega Base Pairs) große Genom und der darin enthaltene hohe Anteil an Regulator-Genen zugrunde. [11], [14], [15] Durch genetische Veränderungen erhält *P. aeruginosa* beispielsweise die Fähigkeit, alternative Quellen zur Energiegewinnung zu nutzen und sich den antimikrobiellen Mechanismen des Wirts zu entziehen. [16] Während somit die ursprüngliche Population diversen Selektionsdrücken innerhalb des Wirts nicht standhalten kann, ist dies den Subpopulationen durch Veränderungen im Genom und den damit einhergehenden Anpassungsvorgängen möglich. [15], [16]

Für den Infektionsprozess sind vor allem Virulenzfaktoren von Bedeutung. Diese Virulenzfaktoren ermöglichen eine Interaktion mit dem Immunsystem des Wirts, die es *P. aeruginosa* erleichtert, sich vor der Detektion und Eradikation durch die Immunabwehr des Wirts zu schützen. [16] Beispielhaft zu nennen ist an dieser Stelle das Quorum-Sensing, das je nach Zelldichte die Genexpression innerhalb der Population und damit die Produktion von Toxinen, Elastasen und des Pigments Pyocyanin reguliert. Ferner ist das Typ III-Sekretionssystem erwähnenswert. Dieses ermöglicht die direkte Sekretion von Toxinen in die Wirtszelle. Auch bedeutsam ist das Lipopolysaccharid-O-Antigen auf der bakteriellen Oberfläche, welches vor einer Lyse durch das Komplementsystem schützt. [17]–[22] Weitere Beispiele für Virulenzfaktoren sind die Produktion von Toxin A, welches durch die Hemmung der Proteinsynthese eine hohe Toxizität aufweist. Ebenso relevant ist die Bildung der Pigmente Pyocyanin und Fluoreszein. Diese sind in der Lage, die Immunabwehr des Wirts zu unterdrücken. [8], [16], [23]–[25]

Neben den Virulenzfaktoren ist die Bildung von Biofilmen für die lange Persistenz von *P. aeruginosa* an Oberflächen oder im Wirtsorganismus verantwortlich. [16] Bei Biofilmen handelt es sich um einen Verband von Bakterienzellen, die von einer Exopolysaccharid-Matrix umgeben sind und sich auf festen Oberflächen, wie beispielsweise auf Kathetern, Gelenkprothesen oder Kontaktlinsen, bilden. [3], [20], [22], [26] Bakterienzellen innerhalb eines Biofilms sind somit vor Hitze, Kälte, pH-Wert-Veränderungen und der Wirkung von chemischen Mitteln, vor allem jedoch gegenüber dem Angriff von

antimikrobiellen Mitteln besser geschützt. [22], [26] Als Grund hierfür werden vor allem die Exopolysaccharid-Matrix als initiale Barriere, der heruntergefahrte Stoffwechsel, das damit verlangsamte Wachstum der Zellen innerhalb des Biofilms sowie auch eine Reihe von anderen Faktoren wie die Expression von Efflux-Pumpen und Porinen diskutiert. Die genaue Wirkungsweise von Biofilmen ist jedoch noch nicht gänzlich verstanden. Fakt ist jedoch, dass die Beseitigung von Biofilmen aufgrund einer bis zu 1000fach höheren Antibiotika-Resistenz der darin lebenden Bakterien oft sehr schwierig ist. Die daraus resultierenden Komplikationen für Patienten konfrontieren das Gesundheitssystem Jahr für Jahr mit hohen Kosten. [26]

Des Weiteren machen die intrinsische Resistenz gegenüber einer Vielzahl von Antibiotika und die Fähigkeit, weitere Resistenzen zu erwerben, das Bakterium so gefährlich und erschweren die Auswahl einer angebrachten antipseudomonalen Therapie. [2], [4], [11], [27] Die intrinsische antibiotische Resistenz kommt hierbei unter anderem durch eine geringe Permeabilität der äußeren Zellmembran, dem Vorhandensein von genetischem Material zur Bildung von Efflux-Pumpen, welche die Medikamente aus der Zelle schleusen, und durch die Produktion einer induzierbaren AmpC-Cephalosporinase zustande. [4], [27] Der Erwerb von sogenannten extrinsischen Resistenzen hingegen findet vor allem durch Mutationen oder horizontalen Gentransfer statt. Dies hat unter anderem die Hochregulation von oben genannten Efflux-Pumpen zur Folge. [4] Das Auftreten von mehreren dieser erworbenen Resistenzmechanismen innerhalb eines Bakteriums führt zu sogenannten multiresistenten (MDR, engl. Multi Drug-resistant) oder extrem Antibiotika-resistenten (XDR, engl. Extensive Drug-resistant) Stämmen. Diese gehen mit einer erhöhten Mortalität, einer Verlängerung des Krankenhausaufenthalts und ansteigenden Krankenhauskosten als Folge einer häufig inadäquaten Antibiotikatherapie einher. [4], [11]

### **1.1.2 Risikokollektiv für *P. aeruginosa* und Klinik**

Während *P. aeruginosa* eine der drei häufigsten Ursachen für opportunistische Infektionen im Menschen ist, bleibt eine Kolonisation für Gesunde meist ohne Konsequenz. [14], [28] Als seltener Bestandteil der menschlichen Mikroflora in

hauptsächlich hospitalisierten Patienten wurde eine Besiedlung mit *P. aeruginosa* hierbei am häufigsten im Darmtrakt (2,6 – 24 %) und seltener in Armbeugen, Perineum, Respirationstrakt und Rachen (0 – 6,6 %) nachgewiesen. [11], [29] Sobald es jedoch zu einer Beeinträchtigung der Immunkompetenz kommt, kann eine Kolonisation innerhalb kürzester Zeit in eine Infektion mit *P. aeruginosa* übergehen. Laut Hoang *et al.* scheint hierbei vor allem die Kolonisation im Darm eine wichtige endogene Infektionsquelle zu sein. [7] Darüber hinaus erhöhen vor allem der Durchbruch von physikalischen Barrieren, wie beispielsweise durch intravenöse Zugänge, Harnkatheter und Endotrachealtuben, das Risiko einer Infektion. [30] Besonders gefährdet sind aus diesem Grund vor allem Patienten mit Aufenthalt auf einer Intensivstation, da dort einerseits invasive Verfahren (Anlage eines Zentralen Venenkatheters, Bronchoskopien, Endoskopien, Intubation zur mechanischen Beatmung etc.) häufiger eingesetzt werden als auf anderen Stationen. Zum anderen handelt es sich um ein meist kritisch krankes Patientenkollektiv mit langer Hospitalisierungsdauer. [1], [4], [7], [11], [29], [31] Zum Risiko-Patientenkollektiv gehören hierbei vor allem Menschen mit Immunschwäche (Neutropenie, medikamentöse Immunsuppression, AIDS etc.). Ebenso dazu zählen Patienten mit großflächigen Brandwunden sowie Patienten mit chronisch pulmonalen, metabolischen, hämatologischen oder malignen Erkrankungen. [1], [7], [32], [33] Eine Sonderstellung nehmen Patienten mit Mukoviszidose ein, da im Erwachsenenalter die große Mehrheit (80 %) der an Mukoviszidose-Erkrankten an einer chronischen Infektion mit *P. aeruginosa* leidet und *P. aeruginosa* in diesem Kollektiv sogar eine der Haupttodesursachen mit einer Sterblichkeitsrate von 50 % im Zeitraum von 5 Jahren nach Beginn der chronischen Infektion darstellt. [20], [22], [29], [33], [34]

Weitere exogene Faktoren, die das Risiko einer Infektion mit *P. aeruginosa* erhöhen, sind zudem auch der Kontakt mit anderen kolonisierten Patienten (über Infektionsträger) oder die Aufnahme von kontaminierter Nahrung bzw. kontaminiertem Wasser. [8], [9], [35] Auch die Übertragung durch medizinisches Personal ist relevant und vor allem bei einer Infektion mit XDR-*P. aeruginosa*-Stämmen als einer der wichtigsten Risikofaktoren einzustufen. [35], [36] Nicht

zuletzt hängt auch der Einsatz von Ciprofloxacin, welches häufig bei neutropenen Patienten als Antibiotikaprophylaxe eingesetzt wird, mit einer Risikoerhöhung für eine Infektion mit *P. aeruginosa* zusammen. [35] Da sich der Nass- und Pfützen-Keim auch gut in feuchter Umgebung ansiedelt und durch die Bildung von Biofilmen dort lang persistieren kann, stellen auch Wasserhähne, Waschbecken, Beatmungs-Geräte, Krankenhausnahrung, Infusionen und andere Umgebungs-Gegenstände mit feuchtem Milieu eine wichtige Expositionsquelle dar. [8], [31], [37]

*P. aeruginosa*-assoziierte Krankheitsbilder sind an erster Stelle vor allem die Beatmungsgerät-assoziierte Pneumonie (VAP, engl. *Ventilator-associated Pneumonia*) und Brandwunden-Infektionen, die mit einer hohen Mortalität einhergehen. [11] Häufig sind außerdem Infektionen der Blutbahn, des Harntraktes, der Haut, von Operationsgebieten oder allgemein von lokalen Wunden. [1], [11], [14], [20], [29]

## **1.2 Genetik und Phylogenie von *P. aeruginosa***

Im Hinblick auf das Ziel dieser Arbeit und zum besseren Verständnis der den Ausbruchsanalysen zugrunde liegenden Verwandtschaft und Evolution von *P. aeruginosa*-Stämmen sind Kenntnisse zum Genom, der Mutationsrate und zur genetischen Distanz zweier *P. aeruginosa*-Stämme essenziell und werden im Folgenden erläutert.

### **1.2.1 Pangenom, Core-Genom und akzessorisches Genom**

Bezüglich des Genoms lassen sich verschiedene Begrifflichkeiten wie das Pan-Genom, das Core-Genom und das akzessorische Genom unterscheiden. Unter dem Pangenom versteht man den kompletten Satz an Genen einer Bakterienart oder eines -stamms. Angesichts neuer Studienerkenntnisse wächst die Gesamtheit der Gene eines Pangenoms mit der Anzahl an analysierten Stämmen. [38] Während jeder einzelne Stamm eine durchschnittliche Anzahl von etwas mehr als 6000 Genen enthält, ist das Pangenom mit einer Anzahl von 16.000 bis 19.000 nicht-redundanten Genen bei 99 und 181 analysierten Stämmen bzw. bis zu über 54.000 Genen bei 619 analysierten Stämmen, somit deutlich größer. [38]–[40]

Jedes Pangenom setzt sich wiederum aus einem Core-Genom und einem akzessorischen Genom zusammen. Beim Core-Genom handelt es sich um die Gensequenzen, die in 100 % der Genome der untersuchten *P. aeruginosa*-Stämme enthalten sind. Diese kodieren hauptsächlich für lebenswichtige Funktionen wie zur Aufrechterhaltung von Zellerhalt und -teilung und für die intrinsische Infektiosität und Resistenzen. [10], [38]–[40] Die Größe des Core-Genoms sinkt mit der Anzahl an analysierten Stämmen und variiert in verschiedenen Studien zwischen 665 und 3814 Genen. [38], [39] Dieses macht somit weniger als ein Viertel der Größe des Pangenoms aus, bei einer Anzahl von 665 Core-Genen und einem Pangenom von 54.272 Genen sogar nur 1 %. [38]–[40] Dies ist nicht zu verwechseln mit dem Anteil des Core-Genoms am Genom jedes einzelnen Stamms. Dieser liegt mit bis  $\geq 90$  % deutlich höher und spricht für eine hohe Konservierung der Gene innerhalb eines *P. aeruginosa*-Stamms. [10], [41], [42] Die übrigen Gensequenzen des Pangenoms, die nicht Teil des Core-Genoms sind, bezeichnet man als akzessorisches Genom, welches für die Faktoren der Virulenz und anderer medizinisch wichtiger Dinge kodiert. Dieser Teil des Pangenoms ist somit primär für die genetische Variabilität zwischen den unterschiedlichen *P. aeruginosa*-Stämmen verantwortlich. Diese genetischen Unterschiede gehen unter anderem auf Genumlagerungen, Austausch von Plasmiden oder Insertionen und Deletionen zurück. [43] Weiter unterteilen lässt sich das akzessorische Genom in einmalige Gene, die in nur einem einzigen analysierten Stamm vorkommen. Diese sind von flexiblen Genen, die in mehr als einem, aber nicht in allen analysierten Stämmen vorhanden sind, abzugrenzen. [40]

Auf molekularer Ebene unterscheiden sich die Core-Genome von nicht miteinander verwandten *P. aeruginosa*-Stämmen in weniger als 1 % der DNA-Sequenzen. Die Sequenzdiversität zwischen verwandten Stämmen ist zwar unbekannt, aber als noch geringer als die von nicht verwandten Stämmen anzunehmen. [10] In Anbetracht dessen und der Tatsache, dass das akzessorische Genom sehr variabel ist, ist die alleinige Verwendung des Core-Genoms zur Untersuchung der genetischen Distanz zweier Stämme – wie auch in dieser Arbeit durchgeführt – gut geeignet.

### 1.2.2 Mutationsrate und SNP-Abstand

Betrachtet man die Anpassungsfähigkeit von *P. aeruginosa* genauer, dann geht ein großer Teil dieser adaptiven Evolution auf das Auftreten von Mutationen innerhalb des Genoms zurück. Obwohl es diverse Erkenntnisse zu den Auswirkungen dieser Mutationen gibt, ist noch wenig über die Rate und Häufigkeit der verschiedenen Arten von Mutationen, sowie die Variabilität des Mutationsprozesses im gesamten Genom bekannt. [44]

Bei den Mutationen handelt es sich um Basenfehlpaarungen im Genom, die in über 99 % der Fälle erkannt und durch ein Mismatch-Reparatur-System korrigiert werden. Ein kleiner Teil dieser sogenannten Prä-Mutationen bleibt jedoch unerkannt und führt folglich zu bleibenden Mutationen. Dabei hat auch der natürliche Selektionsprozess erheblichen Einfluss, welche Mutationen erhalten bleiben und welche entfernt werden. [44] Die Mutationen sind dabei nicht zufällig über das Genom verteilt, sondern hängen von mehreren Faktoren ab und werden gehäuft in Genen festgestellt, die für Virulenzfaktoren wie Efflux-Pumpen und Quorum-Sensing, sowie Beweglichkeit und Chemotaxis kodierten. [44], [45]

Generell umfassen diese Mutationen neben Einzelnukleotid-Insertionen und -Deletionen die sogenannten Einzelnukleotid-Polymorphismen (SNPs, engl. Single Nucleotide Polymorphism). Diese sind eine sehr häufige Ursache für die genetische Variation in der Sequenz von Genomen. [43], [46] Sie treten ca. alle 100 – 300 Basenpaare (bp, engl. Base Pairs) in kodierenden sowie in nicht-kodierenden Genomabschnitten auf und geben Aufschluss über die genetische Distanz und damit Verwandtschaft der untersuchten Stämme. [47] Je weniger SNPs zwischen zwei untersuchten Genomen vorhanden sind, desto geringer deren genetische Distanz und desto enger verwandt sind die dazu gehörigen analysierten Stämme.

Durch die über die Zeit auftretenden Mutationen akkumulieren diese SNPs, wobei für verschiedene Bakterienspezies unterschiedliche Raten festgestellt wurden. Während beispielsweise in *Staphylococcus aureus*-Stämmen ca. 1 SNP in sechs Wochen akkumuliert, liegt die Mutationsrate bei *P. aeruginosa* mit 1,18 SNPs in einem Jahr deutlich niedriger. [37], [48] Bemerkenswert höher

liegen kann diese Rate jedoch in hypermutierenden Stämmen, die im Folgenden als Hypermutator-Stämme bezeichnet werden. Die dadurch höhere Wahrscheinlichkeit für gewinnbringende Mutationen, sorgt für eine noch bessere Anpassungsfähigkeit dieser Stämme. [44] Im Fall von *P. aeruginosa* wurde eine große Anzahl an Hypermutator-Stämme vor allem in Patienten nachgewiesen, die an Mukoviszidose erkrankt waren. [44], [49]

### **1.3 Next-Generation Sequencing**

Nicht nur das theoretische Verständnis der Vorgänge, die sich auf molekularer Ebenen im Genom von *P. aeruginosa* abspielen, sondern auch die Methode, die es überhaupt möglich macht, die genetische Zusammensetzung von isolierten Stämmen zu bestimmen, soll an dieser Stelle genauer beleuchtet werden.

Die Möglichkeit der kompletten Sequenzierung bakterieller Genome ist eine wichtige Technologie, ohne welche die heutzutage durchgeführte Forschung auf molekularer Ebene nicht in gleichem Maße möglich wäre. Den Anfang dieses wissenschaftlichen Fortschritts bildet die von Frederick Sanger 1977 entwickelte Sanger-Technik, die auf einer Kombination von Desoxy- und Didesoxy-markierten Kettenabbruch-Nukleotiden beruht und für welche er 1980 auch den Chemie-Nobelpreis erhielt. [45] Die sogenannten Methoden des Next-Generation Sequencings (NGS) haben die Sequenziertechnik in den letzten Jahren hinsichtlich des Durchsatzes, der Kosten, der Sequenzlänge und der Fehlerrate perfektioniert und so zu einer Revolution dieser Technologie mit Einsparung von Zeit und Kosten geführt. [45] Vor allem der Einsatz von Fluoreszenz-markierten Nukleotiden war hierbei eine bahnbrechende Entwicklung. [50]

Bei der Durchführung von NGS kann man generell zwei mögliche Sequenzierungsarten unterscheiden: die Sequencing-by-Hybridization- (SBH) und die Sequencing-by-Synthesis (SBS)-Methode. An dieser Stelle soll nur auf die häufigere und hier angewandte SBS-Methode eingegangen werden. Bei dieser werden folgende vier Schritte nacheinander durchgeführt: Desoxyribonukleinsäure (DNA, engl. Deoxyribonucleic Acid) -Fragmentation, Bibliotheksvorbereitung (engl. Library Preparation), Sequenzierung und

Bioinformatische Analyse. Im Folgenden werden die einzelnen Schritte kurz erklärt.

**DNA-Fragmentation:** Die zu sequenzierende DNA wird mittels mechanischer, enzymatischer Methoden oder schallgesteuert in 100-300 bp lange DNA-Fragmente geteilt. Aus diesen werden dann für die Analyse relevante Segmente mithilfe von komplementären Sonden (sogenannter Hybridization-Capture-Assay) herausgefiltert oder durch die Polymerase-Kettenreaktion amplifiziert (PCR, engl. Polymerase Chain Reaction) (sogenannter Amplicon-Assay). [51]

**Bibliotheksvorbereitung (Library Preparation):** Zur Probenidentifikation erhält jede zu sequenzierende DNA einen spezifischen Index und Sequenzierungsadapter an den beiden DNA-Enden, die später den Sequenzierungsprimern die Anbindung ermöglichen. [51]

**Sequenzierung:** Die Bibliothek wird auf eine Sequenzierungsmatrix mit zu den Sequenzierungsadaptoren komplementären Oligonukleotiden gebracht. Im Fall der in dieser Studie verwendeten Illumina-Technologie handelt es sich bei der Sequenzierungsmatrix um eine Durchflusszelle aus Glas. Dort findet die sogenannte Bridge-Amplifikation statt, also die Herstellung von tausenden von DNA-Kopien in unmittelbarer Nähe des Originals in einer isothermen Reaktion. Dies führt zur Bildung von Millionen von (Einzelstrang-)DNA-Fragment-Clustern. Die Durchflusszelle wird daraufhin mit spaltbaren fluoreszierenden Nukleotiden und für die DNA-Polymerisation erforderlichen Reagenzien geflutet, sodass zu jedem DNA-Fragment ein komplementärer Strang gebildet wird. Beim Einbau der einzelnen Basen wird - durch einen Laser angeregt - ein je nach Base, unterschiedlich farbiges Licht emittiert und detektiert. Dieser Vorgang wird in jeder Polymerisationsrunde wiederholt. [45], [50], [51]

**Bioinformatische Analyse:** Zuletzt werden die bei der Sequenzierung entstandenen Reads – kurze DNA-Sequenzen – weiter analysiert. Ein gerade in der Bakteriologie beliebtes Verfahren ist die Assemblierung dieser kurzen Reads zu einem gesamten oder teilweise fragmentarischen Genom. [51]

Diese NGS-Methode ist gegenüber der Sanger-Methode eine weitaus schnellere Technik des Sequenzierens und liefert die für phylogenetische



Analysen notwendigen genetischen Informationen. [45] Die im Gegensatz zur Sanger-Methode initial höhere Fehlerrate wird durch die enorme Anzahl an parallelen Sequenzierungen und durch die Übereinstimmung in den Sequenzen ausgeglichen. Nichtsdestotrotz kann es zu Sequenzkontextfehlern kommen, die durch den Einsatz unterschiedlicher Plattformen minimiert werden können. [50] Die HiSeq2500 Plattform (Illumina, San Diego, USA) wird dabei aufgrund ihrer hohen Laufrate und geringen Fehlerrate von nur 0,1 % häufig als Sequenziergerät verwendet. [9], [45]

Zusammenfassend ist das NGS eine moderne Technologie, die den universellen Grundbaustein für die Analyse von allen Arten genetischen Materials und die in der vorliegenden Studie durchgeführten Untersuchung des genetischen Abstands von *P. aeruginosa*-Stämmen darstellt.

#### **1.4 Ausbruchanalysen**

Ausbrüche mit *P. aeruginosa* stellen Krankenhäuser vor eine große Herausforderung. Nicht nur die hohen ökonomischen Kosten, sondern auch die ansteigenden Antibiotikaresistenzraten und die hohe Mortalität bei vor allem Blutstrominfektionen mit *P. aeruginosa* und die weit verbreitete Kontamination der Umwelt legen die Bezeichnung als Problemkeim nahe. [37]

Einige Studien kamen zu der Erkenntnis, dass die durch NGS festgestellte Anzahl an SNPs zweier Stämme ein sinnvolles Vergleichsinstrument zur Bestimmung des genetischen Abstands und im Folgenden zur Analyse von Ausbrüchen und zur Cluster-Einteilung sind. [47], [52]–[54] Sabat *et al.* konnten so mithilfe des SNP-Abstands Ausbruchsstämme von Nicht-Ausbruchsstämmen abgrenzen, während es das Ziel von Stimson *et al.* war, verschiedene Stämme von *Mycobacterium tuberculosis* mithilfe eines SNP-Grenzwertes einem Transmissionscluster zuzuordnen und ein zeitlich nicht weit zurückliegendes Transmissionsereignis als wahrscheinlich einstufen zu können. [47], [52] Auch für weitere Bakterienspezies wurden Ausbruchsanalysen mithilfe von SNP-Grenzwerten durchgeführt. [55]

Dank den Fortschritten in den Sequenzierungstechniken und der Verfügbarkeit auch in kleineren Forschungs- und klinischen Laboren könnten die aus NGS

gewonnenen Daten die Basis für Routine-Diagnostiken und Infektionskontrollmaßnahmen bilden sowie der Ermittlung von Transmissionsrouten und -ursprüngen in Ausbruchsanalysen im klinischen Setting dienen. [37], [52], [53]

## **1.5 Zielsetzung und Hypothesen**

Ziel dieser Arbeit ist die Entwicklung eines Algorithmus zur retrospektiven Analyse der SNP-Abstände mit und ohne Berücksichtigung des Zeitfaktors – also des zeitlichen Abstands – von verschiedenen nosokomialen *P. aeruginosa*-Stämmen. Diese Datenanalyse soll im Folgenden der Entscheidung über den Nutzen eines sinnvollen einheitlichen SNP-Grenzwertes dienen, um zu differenzieren, ab welchem SNP-Abstand eine Transmission als wahrscheinlich einzustufen ist. Diese Unterscheidung ist Grundlage für die weitere Erforschung von Transmissionsfaktoren, die die Übertragung von *P. aeruginosa* innerhalb von Krankenhäusern begünstigen. Außerdem sind dadurch weitere Schlussfolgerungen bezüglich Transmissionsquellen und -routen für Ausbruchsanalysen möglich.

Für die Erforschung von Einflussfaktoren auf die Größe des SNP-Abstands wurden folgende Hypothesen aufgestellt und untersucht:

**Hypothese 1: Der SNP-Abstand zwischen Stämmen aus unterschiedlichen Probenmaterialien ist signifikant größer als jener zwischen Stämmen aus gleichen Probenmaterialien**

**a) SNP-Abstand ohne Berücksichtigung des Zeitfaktors**

**b) SNP-Abstand mit Berücksichtigung des Zeitfaktors**

An den verschiedenen Stellen des menschlichen Körpers, die von *P. aeruginosa* besiedelt werden können, könnten sich die Stämme isoliert entwickeln, sodass hypothetisch der SNP-Abstand zwischen Stämmen, die in unterschiedlichen Patientenmaterialien nachgewiesen wurden, größer wäre als jener zwischen Stämmen aus gleichen Patientenmaterialien. Da diese Stämme so bei zeitlich gleichem Abstand einen größeren SNP-Abstand aufweisen könnten, wäre hypothetisch auch der SNP-Abstand mit Berücksichtigung des

Zeitfaktors in Stämmen, die in unterschiedlichen Patientenmaterialien nachgewiesen wurden, größer als jener zwischen Stämmen aus gleichen Patientenmaterialien.

**Hypothese 2: Unter Berücksichtigung von nur Indexstamm-Verbindungen ist der SNP-Abstand zwischen Stämmen aus unterschiedlichen Probenmaterialien signifikant größer als jener zwischen Stämmen aus gleichen Probenmaterialien**

**a) SNP-Abstand ohne Berücksichtigung des Zeitfaktors**

**b) SNP-Abstand mit Berücksichtigung des Zeitfaktors**

Dieser Untersuchung liegt die Hypothese zugrunde, dass ein großer SNP-Abstand zweier Stämme, deren Nachweisdatum zeitlich eng beieinander liegt, dadurch zu erklären ist, dass sich die Stämme aus einem gemeinsamen Ursprungsstamm (Most Recent Common Ancestor) und nicht aus dem jeweiligen anderen Stamm entwickelt haben und sie somit getrennten Ästen im phylogenetischen Baum zuzuordnen sind. Aus diesem Grund wurde der im Cluster zeitlich erstnachgewiesene Stamm als dieser Ursprungsstamm angenommen und im Folgenden als Indexstamm bezeichnet. Die Analyse wurde mit und ohne Berücksichtigung des Zeitfaktors durchgeführt.

**Hypothese 3: Die Clustergröße korreliert mit dem SNP-Abstand**

Beinhalten die Cluster eine größere Anzahl an Stämmen, dann ist es wahrscheinlich, dass entweder mehr Stämme mit größeren zeitlichen Abständen oder mehr Stämme aus verschiedenen Probenmaterialien enthalten sind, sodass hypothetisch ein größerer SNP-Abstand bei zunehmender Clustergröße zu erwarten wäre, namentlich als Ausdruck einer durch die Messzahl bedingten Heterogenität. Darüber hinaus kann man bei einer größeren Anzahl an Stämmen im Cluster auch durch die Unterschiede im Resistenzmuster (und damit Unterschiede im Genom) von einem größeren medianen SNP-Abstand bei zunehmender Clustergröße ausgehen.

#### **Hypothese 4: Der maximale zeitliche Abstand korreliert mit dem SNP-Abstand**

Geht man von einer natürlichen Mutationsrate von *P. aeruginosa* aus, dann wäre ein SNP-Abstand zweier Isolate umso größer, je weiter diese zeitlich auseinander liegen, sodass hypothetisch eine positive Korrelation des medianen SNP-Abstands mit dem maximalen zeitlichen Abstand zu erwarten wäre.

#### **Hypothese 5: Die Anzahl der verschiedenen Probenmaterialien korreliert mit dem SNP-Abstand**

Geht man von einem größeren SNP-Abstand zwischen zwei Stämmen aus, die in unterschiedlichen Probenmaterialien nachgewiesen wurden, als jener zwischen Stämmen aus gleichen Probenmaterialien (vergleiche Hypothese 1), dann wäre hypothetisch auch ein größerer medianer SNP-Abstand in den Clustern, die Stämme aus einer größeren Anzahl an unterschiedlichen Probenmaterialien beinhalten, zu erwarten.

## 2 Material und Methoden

### 2.1 Studienaufbau und erhobene Daten

#### 2.1.1 Herkunft der Daten

Der vorliegenden Arbeit liegt eine prospektive Kohortenstudie zugrunde. Es wurden die Daten von 463 volljährigen Patientinnen und Patienten in anonymisierter Form verwendet.

Die Daten wurden im Zeitraum vom 01.07.2012 bis zum 01.01.2017 erhoben. Die eingeschlossenen Patientinnen und Patienten befanden sich auf den Stationen 65, 85 und 92 der Abteilung Innere Medizin II in der Medizinischen Klinik des Universitätsklinikums Tübingen. Diese Stationen umfassen die Fachgebiete Onkologie, Hämatologie, Immunologie, Rheumatologie und Pulmologie. Eine Hochrisiko-Konstellation gegenüber schweren Infektionen mit *P. aeruginosa* in diesem Patientenkollektiv ist anzunehmen. [11] Aus diesem Patientenkollektiv wurden variable Probenmaterialien mit positivem Nachweis auf *P. aeruginosa* aus Screening-Kulturen in die Studie eingeschlossen. Dieses Screening fand wie folgt statt: Routinemäßig wurden ein Rektal- und ein Rachenabstrich am ersten Tag des stationären Aufenthalts (Aufnahmetag) entnommen. Weitere Abnahmen desselben Materials fanden wöchentlich statt. Bei Verdacht auf Befall einer anderen Prädilektionsstelle mit *P. aeruginosa* sowie zur Verlaufskontrolle fanden Probenentnahmen auch in einem kürzeren Zeitintervall und an anderen Körperstellen statt.

Aus den oben erwähnten Patienten und Bakterienkulturen wurden zwei Listen (für den Zeitraum vom 01.07.2012 bis zum 01.02.2016 und vom 01.02.2016 bis zum 01.01.2017) mithilfe der Hybase Software (Tieto GmbH, Eschborn, Deutschland) und des Laboratoriumsinformationssystems SWISLAB (Swisslab GmbH, Berlin, Deutschland) erstellt. Diese Listen beinhalteten den Namen und das Geburtsdatum des Patienten, die Labornummer und das Material der entnommenen Proben, die Station auf der der Patient zum Abnahmezeitpunkt untergebracht war, sowie seine Liegedauer, den detektierten

Erreger mit Status und das Ergebnis der Antibiotikaresistenztestung. Für die Erstellung von weiteren Listen zur Gegenüberstellung der Daten wurde die Datenbank Microsoft Excel für Office 365 MSO Version 2002 (Microsoft Corporation, Redmond, USA) verwendet.

Die Studie wurde der Ethik-Kommission der Medizinischen Fakultät des UKT vorgelegt und wurde von dieser zugelassen (Projekt-Nummer: 016/2015B02).

Eine Übersicht der Reihenfolge des Arbeitsablaufs zum finalen Einschluss in die Studie wird in den folgenden Kapiteln (Kap.) genauer erläutert. Kurz zusammengefasst: Nachdem ein *P. aeruginosa*-Stamm in Screening-Abstrichen nachgewiesen wurde, wurde dieser auf Antibiotikaempfindlichkeit untersucht. Wenn ein *P. aeruginosa*-Stamm die Einschlusskriterien für die Studie erfüllte, wurden Identifikationsnummern sowohl für den Stamm als auch dem jeweiligen Patienten vergeben. Das Genom der Stämme wurde mittels NGS sequenziert und bioinformatisch analysiert, bis hin zu einer phylogenetischen Analyse der Stämme.

### **2.1.2 Einschlusskriterien für Patienten und Stämme (ID-Vergabe)**

Das zeitlich erste *P. aeruginosa*-Isolat eines stationären Patienten wurde in die Studie eingeschlossen, unabhängig davon, in welchem Probenmaterial das Isolat detektiert wurde und welches Antibiogramm vorlag (siehe Abb. 2). Weitere Isolate desselben Patienten wurden dann eingeschlossen, wenn (i) sie in einem weiteren Material nachgewiesen wurden (Ausnahme: Rektalabstriche und Stuhl wurden als gleiche Materialien betrachtet), oder (ii) wenn zwar schon ein Stamm im gleichen Material nachgewiesen und eingeschlossen wurde, aber das Antibiogramm von dem bereits eingeschlossenen Isolat abwich (siehe Abb. 2). Ein Antibiogramm war definitionsgemäß abweichend, wenn sich der Stamm in mindestens einer kompletten Titerstufe (von sensibel (S) auf resistent (R) oder umgekehrt) in mindestens einem Antibiotikum von den bereits eingeschlossenen Isolaten unterschied. Die Reihenfolge der Isolat-ID-Vergabe erfolgte innerhalb eines Patienten zeitlich nach Datum der Probenabnahme.



Nachweis verdächtigen Wachstums (grünliche Kolonien) erfolgte eine Speziesidentifizierung durch ein lineares Matrix-assisted Laser Desorption Ionization Time-of-flight (MALDI-TOF)-Massenspektrometer (bioMérieux, Marcy l'Etoile, Frankreich), welche durch ein weiteres MALDI-TOF Massenspektrometer (Microflex LT, Bruker Daltonics, Deutschland) bestätigt wurde. In die Studie eingeschlossen wurden alle Proben mit im MALDI-TOF bestätigtem Nachweis von *P. aeruginosa*.

Die in vitro-Antibiotikaresistenztestung fand automatisiert auf dem VITEK (bioMérieux, Marcy l'Etoile, Frankreich) statt. Bei unklaren Befunden wurde die minimale Hemmkonzentration mithilfe von antibiotischen Etest-Streifen (M.I.C Evaluator-Teststreifen, Liofilchem, Italien) bestimmt. Ergebnisse der Testungen wurden nach den Leitlinien des European Committee on Antimicrobial Susceptibility Testing (EUCAST Version 8.0, 2018) mittels eines Plattendiffusionstests ausgewertet.

Für das weitere Procedere wurden die Stämme bei -80 °C gelagert.

### **2.2.2 Next-Generation Sequencing (NGS)**

Mithilfe des Focused Ultrasonicator (Covaris, Woburn, USA) wurde die genomische DNA der *P. aeruginosa*-Isolate in DNA-Fragmente mit einer Größe von 450 - 550 bp geschnitten. DNA-Bibliotheken wurden mit dem NEXTflex™ DNA Sequencing Kit (Bioo Scientific, Austin, USA) erstellt. Diese wurden auf einer HiSeq2500 Plattform (Illumina, San Diego, USA) mit einer Leseweite von 2 x 125 bp im Rahmen einer externen Auftragsvergabe sequenziert (CeMeT, Tübingen, Germany).

Eine Matrix der vorhandenen und fehlenden Gene aller *P. aeruginosa*-Isolate (Presence-Absence-Matrix) wurde mithilfe von Roary Version 1.006924 [56] erzeugt und eine minimale relative Übereinstimmung der Identität von 95 % für blastp vorausgesetzt.

Mithilfe der A5 pipeline Version 20,140,604 [57] und SPAdes Version 3.7.0 [58] fand ein Zusammenbau (Assembly) der Sequenzstücke zu fast vollständigen



bakteriellen Genomen statt. Die zusammengesetzten Gerüste (Assembled Scaffolds) wurden mithilfe von Prokka Version 1.11 [59] annotiert. Um ein Core-Genom (Erklärung siehe Kap. 1.3.1) für *P. aeruginosa* zu generieren, wurde die Software Spine Version 0.1.2 [41] verwendet. Das entstandene Core-Genom der hier verwendeten *P. aeruginosa*-Stämme bestand aus einer Länge von 5.259.603 bp.

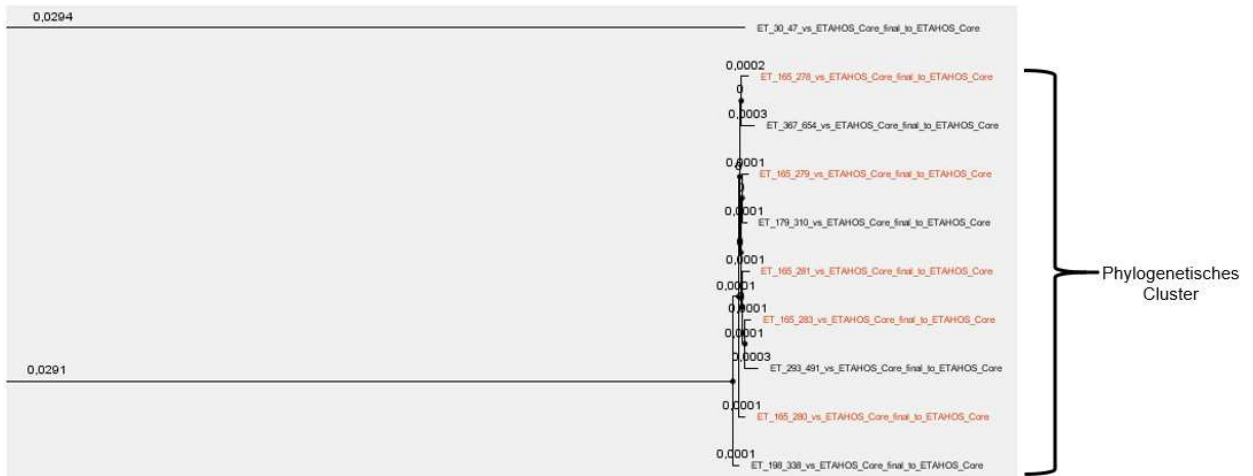
Um zu überprüfen, ob es sich bei den analysierten Stamm-DNAs um ein reines *P. aeruginosa*-Genom ohne Kontamination handelte, wurde der GC-Gehalt bestimmt. Stover *et al.*, 2000 geben einen GC-Gehalt des Genoms von *P. aeruginosa* von durchschnittlich 66,6 % an. [14] Aus diesem Grund wurden alle Sequenzen, deren GC-Gehalt  $< 65\%$  oder  $\geq 67\%$  betrug, von dieser Studie ausgeschlossen. Des Weiteren wurden Stamm-DNAs ausgeschlossen, die aus  $\geq 400$  Scaffolds bestanden, da hier das Assemblieren nicht zu Genüge funktionierte.

Mittels SPANDx Version 3.1 [61] wurden Core-Genom SNPs festgestellt. Mithilfe dieser genetischen Variationen wurden die *P. aeruginosa*-Stämme auf ihre phylogenetische Verwandtschaft hin mit RAxML Version 8.2.6 [62] untersucht. Für diesen Vorgang wurde ein GTR-Substitutionsmodell sowie eine Gamma-Distribution mit 1000 Bootstraps verwendet. Der entstehende Maximum-Likelihood-Baum zeigt genetische Verwandtschaft an.

## **2.3 Phylogenetische Analyse**

### **2.3.1 Einteilung in Patientencluster**

Zur Übersicht über den genetischen Unterschied der eingeschlossenen Stämme untereinander wurde mithilfe des Programms FigTree v1.4.3 [63] aus dem SNP-Alignment ein phylogenetischer Baum erstellt, durch welchen ersichtlich war, in welche phylogenetischen Cluster die Stämme eingeordnet werden konnten (siehe Abb. 2).



**Abb. 2:** Phylogenetisches Cluster in FigTree

Als phylogenetisches Cluster wurden die Stämme definiert, die einen gemeinsamen Ursprungstamm (Most Recent Common Ancestor) in unmittelbarer Nähe hatten (siehe Abb. 2). Für dieses Entscheidung, ob es sich um ein phylogenetisches Cluster handelte, gab es keine absoluten Kriterien. Es handelte sich um eine Einteilung anhand der optischen Darstellung im phylogenetischen Baum. Diese Einteilung diente im weiteren Verlauf der Untersuchung, ab welchem SNP-Abstand zwischen zwei Stämmen man von zwei phylogenetisch verschiedenen Stämmen ausgehen konnte und beispielsweise eine Transmission von Stämmen zwischen Patienten wahrscheinlich und nicht auf Mutationen von bakterieller DNA während einer Kolonisation bzw. Infektion eines Patienten zurückzuführen war.

Um dieser Frage zur natürlichen Mutationsrate eines *P. aeruginosa*-Stammes in einem Patienten im Detail nachzugehen, wurden *P. aeruginosa*-Stämme, die aus dem gleichen Patienten stammten, zu einem sogenannten Patientencluster zusammengefügt. Für die Bestimmung der Mutationsrate fungierten jene Patientencluster als Grundlage, die die folgenden vier Bedingungen erfüllten:

- 1) Die Stämme befanden sich in einem phylogenetischen Cluster.
- 2) Die Stämme wurden beim gleichen Patienten detektiert (in Abb. 3 farblich markierte Stämme).

- 3) Das Patientencluster bestand aus mindestens vier Stämmen.
- 4) Der zeitliche Abstand zwischen zuerst und zuletzt detektiertem Stamm betrug mindestens 20 Tage.

## 2.4 Berücksichtigung des Zeitfaktors

Als neuer Ansatz für die statistischen Untersuchungen wurde neben dem absoluten SNP-Abstand (Anzahl an SNPs) die Cluster-Mutationsrate (in SNPs/Jahr) für jede Verbindung zwischen zwei *P. aeruginosa*-Stämmen gebildet, um auch den zeitlichen Faktor zu berücksichtigen. Dies ist relevant, da der zeitliche Abstand zwischen den *P. aeruginosa*-Stämmen nicht standardisiert war und somit auch die SNP-Abstände deutlich variierten, da ein umso größerer SNP-Abstand bei zunehmendem zeitlichem Abstand zu erwarten war. Hypothetisch wäre diese Cluster-Mutationsrate dann zwischen allen Stämmen ähnlich groß und somit besser vergleichbar, unabhängig davon, welchen zeitlichen Abstand die zwei Isolate aufwiesen. Für diese Cluster-Mutationsrate wird im Folgenden auch synonym der Begriff SNP-Abstand/Zeit verwendet. Hierbei wurde der Quotient aus dem SNP-Abstand zwischen zwei Stämmen und dem jeweils dazu gehörigen zeitlichen Abstand der Isolation in Tagen gebildet und auf ein Jahr (365 Tage) hochgerechnet. Dies ist in folgender Formel nochmals zusammengefasst:

$$SNP - Abstand/Zeit = \frac{SNP - Abstand\ zwischen\ Stamm\ X\ und\ Y}{Zeitlicher\ Abstand\ zwischen\ Stamm\ X\ und\ Y} \times 365\ (Tage)$$

Quotienten zwischen Stämmen, die am selben Tag detektiert wurden, konnten nicht berücksichtigt werden, da eine Division durch 0 nicht zulässig war.

## 2.5 Statistische Analyse

Für die statistischen Analysen wurde das Programm IBM SPSS Statistics 25 Version 1.1 für Microsoft Windows (SPSS GmbH, München, Deutschland) verwendet. Für die deskriptive Beschreibung der Daten wurden absolute und relative Häufigkeit, Mittelwert, Median und Standardabweichung bestimmt.

Zur Untersuchung von Zusammenhängen wurde eine nicht-parametrische Rangkorrelationsanalyse (Spearman-Rho) angewendet. Ein p-Wert  $\leq 0.05$

wurde dabei für den Spearman-Korrelationskoeffizienten  $\rho$  als statistisch signifikant festgelegt. Zusätzlich wurden die Korrelationsanalysen in Streudiagrammen dargestellt.

Für den Vergleich der SNP-Abstände und SNP-Abstände/Zeit zwischen gleichem und ungleichem Probenmaterial wurde der nicht-parametrische Mann-Whitney-Test angewendet. Eine zweiseitige Irrtumswahrscheinlichkeit von 5 % wurde auch hier als Signifikanzniveau angenommen.

## **3 Ergebnisse**

### **3.1 Eingeschlossene Patienten und Stämme**

#### **3.1.1 Ausgangslage und Ausschlussverfahren**

In den Zeiträumen vom 01.07.2012 bis zum 01.02.2016 und vom 01.02.2016 bis zum 01.01.2017 (54 Monate) wurden insgesamt 1740 Stämme der Spezies *P. aeruginosa* in Patientenmaterialien nachgewiesen. (1360 im ersten Zeitraum und 380 im zweiten).

Es fand dabei eine Detektion in insgesamt 32 verschiedenen Probenmaterialien statt. Mit einer Anzahl von 1053 (60,52 %) wurde *P. aeruginosa* am häufigsten in Rektalabstrichen detektiert. Die zweithäufigste Detektion fand in Rachenabstrichen mit einer Anzahl von 315 (18,10 %) statt. Eine detaillierte Zusammensetzung der verschiedenen Probenmaterialien findet sich in Tab. 1. Unter dem Begriff ‚Respiratorisches Material‘ wurden die Probenmaterialien Sputum, Bronchialalveoläre Lavage, Tracheal- und Bronchialsekret zusammengefasst.

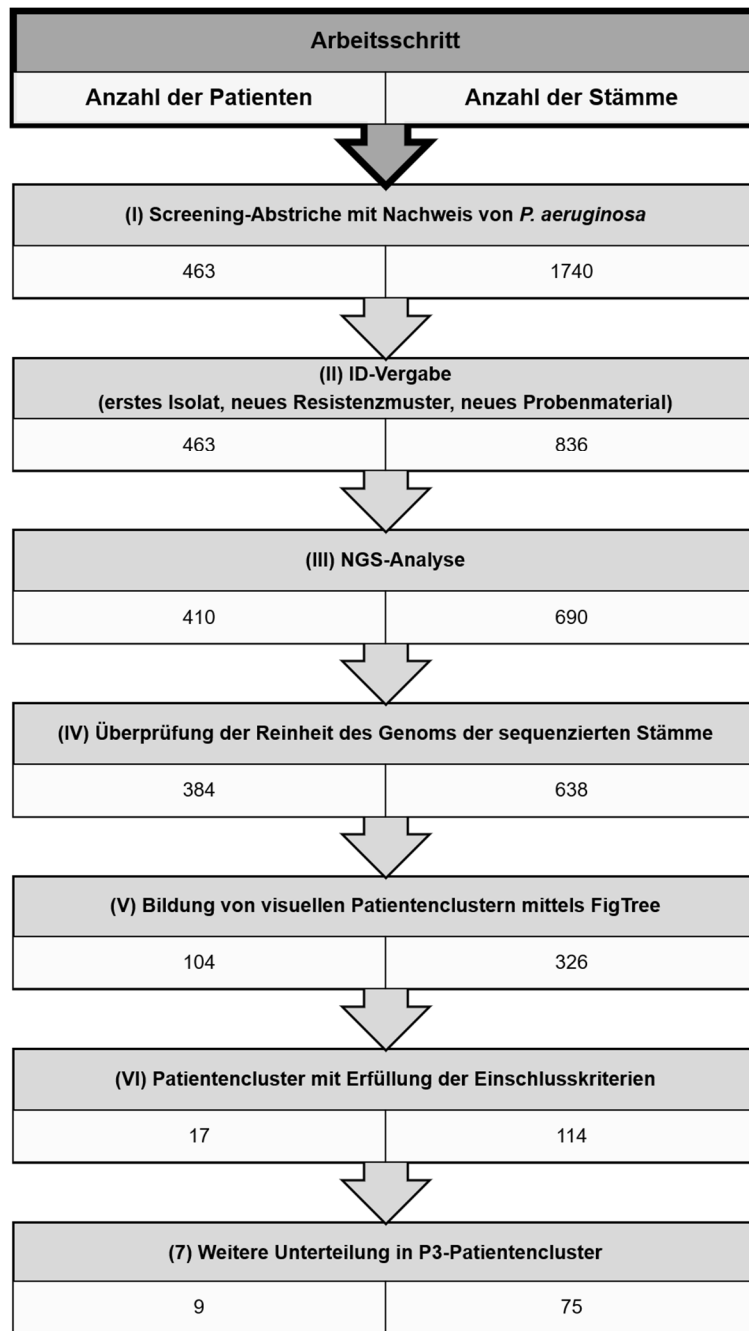
**Tab. 1:** Anzahl der Probenmaterialien mit Nachweis von *P. aeruginosa* mit absoluter und relativer Häufigkeit, aufgeschlüsselt nach Zeitraum.

Probenmaterial	2012 - 2016 N = 1360		2016 - 2017 N = 380		Gesamt N = 1740	
	Absolut	Relativ [%]	Absolut	Relativ [%]	Absolut	Relativ [%]
Rektalabstrich	781	57,43	272	71,58	1053	60,52
Rachenabstrich	264	19,41	51	13,42	315	18,10
Stuhl	183	13,46	26	6,84	209	12,01
Mittelstrahlurin	36	2,65	6	1,58	42	2,41
Respiratorisches Material	33	2,42	8	2,11	41	2,36
Blutkultur	9	0,66	7	1,84	16	0,92
Wundabstrich	11	0,81	4	1,05	15	0,86
Abstrich	8	0,59	-	-	8	0,46
Katheterurin	4	0,29	2	0,53	6	0,34
Leistenabstrich	4	0,29	2	0,53	6	0,34
Ulcusabstrich	4	0,29	-	-	4	0,23
Analabstrich	3	0,22	-	-	3	0,17
Zungenabstrich	3	0,22	-	-	3	0,17
Abstrich Katheter- Austritt	2	0,15	-	-	2	0,11
Hautabstrich	2	0,15	-	-	2	0,11
Zentraler Venenkatheter	-	-	2	0,53	2	0,11
Abstrich Hoden	1	0,07	-	-	1	0,06
Abstrich Mundschleimhaut	1	0,07	-	-	1	0,06
Abstrich Ulcus cruris	1	0,07	-	-	1	0,06
Ascitespunktat	1	0,07	-	-	1	0,06
Blasengalle	1	0,07	-	-	1	0,06
Drainage	1	0,07	-	-	1	0,06
Katheter	1	0,07	-	-	1	0,06
Lungengewebe	1	0,07	-	-	1	0,06
PEG-Eintrittsstelle	1	0,07	-	-	1	0,06
Punktat	1	0,07	-	-	1	0,06
Redonkatheter	1	0,07	-	-	1	0,06
Urethralabstrich	1	0,07	-	-	1	0,06
Vaginalabstrich	1	0,07	-	-	1	0,06

In Abb. 3 wird der Arbeitsablauf und der Ausschluss der *P. aeruginosa*-Stämme etappenweise in den Arbeitsschritten (I) – (VI) dargestellt:

- (I) Im Rahmen des Screenings (siehe Kap. 2.1.1) wurden 1740 *P. aeruginosa*-Stämme bei 463 Patienten nachgewiesen.
- (II) Mithilfe der in Kap. 2.1.2 beschriebenen Kriterien wurden zunächst 836 Isolate von 463 Patienten eingeschlossen.
- (III) Die NGS-Analyse wurde an 690 *P. aeruginosa*-Stämmen von 410 Patienten durchgeführt. Zuvor wurden 146 *P. aeruginosa*-Stämme von 53 Patienten ausgeschlossen, da sie nicht korrekt eingefroren/aufbereitet, ambulant statt stationär detektiert oder fälschlicherweise als *P. aeruginosa*-Stamm deklariert worden waren.
- (IV) Nach Überprüfung der Reinheit des Genoms der sequenzierten Stämme mittels der in Kap. 2.2.2 erläuterten Kriterien wurden weitere 52 Stämme von 26 Patienten ausgeschlossen.
- (V) Im folgenden Schritt ergab sich mittels FigTree v1.4.3 [63] eine Einteilung von 326 Stämmen von 104 Patienten in 104 visuelle Patientencluster.
- (VI) Weitere 212 Stämme von 87 Patienten wurden ausgeschlossen, da sie die Einschlusskriterien aus Kap. 2.3.1 nicht erfüllten und weniger als vier Stämme pro Cluster oder einen maximalen zeitlichen Abstand < 20 Tage aufwiesen.

Für die finale Auswertung im Rahmen von Patientenclustern standen somit noch 114 (6,6 %) der ursprünglichen *P. aeruginosa*-Stämme von 17 Patienten (3,7 %) zur Verfügung. Demnach konnten 17 Patientencluster für die meisten Analysen verwendet werden.



**Abb. 3:** Arbeitsablauf mit Anzahl der eingeschlossenen Patienten und *P. aeruginosa*-Stämme.

Für die Feststellung von Unterschieden zwischen gleichem und ungleichem Probenmaterial wurden nur Cluster berücksichtigt, die mindestens drei Paar-Vergleiche in gleichen und mindestens drei Paar-Vergleiche in ungleichen Probenmaterialien ermöglichten, da bei einer statistischen Analyse von nur zwei Paar-Vergleichen keine aussagekräftigen Ergebnisse zu erwarten wären (z.B. Median). (7) Dies führte zum Ausschluss von acht weiteren Patientenclustern



und somit zu einer Anzahl von neun P3-Patientenclustern. Diese umfassten 75 Stämme.

### **3.1.2 Demographische Daten der zugrunde liegenden Studienpopulation der Patientencluster**

Zur besseren Übersicht wurden im Folgenden die Daten der 17 Patientencluster und der neun P3-Patientencluster in jedem Unterkapitel gemeinsam nacheinander dargestellt, auch wenn für die späteren statistischen Analysen nur jeweils eine der beiden Clustergruppierungen (Patientencluster oder P3-Patientencluster) herangezogen wurde.

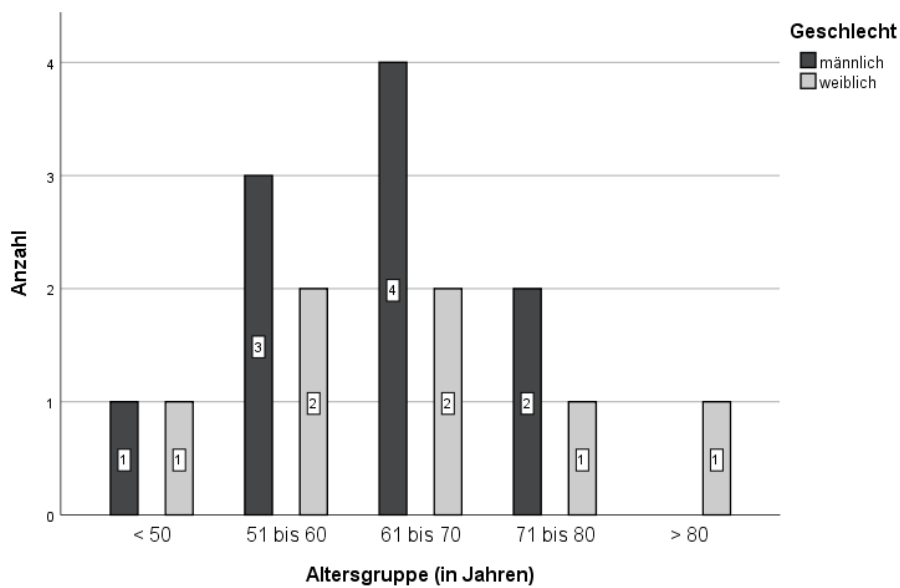
Die Bezeichnungen der 17 Patientencluster (Cluster 1 bis 17) entsprechen nicht den Bezeichnungen des ursprünglichen Datensatzes (siehe Anhang). In Tab. 2 kann entnommen werden, welches Cluster welcher Patienten-ID entsprach.

**Tab. 2:** Übersicht über die den Clustern entsprechenden Patienten-IDs

<b>Cluster</b>	<b>Patienten-ID</b>
<b>1*</b>	<b>244</b>
<b>2*</b>	<b>365</b>
3	190
<b>4*</b>	<b>15</b>
5	142
<b>6*</b>	<b>413</b>
<b>7*</b>	<b>358</b>
<b>8*</b>	<b>295</b>
9	301
<b>10*</b>	<b>54</b>
<b>11*</b>	<b>127</b>
12	247
13	206
<b>14*</b>	<b>258</b>
15	166
16	131
17	149

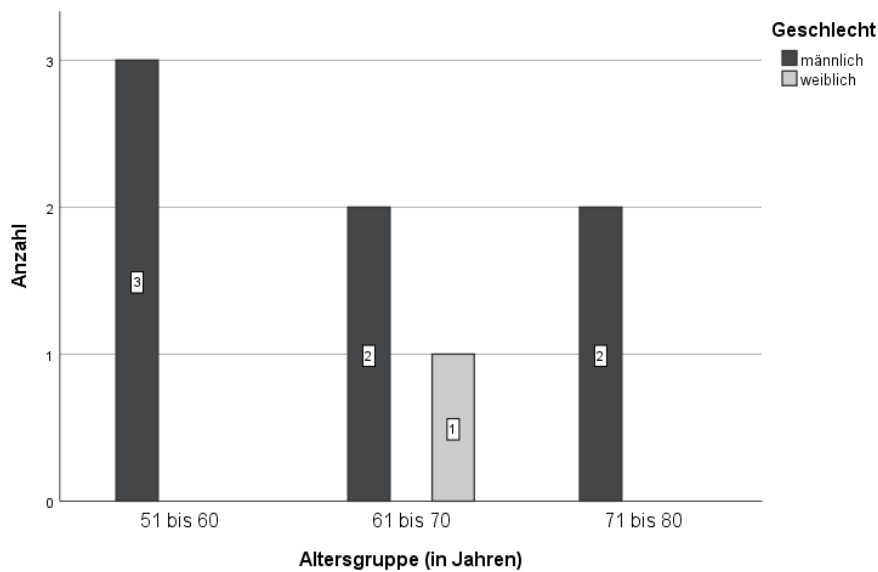
\* Es handelt sich um ein P3-Patientencluster

Die Altersspanne der eingeschlossenen Patienten der 17 Patientencluster reichte von 47 bis 86 Jahren und das mediane Alter betrug 64 (IQR 54-70) Jahre. Von den 17 Patienten war über die Hälfte männlich (58,82 %). Eine genaue Altersverteilung findet sich in Abb. 4. Es handelte sich jeweils um das Alter bei Erstdetektion von *P. aeruginosa*.



**Abb. 4:** Altersverteilung der zugrunde liegenden Studienpopulation der 17 Patientencluster, aufgeschlüsselt nach Altersgruppe.

Die Altersspanne der eingeschlossenen Patienten der neun P3-Patientencluster reichte von 51 bis 74 Jahren und das mediane Alter betrug 66 (IQR 56-70) Jahre. Von den neun Patienten waren acht (88,89 %) männlich und nur ein Patient (11,11 %) weiblich. Eine genaue Altersverteilung findet sich in Abb. 5. Wieder handelte es sich jeweils um das Alter bei Erstdetektion von *P. aeruginosa*.

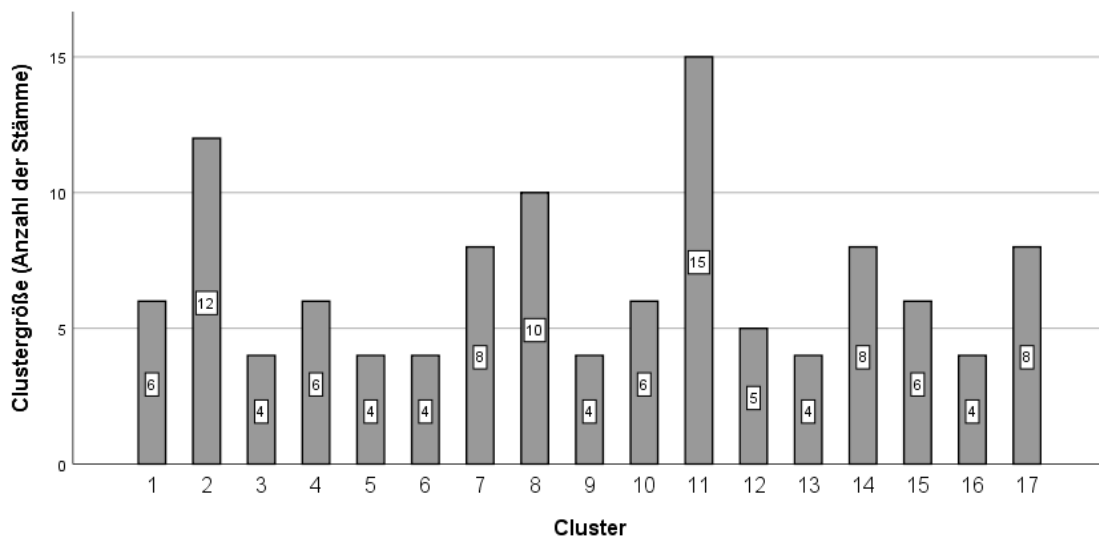


**Abb. 5:** Altersverteilung der zugrunde liegenden Studienpopulation der neun P3-Patientencluster, aufgeschlüsselt nach Altersgruppe

### 3.1.3 Clustergröße der Patientencluster

In den 17 Patientenclustern betrug die Clustergröße im Median 6 (IQR 4-8) Stämme pro Cluster. Das größte Patientencluster beinhaltete 15 Stämme (Cluster 11), gefolgt von Cluster 2 mit zwölf und Cluster 8 mit zehn Stämmen, während die kleinsten Patientencluster eine Anzahl von vier Stämmen besaßen (Cluster 3, 5, 6, 9, 13 und 16).

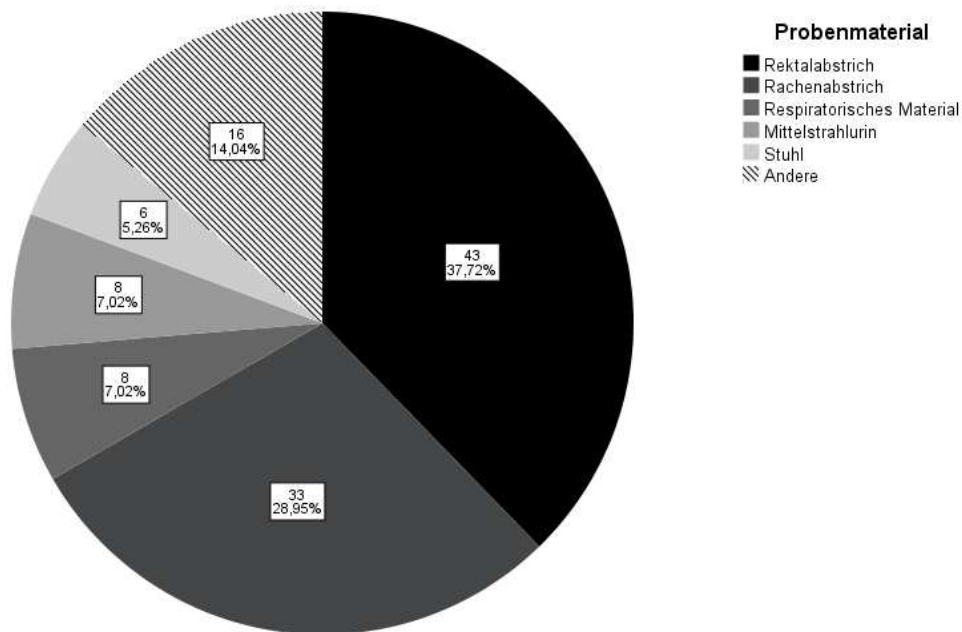
In den neun P3-Patientenclustern betrug die Clustergröße im Median 8 (IQR 6-11) Stämme pro Cluster. Das größte Patientencluster beinhaltete 15 Stämme (Cluster 11), gefolgt von Cluster 2 mit zwölf und Cluster 8 mit zehn Stämmen, während das kleinste Patientencluster wieder eine Anzahl von vier Stämmen besaß (Cluster 6). Graphisch dargestellt ist die Clustergröße in Abb. 6.



**Abb. 6:** Clustergröße (Anzahl der Stämme pro Cluster) der 17 Patientencluster, aufgeschlüsselt nach Cluster.

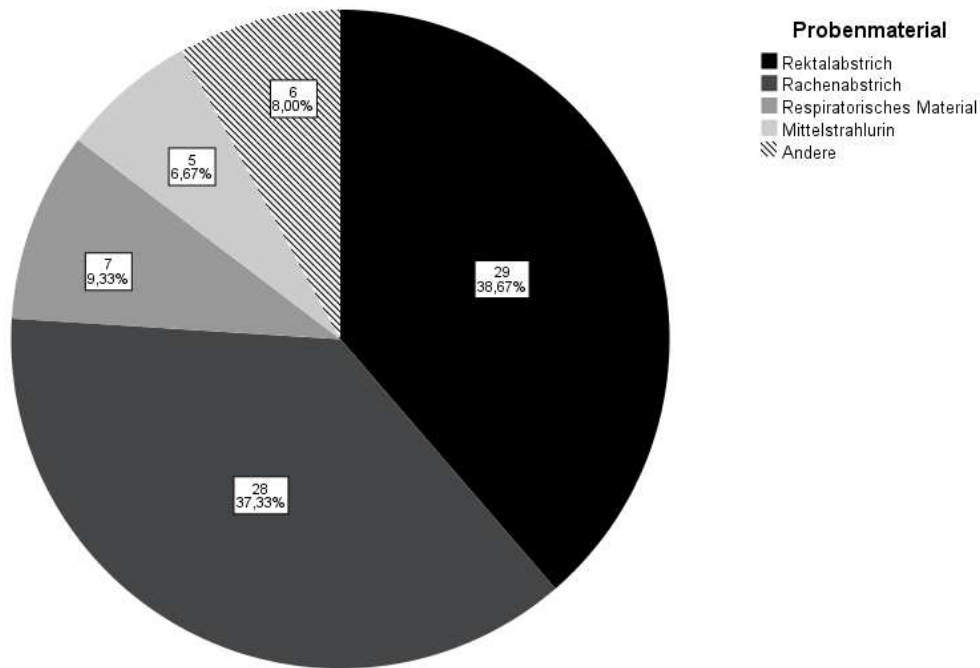
### 3.1.4 Probenmaterialien der Patientencluster

Die 114 Stämme der 17 Patientencluster wurden in 16 verschiedenen Probenmaterialien detektiert. Das Probenmaterial Rektalabstrich war am häufigsten vertreten (37,72 %), gefolgt von den Probenmaterialien Rachenabstrich (28,95 %), Respiratorisches Material (7,02 %) und Mittelstrahlurin (7,02 %). Eine genaue Zusammensetzung der verwendeten Probenmaterialien der 17 Patientencluster findet sich in Abb. 7.



**Abb. 7:** Probenmaterialien der 17 Patientencluster mit absoluter und relativer Häufigkeit. Probenmaterialien, die anteilmäßig < 5 % vertreten sind, sind unter „Andere“ zusammengefasst (Ulcusabstrich, Blutkultur, Abstrich, Leistenabstrich, Wundabstrich, Abstrich Hoden, Abstrich Mundschleimhaut, Zentraler Venenkatheter, Katheterurin, Abstrich Ulcus cruris).

Die 75 Stämme der neun P3-Patientencluster wurden in acht verschiedenen Probenmaterialien detektiert. Das Probenmaterial Rektalabstrich war auch hier am häufigsten vertreten (38,67 %), gefolgt von den Probenmaterialien Rachenabstrich (37,33 %), Respiratorisches Material (9,33 %) und Mittelstrahlurin (6,67 %). Eine genaue Zusammensetzung der verwendeten Probenmaterialien der vorläufigen Patientencluster findet sich in Abb. 8.



**Abb. 8:** Probenmaterialien der neun P3-Patientencluster mit absoluter und relativer Häufigkeit. Probenmaterialien, die anteilmäßig < 5 % vertreten sind, sind unter „Andere“ zusammengefasst (Stuhl, Blutkultur, Zentraler Venenkatheter, Katheterurin).

### 3.1.5 Tabellarische Übersicht über die demographischen Daten, die Clustergröße und die Probenmaterialien der Patientencluster

Zur abschließenden Übersicht über die demographischen Daten, die Clustergröße und die Probenmaterialien der 17 Patientencluster sowie der neun P3-Patientencluster werden die Daten in Tab. 3 und Tab. 4 veranschaulicht.

**Tab. 3:** Demographische Daten und Clustergröße mit Median (IQR) der Patientencluster

	Patientencluster (N=17)	P3-Patientencluster (N=9)
Anzahl der Patienten (Anzahl der weiblichen Patienten)	17 (7)	9 (1)
Alter [Jahre]	64 (54-70)	66 (56-70)
Anzahl der Stämme Clustergröße [Anzahl der Stämme pro Cluster]	114 6 (4-8)	75 8 (6-11)

**Tab. 4:** Absolute und relative Häufigkeit der verschiedenen Probenmaterialien der Patientencluster

Probenmaterial	Stämme (N=114) der Patientencluster		Stämme (N=75) der P3-Patientencluster	
	Absolut	Relativ [%]	Absolut	Relativ [%]
Rektalabstrich	43	37,72	29	38,67
Rachenabstrich	33	28,95	28	37,33
Mittelstrahlurin	8	7,02	5	6,67
Respiratorisches Material	8	7,02	7	9,33
Stuhl	6	5,26	3	4,00
Ulcusabstrich	3	2,63	-	-
Blutkultur	2	1,75	1	1,33
Abstrich	2	1,75	-	-
Leistenabstrich	2	1,75	-	-
Wundabstrich	2	1,75	-	-
Abstrich Hoden	1	0,88	-	-
Abstrich Mundschleimhaut	1	0,88	-	-
Zentraler Venenkatheter	1	0,88	1	1,33
Katheterurin	1	0,88	1	1,33
Abstrich Ulcus cruris	1	0,88	-	-

## 3.2 Darstellung des SNP-Abstands in den Patientenclustern

### 3.2.1 SNP-Abstand ohne Berücksichtigung des Zeitfaktors

In diesen Ergebnissen wurden die zusammenfassenden Statistiken nach Patientencluster aufgeschlüsselt miteinander verglichen. Der Zeitfaktor, also der zum SNP-Abstand zweier Stämme zugehörige zeitliche Abstand ihrer Isolation in Tagen, wurde dabei nicht berücksichtigt. Dieser wird in Kap. 3.2.2 mit einbezogen.

Die 17 Patientencluster wurden nach aufsteigendem Mittelwert des SNP-Abstands (Anzahl der SNPs) sortiert. Cluster 1 war somit das Cluster mit dem



kleinsten mittleren SNP-Abstand mit 40,40 (95%KI 33,26-47,54) SNPs, gefolgt von Cluster 2 mit 41,85 (95%KI 33,1-50,59) SNPs und Cluster 3 mit 59,50 (95%KI 49,12-69,88) SNPs. Die größten Mittelwerte des SNP-Abstands fanden sich in Cluster 15 mit 560,07 (95%KI 388,71-731,43) SNPs, Cluster 16 mit 566,67 (95%KI 488,55-644,78) SNPs und Cluster 17 mit 583,14 (95%KI 411,24-755,04) SNPs. Betrachtet man den Median, ließ sich der niedrigste Wert in Cluster 2 mit 28 (IQR 23-33,25) SNPs erkennen, während der höchste in Cluster 15 mit 765 (IQR 197-811) SNPs zu finden war.

Betrachtet man den minimalen und maximalen SNP-Abstand, so ließ sich der kleinste SNP-Abstand in Cluster 2 mit einem SNP-Abstand von 15 SNPs beobachten, während der größte Wert mit 1423 SNPs zu Cluster 17 gehörte. Die *P. aeruginosa*-Stämme von Cluster 13 wurden alle im gleichen Probenmaterial (Rektalabstrich) nachgewiesen. Cluster 2 enthielt mit sieben die höchste Anzahl an verschiedenen Probenmaterialien (Blutkultur, Katheterurin, Mittelstrahlurin, Rachenabstrich, Rektalabstrich, Stuhl, zentraler Venenkatheter). Der zeitlich größte Abstand der Isolation von Stämmen innerhalb eines Clusters betrug 417 Tage und wurde in Cluster 15 festgestellt. Der zeitlich geringste Abstand war mit 25 Tagen in Cluster 16 enthalten.

Eine übersichtliche Darstellung aller Werte zeigt Tab. 5, graphisch dargestellt werden die Werte in Abb. 9.

**Tab. 5:** Übersicht über SNP-Abstand, Clustergröße, Anzahl der verschiedenen Probenmaterialien und über den maximalen zeitlichen Abstand der 17 Patientencluster. Die neun P3-Patientencluster sind fett gedruckt, grau hinterlegt und mit einem ‚\*‘ gekennzeichnet.

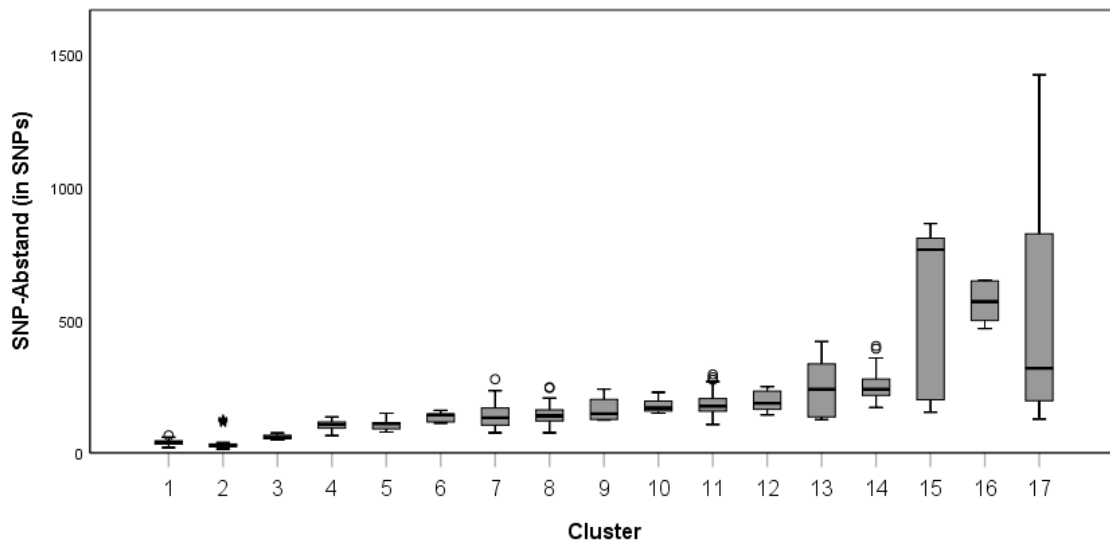
Cluster	SNP-Abstand <sup>a</sup>				Clustergröße <sup>b</sup>	Anzahl verschiedener Probenmaterialien	Maximaler zeitlicher Abstand im Cluster <sup>c</sup>
	M (95%KI)	Me (IQR)	Min	Max			
1*	40,40 (33,26-47,54)	39,00 (30-45)	20	66	6	2	105
2*	41,85 (33,1-50,59)	28,00 (23-33,25)	15	128	12	7	100
3	59,50 (49,12-69,88)	57,50 (51,75-68,25)	48	75	4	2	28
4*	104,13 (93,66-114,6)	107,00 (92-116)	65	135	6	3	343
5	108,50 (82,94-134,06)	110,00 (87-122,75)	78	149	4	1	119
6*	136,67 (116,95-156,93)	143,00 (115-149,5)	112	160	4	2	44
7*	142,07 (123,13-161,02)	132,00 (102-171,75)	76	277	8	4	372
8*	142,80 (132,15-153,45)	139,00 (119,5-163)	76	247	10	3	403
9	164,17 (114,42-213,91)	147,50 (125,25-210,75)	123	240	4	2	49
10*	177,93 (163,08-192,78)	169,00 (157-207)	149	228	6	2	82
11*	183,18 (175,96-190,4)	176,00 (156,5-204,5)	107	295	15	3	97
12	192,60 (165,27-219,93)	186,50 (159,25-232,25)	143	249	5	4	170
13	248,67 (115,34-382)	239,00 (132,5-356)	125	419	4	1	168
14*	252,07 (228,45-275,69)	239,00 (216,25-279,75)	171	402	8	3	42
15	560,07 (388,71-731,43)	765,00 (197-811)	153	863	6	4	417
16	566,67 (488,55-644,78)	569,00 (490,5-647,5)	468	649	4	2	25
17	583,14 (411,24-755,04)	318,00 (194,5-839,25)	127	1423	8	5	79

\* Es handelt sich um ein P3-Patientencluster

<sup>a</sup> Mittelwert (M), Median (Me), Minimal- (Min) und Maximalwert (Max) der SNP-Abstände (in SNPs) eines Clusters

<sup>b</sup> Entspricht der Anzahl an Stamm-IDs eines Clusters

<sup>c</sup> Entspricht dem zeitlichen Abstand zwischen zeitlich erster und letzter Stamm-ID eines Clusters in Tagen



**Abb. 9:** Verteilung des mittleren SNP-Abstands (in SNPs) innerhalb der 17 Patientencluster, aufgeschlüsselt nach Cluster. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

### 3.2.2 SNP-Abstand mit Berücksichtigung des Zeitfaktors

In diesen Ergebnissen wurden die zusammenfassenden Statistiken nach Patientencluster aufgeschlüsselt miteinander verglichen. Der Zeitfaktor, also der zeitliche Abstand zwischen der Isolation von zwei Isolaten, wurde hierbei berücksichtigt.

Es wurde die Nummerierung der Cluster von Kap. 3.2.1 übernommen.

Cluster 3 war das Cluster mit dem kleinsten mittleren SNP-Abstand/Zeit mit 762,59 (95%KI 635,28-889,89) SNPs/Jahr, gefolgt von Cluster 1 mit 791,50 (95%KI 170,02-1412,98) SNPs/Jahr und Cluster 2 mit 969,69 (95%KI 493,35-1446,03) SNPs/Jahr. Die größten Mittelwerte des SNP-Abstands/Zeit fanden sich in Cluster 14 mit 13542,53 (95%KI 4192,52-22892,54) SNPs/Jahr, Cluster 17 mit 15491,37 (95%KI 3892,08-27090,67) SNPs/Jahr und Cluster 16 mit 22613,59 (95%KI -16533,06-61760,23) SNPs/Jahr. Betrachtet man den Median, ließ sich der niedrigste Wert in Cluster 8 mit 232 (IQR 153,47-1467,24) SNPs/Jahr, gefolgt von Cluster 1 mit 272,53 (IQR 149,1-915,76) SNPs/Jahr,

erkennen, während der höchste in Cluster 16 mit 8336,6 (IQR 7517,67-44847,99) SNPs/Jahr zu finden ist.

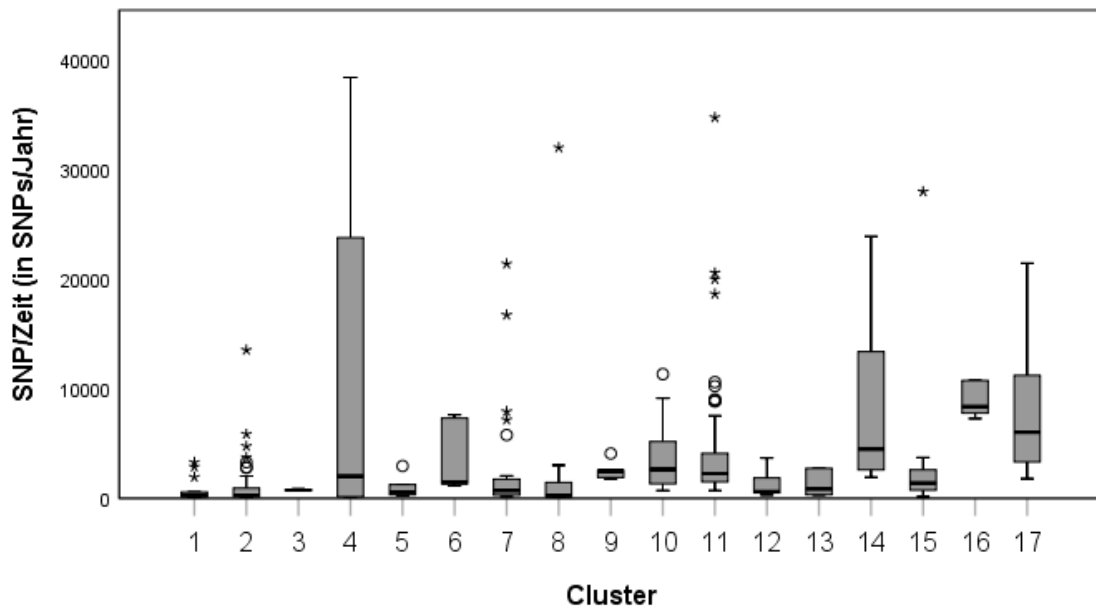
Eine übersichtliche Darstellung aller Werte zeigt Tab. 6, graphisch dargestellt werden die Werte in Abb. 10.

**Tab. 6:** Übersicht über den SNP-Abstand unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) der 17 Patientencluster. Die neun P3-Patientencluster sind fett gedruckt und mit einem ‚\*‘ gekennzeichnet und grau hinterlegt.

Cluster	SNP-Abstand/Zeit <sup>a</sup>			
	M (95%KI)	Me (IQR)	Min	Max
<b>1*</b>	<b>791,50</b> <b>(170,02-1412,98)</b>	<b>272,53</b> <b>(149,10-915,76)</b>	81,94	3285,00
<b>2*</b>	<b>969,69</b> <b>(493,35-1446,03)</b>	<b>292,00</b> <b>(135,47-941,59)</b>	67,73	13505,00
3	762,59 (635,28-889,89)	749,55 (694,15-844,06)	690,89	860,36
<b>4*</b>	<b>10632,68</b> <b>(1292,26-19973,11)</b>	<b>2018,24</b> <b>(138,49-26645,00)</b>	123,44	38325,00
5	982,49 (-94,58-2059,56)	544,32 (313,93-1682,20)	271,14	2946,07
<b>6*</b>	<b>3764,94</b> <b>(-423,58-7953,47)</b>	<b>1440,27</b> <b>(1235,80-7456,43)</b>	1144,32	7612,86
<b>7*</b>	<b>2822,22</b> <b>(786,35-4858,09)</b>	<b>709,72</b> <b>(284,58-1758,64)</b>	152,00	21352,50
<b>8*</b>	<b>3640,63</b> <b>(183,84-7097,42)</b>	<b>232,36</b> <b>(153,47-1467,24)</b>	76,84	51830,00
9	2548,10 (1688,53-3407,68)	2478,54 (1877,50-2936,59)	1787,76	4081,36
<b>10*</b>	<b>3929,26</b> <b>(2071,42-5787,10)</b>	<b>2641,90</b> <b>(1040,93-5944,29)</b>	703,29	11315,00
<b>11*</b>	<b>4774,78</b> <b>(3175,22-6374,34)</b>	<b>2253,48</b> <b>(1505,63-4161,00)</b>	705,99	56210,00
12	6942,30 (-6378,73-20263,33)	609,82 (483,22-2315,26)	313,17	59860,00
13	11039,21 (-14734,23-36812,65)	865,11 (320,46-17337,50)	304,17	61137,50
<b>14*</b>	<b>13542,53</b> <b>(4192,52-22892,54)</b>	<b>4493,28</b> <b>(2565,43-14660,83)</b>	1920,60	113515,00
15	3373,97 (-440,30-7188,25)	1383,16 (552,95-2611,15)	148,80	27922,50
16	22613,59 (-16533,06-61760,23)	8336,60 (7517,67-44847,99)	7270,80	78961,67
17	15491,37 (3892,08-27090,67)	6022,50 (3251,82-11588,75)	1784,44	145087,50

\* Es handelt sich um ein P3-Patientencluster

<sup>a</sup> Mittelwert (M), Median (Me), Minimal- (Min) und Maximalwert (Max) der SNP-Abstände unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) eines Clusters



**Abb. 10:** Verteilung des mittleren SNP-Abstands unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) innerhalb der 17 Patientencluster, aufgeschlüsselt nach Cluster. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Die Ausreißer sind zum Zweck der besseren Skalierung nicht vollständig dargestellt.

### 3.3 Unterschiede des SNP-Abstands zwischen gleichem und ungleichem Probenmaterial

Zur Untersuchung von Hypothese 1a) bzw. 1b), ob ein signifikanter Unterschied im SNP-Abstand (in SNPs) bzw. im SNP-Abstand/Zeit (in SNPs/Jahr) zwischen gleichem und ungleichem Probenmaterial eines Clusters bestand, wurden die neun P3-Patientencluster verwendet.

#### 3.3.1 SNP-Abstand ohne Berücksichtigung des Zeitfaktors

In diesen Ergebnissen zur Untersuchung von Hypothese 1a) wurden die zusammenfassenden Statistiken nach Patientencluster aufgeschlüsselt miteinander verglichen. Der Zeitfaktor, also der zum SNP-Abstand zweier Stämme zugehörige zeitliche Abstand ihrer Isolation in Tagen, wurde dabei nicht berücksichtigt. Dieser wird in Kap. 3.3.2 miteinbezogen. Es wurden jeweils

die SNP-Abstände, die zwischen Stämmen im gleichen Probenmaterial nachgewiesen wurden, mit den Stämmen, die in verschiedenen Probenmaterialien nachgewiesen wurden, verglichen. Eine übersichtliche Darstellung aller Werte zeigt Tab. 7, graphisch dargestellt werden die Werte in den Abb. 11-19.

Aufgrund der besseren Skalierung und da nur Werte innerhalb eines Clusters verglichen wurden, wurde für jedes Cluster ein eigenes Diagramm erstellt.

**Tab. 7:** Übersicht über den SNP-Abstand (in SNPs) der neun P3-Patientencluster, aufgeschlüsselt nach Cluster sowie gleichem und ungleichem Probenmaterial

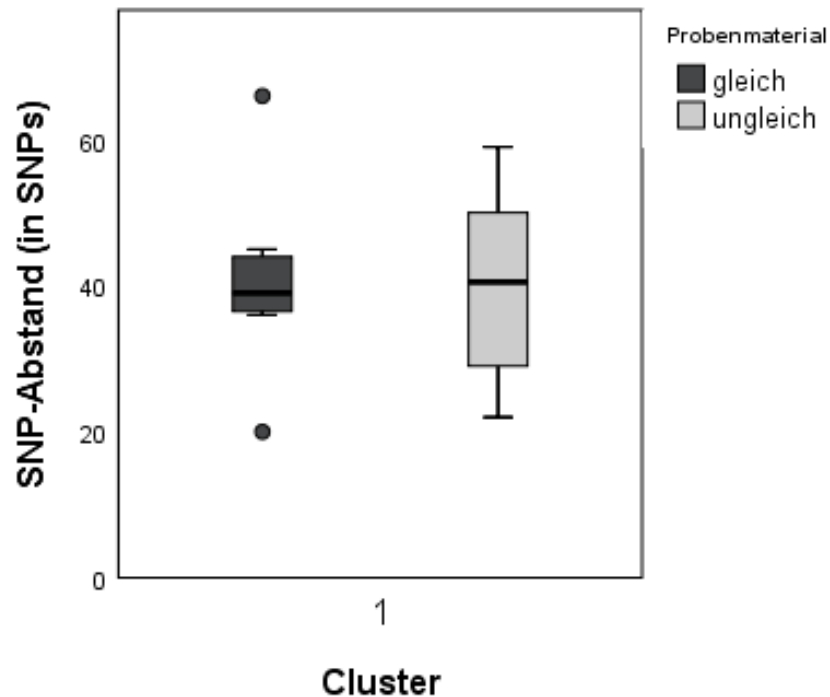
Cluster	Probenmaterial	N	SNP-Abstand <sup>a</sup>			
			M (95%KI)	Me (IQR)	Min	Max
1	Gleich	7	40,86 (28,16-53,55)	39,00 (36-45)	20	66
	Ungleich	8	40,00 (29,09-50,91)	40,50 (28,5-52,5)	22	59
	Insgesamt	15	40,40 (33,26-47,54)	39,00 (30-45)	20	66
2	Gleich	8	40,13 (12,45-67,8)	29,00 (24,25-36,75)	22	121
	Ungleich	58	42,09 (32,58-51,6)	28,00 (23-33,25)	15	128
	Insgesamt	66	41,85 (33,1-50,59)	28,00 (23-33,25)	15	128
4	Gleich	4	85,50 (51,11-118,89)	80,50 (68,75-107,25)	65	116
	Ungleich	11	110,91 (102,13-119,69)	107,00 (103-122)	92	135
	Insgesamt	15	104,13 (93,66-114,6)	107,00 (92-116)	65	135
6	Gleich	3	150,67 (130,59-170,75)	146,00*	146	160
	Ungleich	3	122,67 (85,05-160,29)	116,00*	112	140
	Insgesamt	6	136,67 (116,95-156,93)	143,00 (115-149,5)	112	160
7	Gleich	7	170,14 (116,48-223,81)	175,00 (112-184)	100	277
	Ungleich	21	132,71 (113,17-152,26)	125,00 (95-159)	76	234
	Insgesamt	28	142,07 (123,13-161,02)	132,00 (102-171,75)	76	277
8	Gleich	17	141,06 (125,42-156,7)	139,00 (115,5-166)	98	206
	Ungleich	28	143,86 (128,86-158,86)	139,50 (121-163)	76	247
	Insgesamt	45	142,80 (132,15-153,45)	139,00 (119,5-163)	76	247
10	Gleich	10	184,00 (164,39-203,61)	176,00 (162,25-217,5)	152	228
	Ungleich	5	165,80 (136,66-194,94)	158,00 (151,5-184)	149	207
	Insgesamt	15	177,93 (163,08-192,78)	169,00 (157-207)	149	228
11	Gleich	32	187,09 (171,42-202,77)	175,00 (155,75-208,75)	135	295
	Ungleich	73	181,47 (173,43-189,5)	180,00 (156,5-202,5)	107	275
	Insgesamt	105	183,18 (175,96-190,4)	176,00 (156,5-204,5)	107	295
14	Gleich	15	259,53 (221,93-297,14)	240,00 (207-311)	171	402
	Ungleich	13	243,46 (211,38-275,54)	231,00 (218,5-260,5)	176	392
	Insgesamt	28	252,07 (228,45-275,69)	239,00 (216,25-279,75)	171	402

\* Kein IQR, da die Anzahl (N) an Werten zu gering ist

<sup>a</sup> Mittelwert (M), Median (Me), Minimal- (Min) und Maximalwert (Max) der SNP-Abstände (in SNPs) eines Clusters

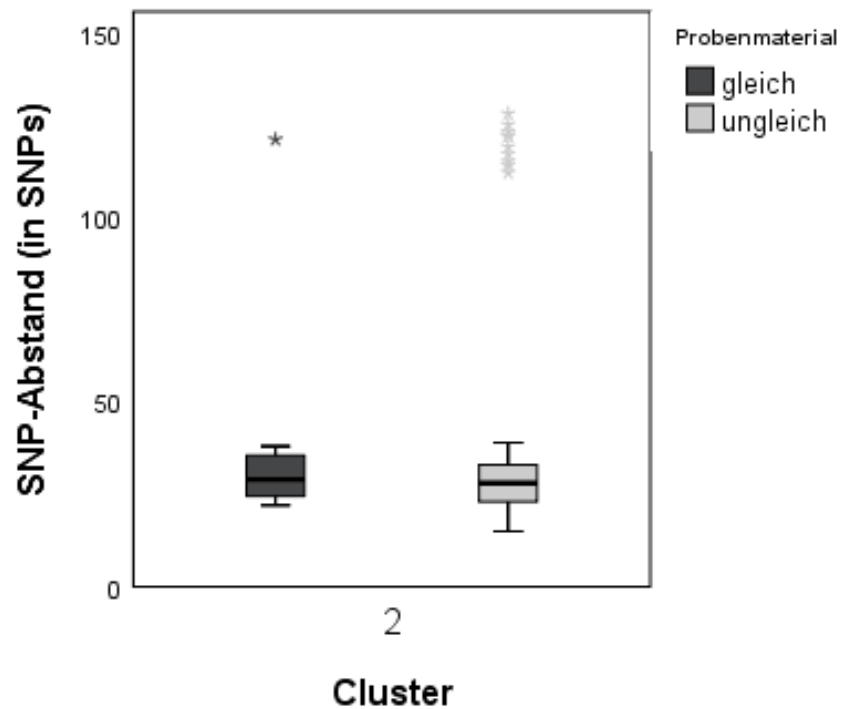


Für das Cluster 1 ergab sich bei einem medianen SNP-Abstand von 39 (IQR 36-45) SNPs bei gleichem und einem medianen SNP-Abstand von 40,5 (IQR 28,5-52,5) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 28$ ,  $Z = 0$ ,  $p = 1,000$ ).



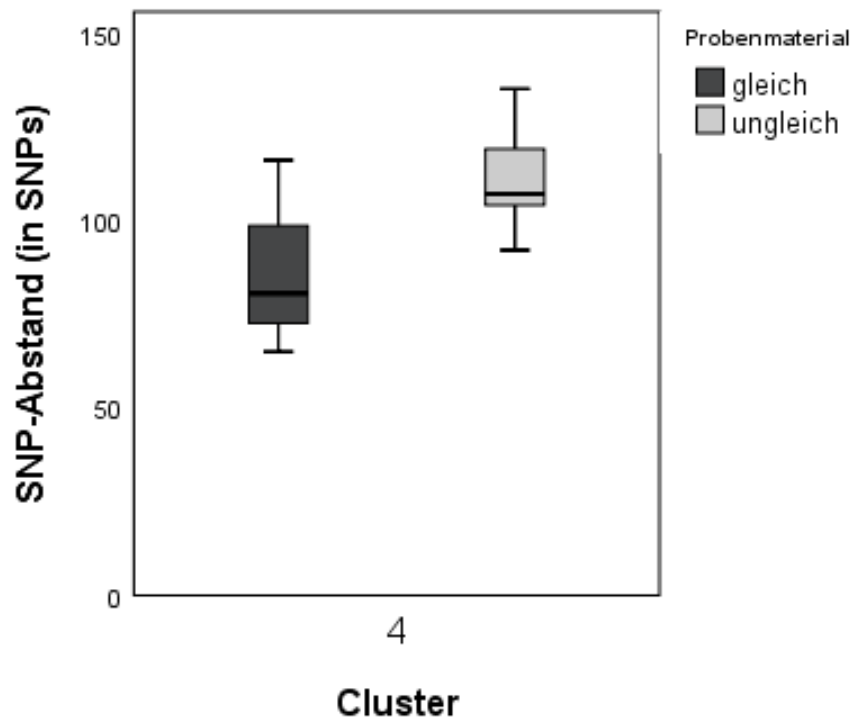
**Abb. 11:** SNP-Abstände von Cluster 1, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 2 ergab sich bei einem medianen SNP-Abstand von 29 (IQR 24,25-36,75) SNPs bei gleichem und einem medianen SNP-Abstand von 28 (IQR 23-33,25) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 203,5$ ,  $Z = -0,561$ ,  $p = 0,575$ ).



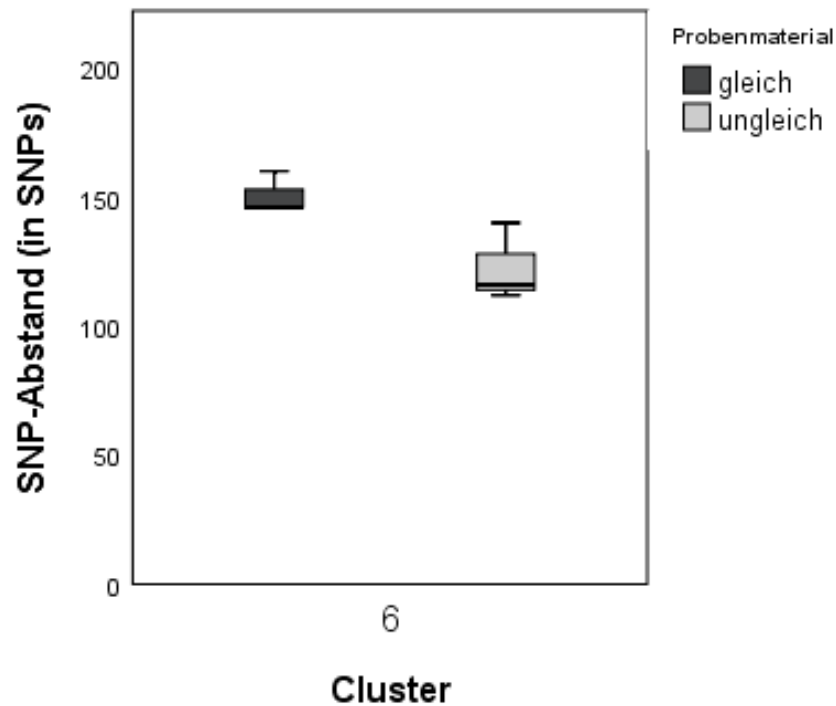
**Abb. 12:** SNP-Abstände von Cluster 2, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 4 ergab sich bei einem medianen SNP-Abstand von 80,5 (IQR 68,75-107,25) SNPs bei gleichem und einem medianen SNP-Abstand von 107 (IQR 103-122) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 7,5$ ,  $Z = -1,896$ ,  $p = 0,056$ ).



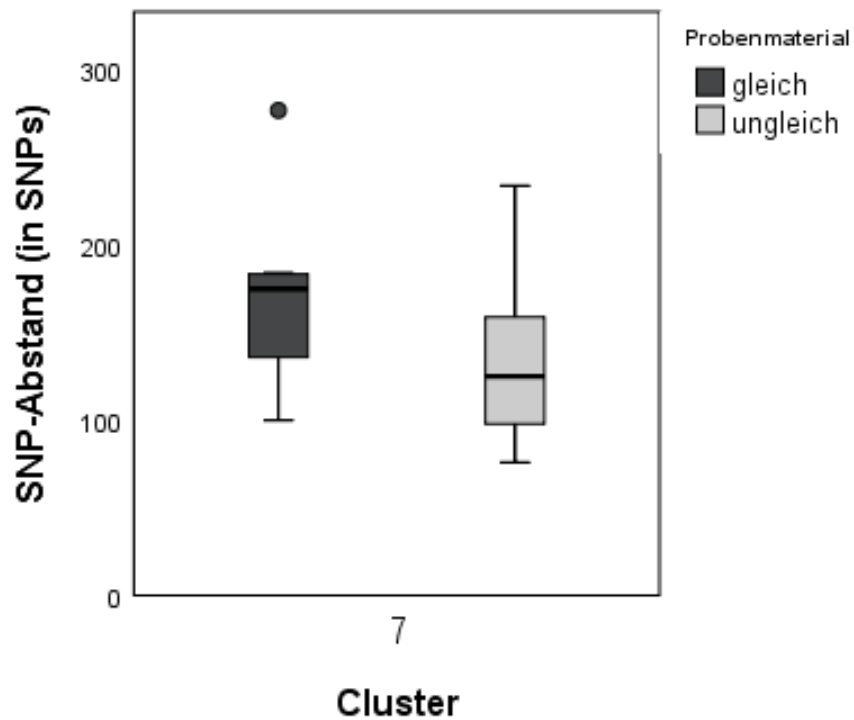
**Abb. 13:** SNP-Abstände von Cluster 4, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 6 ergab sich bei einem medianen SNP-Abstand von 146 SNPs bei gleichem und einem medianen SNP-Abstand von 116 SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 0$ ,  $Z = -1,993$ ,  $p = 0,100$ ).



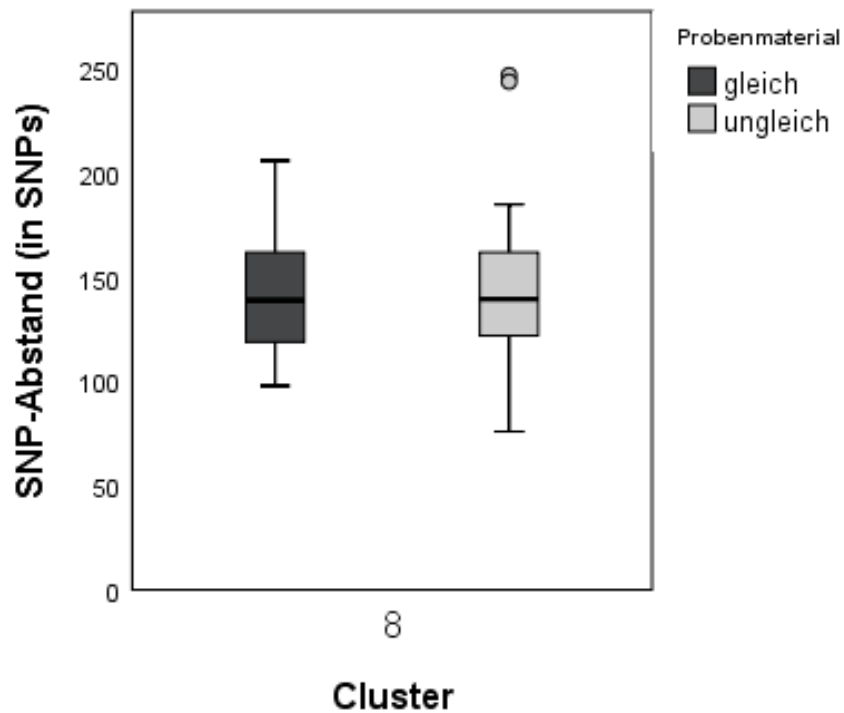
**Abb. 14:** SNP-Abstände von Cluster 6, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 7 ergab sich bei einem medianen SNP-Abstand von 175 (IQR 112-184) SNPs bei gleichem und einem medianen SNP-Abstand von 125 (IQR 95-159) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 41,5$ ,  $Z = -1,698$ ,  $p = 0,090$ ).



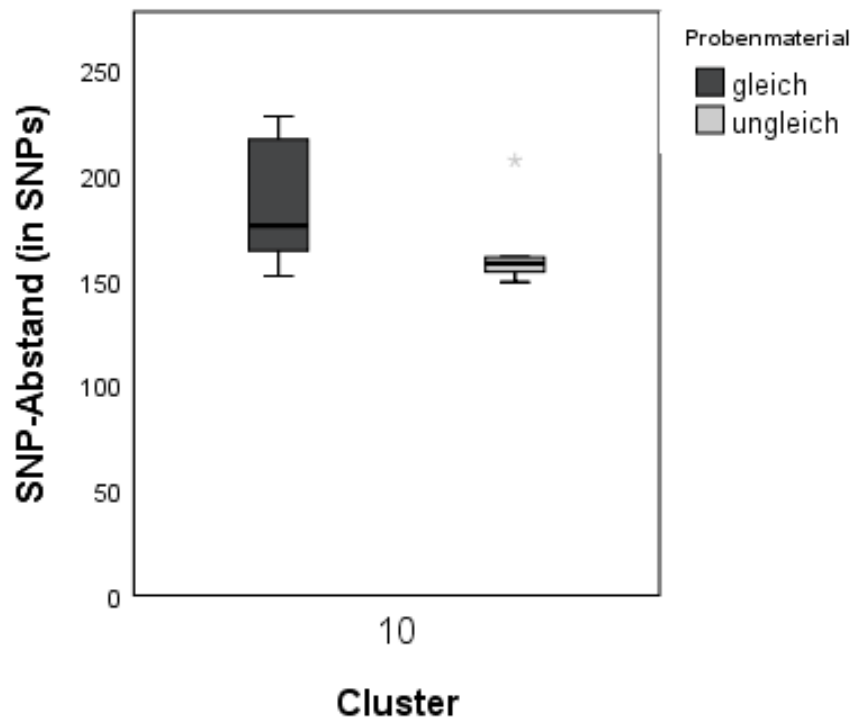
**Abb. 15:** SNP-Abstände von Cluster 7, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 8 ergab sich bei einem medianen SNP-Abstand von 139 (IQR 115,5-166) SNPs bei gleichem und einem medianen SNP-Abstand von 139,5 (IQR 121-163) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 232,5$ ,  $Z = -0,129$ ,  $p = 0,898$ ).



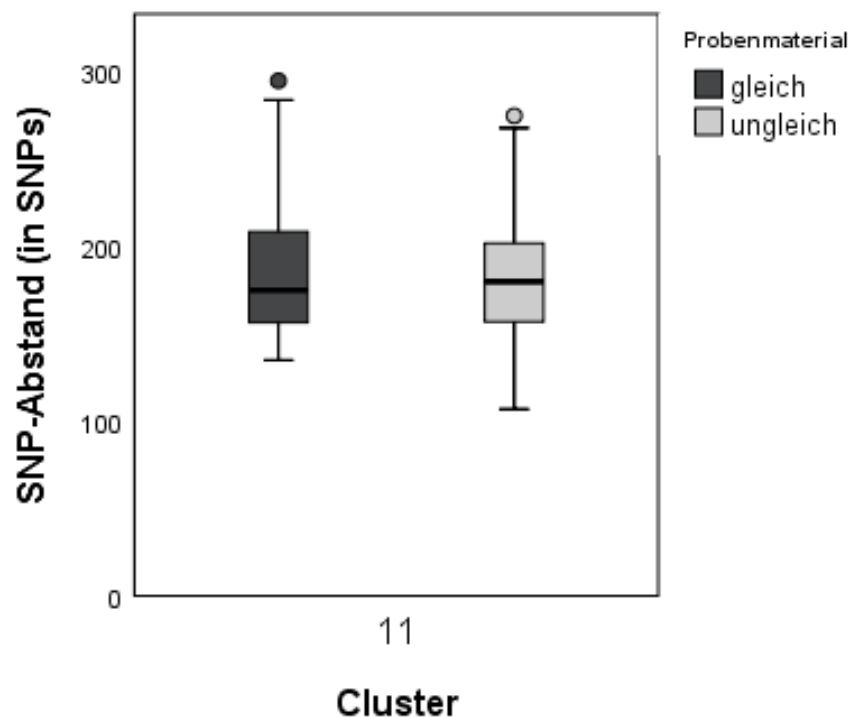
**Abb. 16:** SNP-Abstände von Cluster 8, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 10 ergab sich bei einem medianen SNP-Abstand von 176 (IQR 162,25-217,5) SNPs bei gleichem und einem medianen SNP-Abstand von 158 (IQR 151,5-184) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 12$ ,  $Z = -1,592$ ,  $p = 0,129$ ).



**Abb. 17:** SNP-Abstände von Cluster 10, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

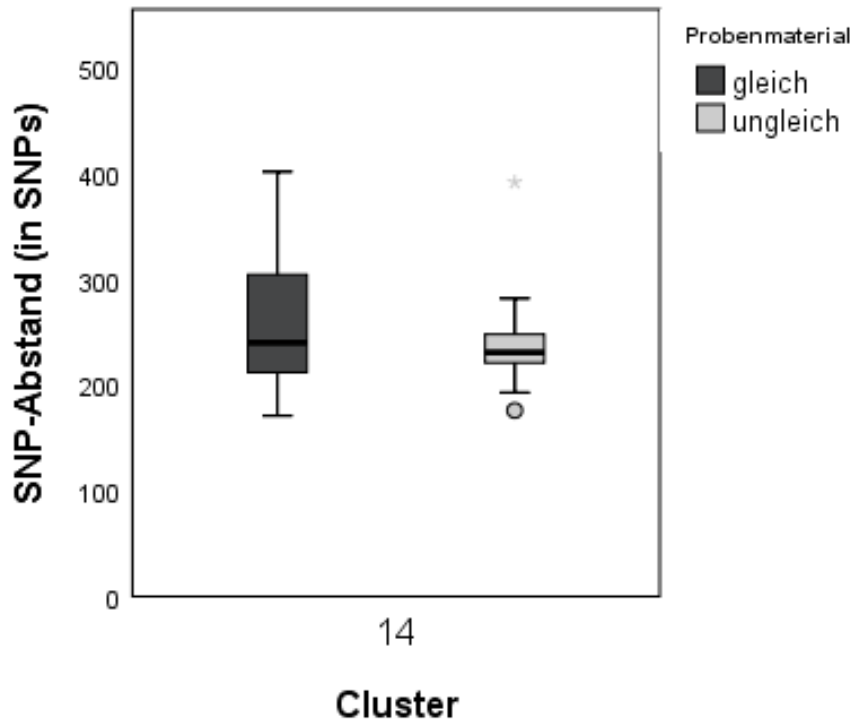
Für das Cluster 11 ergab sich bei einem medianen SNP-Abstand von 175 (IQR 155,75-208,75) SNPs bei gleichem und einem medianen SNP-Abstand von 180 (IQR 156,5-202,5) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 1139,5$ ,  $Z = -0,198$ ,  $p = 0,843$ ).



**Abb. 18:** SNP-Abstände von Cluster 11, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 14 ergab sich bei einem medianen SNP-Abstand von 240 (IQR 207-311) SNPs bei gleichem und einem medianen SNP-Abstand von 231 (IQR 218,5-260,5) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 89$ ,  $Z = -0,438$ ,  $p = 0,683$ ).





**Abb. 19:** SNP-Abstände von Cluster 14, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Zusammenfassend ließ sich in keinem der neun P3-Patientencluster ein signifikanter Unterschied zwischen den SNP-Abständen zwischen gleichen und ungleichem Probenmaterial nachweisen und Hypothese 1a) somit nicht bestätigen. Eine zusammenfassende Übersicht der Ergebnisse des Mann-Whitney-U-Tests ist in Tab. 8 dargestellt.

**Tab. 8:** Werte des Mann-Whitney-U-Tests zum Vergleich des SNP-Abstands bei gleichem und ungleichem Probenmaterial der neun P3-Patientencluster

Cluster	SNP-Abstand				Interpretation
	U	Z	Asymptomatische Signifikanz	Exakte Signifikanz	
1	28,000	-0,000	1,000	1,000	Nicht signifikant
2	203,500	-0,561	0,575	-	Nicht signifikant
4	7,500	-1,896	0,058	0,056	Nicht signifikant
6	0,000	-1,993	0,046	0,100	Nicht signifikant
7	41,500	-1,698	0,089	0,090	Nicht signifikant
8	232,500	-0,129	0,898	-	Nicht signifikant
10	12,000	-1,592	0,111	0,129	Nicht signifikant
11	1139,500	-0,198	0,843	-	Nicht signifikant
14	88,000	-0,438	0,662	0,683	Nicht signifikant

### 3.3.2 SNP-Abstand mit Berücksichtigung des Zeitfaktors

In diesen Ergebnissen zur Untersuchung von Hypothese 1b) werden die zusammenfassenden Statistiken nach Patientencluster aufgeschlüsselt miteinander verglichen. Der Zeitfaktor, also der zeitliche Abstand zwischen der Isolation von zwei Isolaten, wurde dabei berücksichtigt. Es wurden jeweils die SNP-Abstände/Zeit (in SNPs/Jahr), die zwischen Stämmen im gleichen Probenmaterial detektiert wurden, mit den Stämmen, die in verschiedenen Probenmaterialien detektiert wurden, verglichen. Es muss an dieser Stelle angemerkt werden, dass jene Werte des 95%-Konfidenzintervalls, die im negativen Bereich liegen, nicht realistisch sind, da nicht mit einem über die Zeit geringer werdenden SNP-Abstand zu rechnen ist. Eine übersichtliche

Darstellung aller Werte zeigt Tab. 9, graphisch dargestellt werden die Werte in den Abb. 20-28. Aufgrund der besseren Skalierung und da nur Werte innerhalb eines Clusters verglichen werden sollen, wurde für jedes Cluster eine eigene Grafik erstellt.

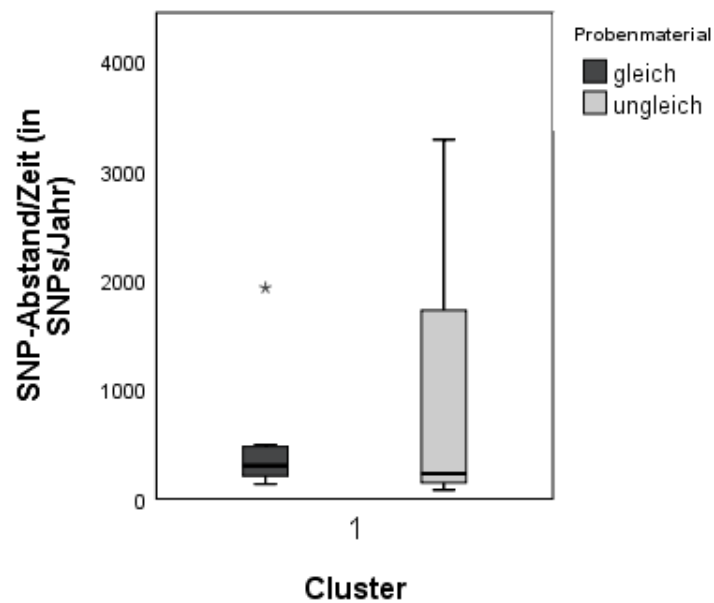
**Tab. 9:** Übersicht über den SNP-Abstand unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) der neun P3-Patientencluster, aufgeschlüsselt nach gleichem und ungleichem Probenmaterial.

Cluster	Probenmaterial	N	SNP-Abstand/Zeit*			
			M (95%KI)	Me (IQR)	Min	Max
1	Gleich	7	534,18 (-48,34-116,70)	304,17 (176,61-495,00)	134,08	1929,29
	Ungleich	7	1048,82 (-245,65-2343,29)	231,56 (147,97-2867,86)	81,94	3285,00
	Insgesamt	14	791,50 (170,02-1412,98)	272,53 (149,10-915,76)	81,94	3285,00
2	Gleich	8	423,92 (81,30-766,54)	262,76 (117,07-720,11)	88,24	1251,43
	Ungleich	58	1044,97 (505,25-1584,68)	292,00 (135,47-999,78)	67,73	13505,00
	Insgesamt	66	969,69 (493,35-1446,03)	292,00 (135,47-941,59)	67,73	13505,00
4	Gleich	3	17804,48 (-20918,66-56527,62)	23725,00*	123,44	29565,00
	Ungleich	10	8481,14 (-2653,23-19615,51)	1996,76 (139,44-11026,04)	129,88	38325,00
	Insgesamt	13	10632,68 (1292,26-19973,11)	2018,24 (13849-26645,00)	123,44	38325,00
6	Gleich	3	3460,13 (-5474,83-12395,10)	1440,27*	1327,27	7612,86
	Ungleich	2	4222,16 (-34885,48-43329,80)	4222,16*	1144,32	7300,00
	Insgesamt	5	3764,94 (-423,58-7953,47)	1440,27 (1235,80-7456,43)	1144,32	7612,86
7	Gleich	7	2744,26 (-2949,59-8438,12)	343,89 (264,41-709,72)	167,82	16699,75
	Ungleich	20	2849,50 (543,77-5155,23)	1192,80 (298,23-1974,61)	152,00	21352,50
	Insgesamt	27	2822,22 (786,35-4858,09)	709,72 (284,58-1758,64)	152,00	21352,50
8	Gleich	17	5464,00 (-1489,10-12417,09)	297,19 (150,47-1899,93)	120,51	48910,00
	Ungleich	27	2492,58 (-1417,89-6403,06)	202,08 (155,81-1042,86)	76,84	51830,00
	Insgesamt	44	3640,63 (183,84-7097,42)	232,36 (153,47-1467,24)	76,84	51830,00
10	Gleich	10	5383,14 (3075,36-7690,94)	4249,64 (2558,26-8694,82)	2283,86	11315,00
	Ungleich	5	1021,50 (588,38-1454,63)	891,56 (783,74-1324,24)	703,29	1607,55
	Insgesamt	15	3929,26 (2071,42-5787,10)	2641,90 (1040,93-5944,29)	703,29	11315,00
11	Gleich	31	3223,28 (2337,78-4108,78)	2294,29 (1452,84-3771,67)	932,29	9020,71
	Ungleich	72	5442,79 (3185,60-7699,97)	2199,13 (1516,93-4460,51)	705,99	56210,00
	Insgesamt	103	4774,78 (3175,22-6374,34)	2253,48 (1505,63-4161,00)	705,99	56210,00
14	Gleich	15	14203,05 (-1326,03-29732,14)	5037,00 (2534,14-10793,57)	2085,71	113515,00
	Ungleich	12	12716,88 (1381,23-24052,53)	3735,49 (2586,37-17243,21)	1920,60	64240,00
	Insgesamt	27	13542,53 (4192,52-22892,54)	4493,28 (2565,43-14660,83)	1920,60	113515,00

\* Kein IQR, da die Anzahl (N) an Werten zu gering ist

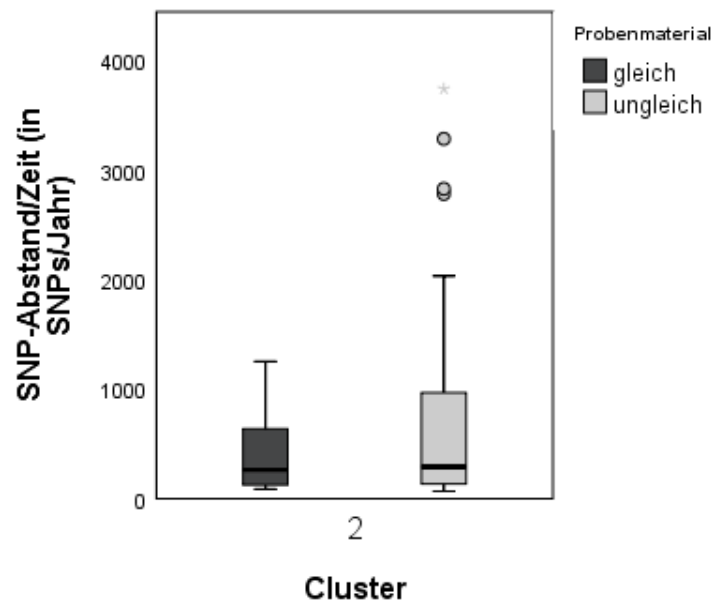
° Anzahl der Werte (N), Mittelwert (M), Median (Me), Minimal- (Min) und Maximalwert (Max) der SNP-Abstände unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) eines Clusters

Für das Cluster 1 ergab sich bei einem medianen SNP-Abstand/Zeit von 304,17 (IQR 176,61-495) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 231,56 (IQR 147,97-2867,86) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 24$ ,  $Z = -0,064$ ,  $p = 1,000$ ).



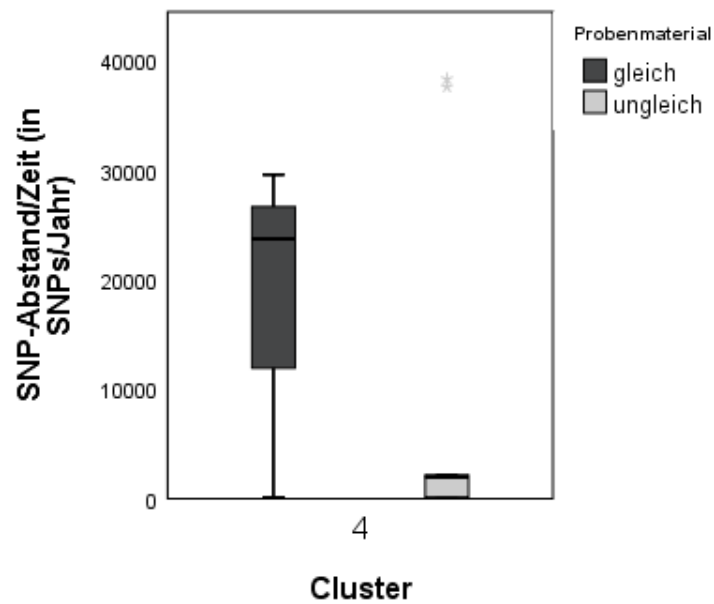
**Abb. 20:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 1, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 2 ergab sich bei einem medianen SNP-Abstand/Zeit von 262,76 (IQR 117,07-720,11) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 292 (IQR 135,47-999,78) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 199$ ,  $Z = -0,648$ ,  $p = 0,517$ ).



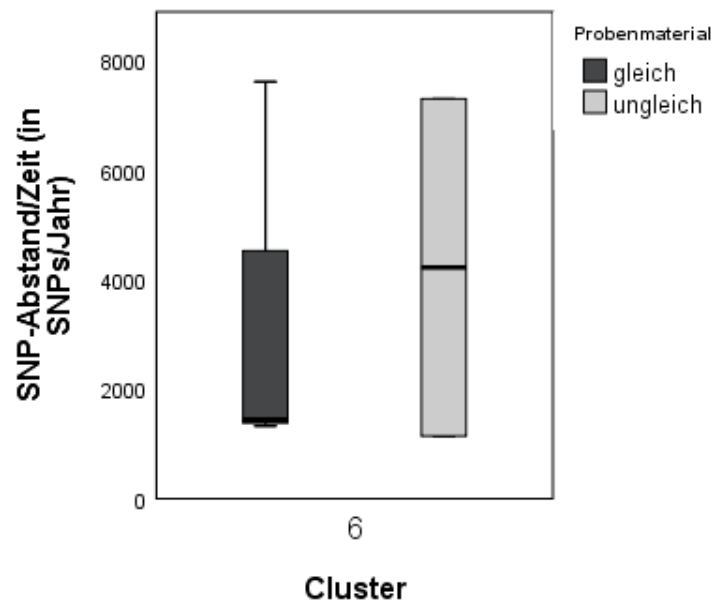
**Abb. 21:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 2, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.

Für das Cluster 4 ergab sich bei einem medianen SNP-Abstand/Zeit von 23725 SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 1996,76 (IQR 139,44-11026,04) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 14$ ,  $Z = -0,169$ ,  $p = 0,937$ ).



**Abb. 22:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 4, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

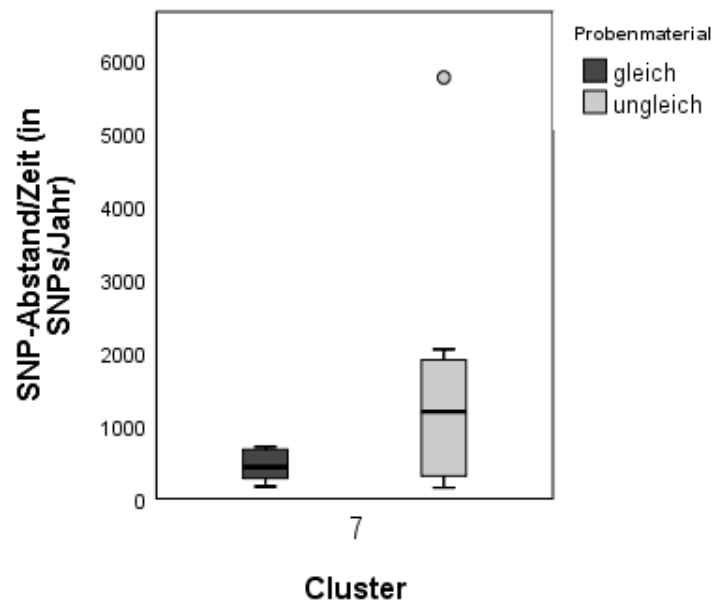
Für das Cluster 6 ergab sich bei einem medianen SNP-Abstand/Zeit von 1440,27 SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 4222,16 SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 2,000$ ,  $Z = -0,577$ ,  $p = 0,800$ ).



**Abb. 23:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 6, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

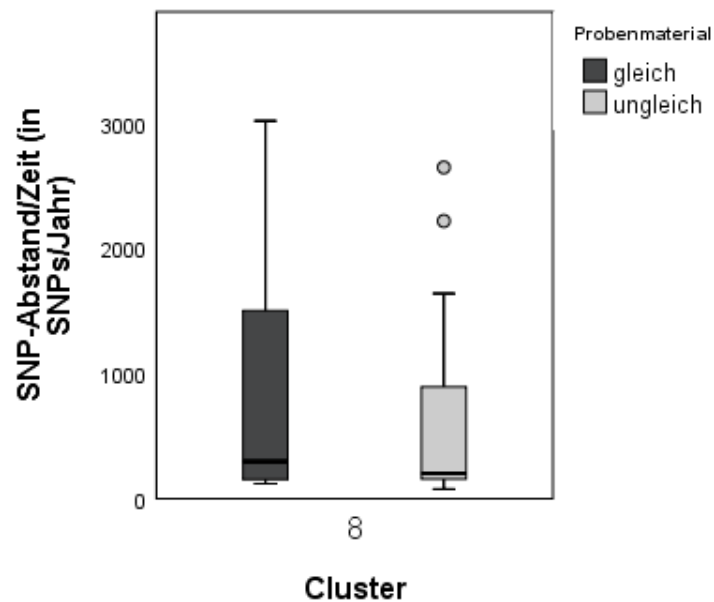
Für das Cluster 7 ergab sich bei einem medianen SNP-Abstand/Zeit von 343,89 (IQR 264,41-709,72) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 1192,8 (IQR 298,23-1974,61) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 52$ ,  $Z = -0,995$ ,  $p = 0,341$ ).





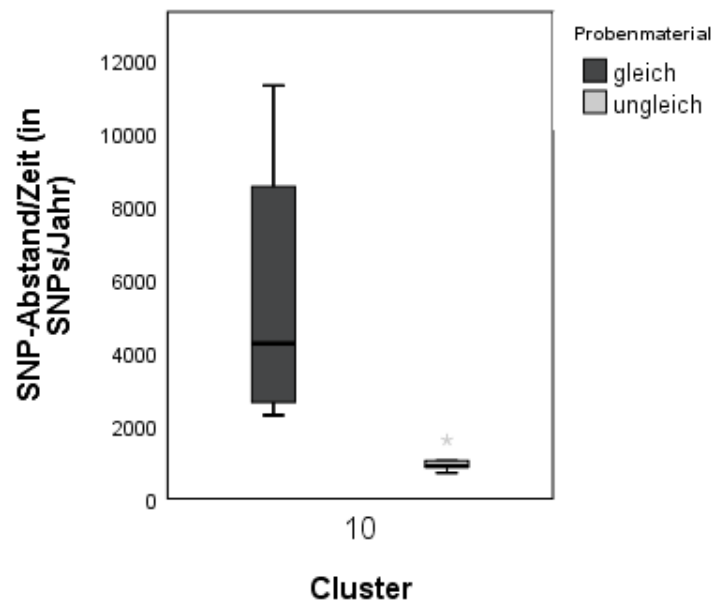
**Abb. 24:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 7, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.

Für das Cluster 8 ergab sich bei einem medianen SNP-Abstand/Zeit von 297,19 (IQR 150,47-1899,93) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 202,08 (IQR 155,81-1042,86) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 196$ ,  $Z = -0,807$ ,  $p = 0,419$ ).



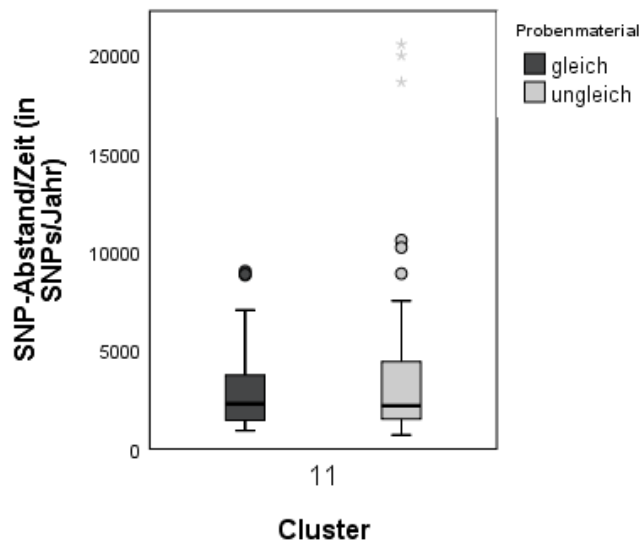
**Abb. 25:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 8, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.

Für das Cluster 10 ergab sich bei einem medianen SNP-Abstand/Zeit von 4249,64 (IQR 2558,26-8694,82) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 891,56 (IQR 783,74-1324,24) SNPs/Jahr bei ungleichem Probenmaterial ein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 0$ ,  $Z = -3,062$ ,  $p = 0,001$ ).



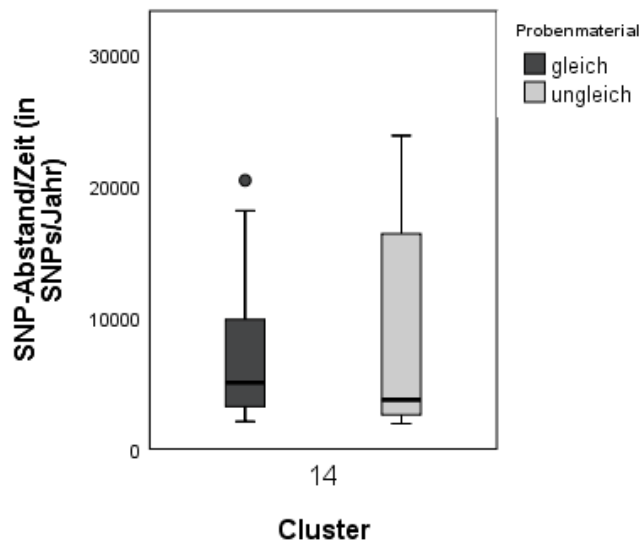
**Abb. 26:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 10, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

Für das Cluster 11 ergab sich bei einem medianen SNP-Abstand/Zeit von 2294,29 (IQR 1452,84-3771,67) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 2199,13 (IQR 1516,93-4460,51) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 1112$ ,  $Z = -0,029$ ,  $p = 0,977$ ).



**Abb. 27:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 11, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.

Für das Cluster 14 ergab sich bei einem medianen SNP-Abstand/Zeit von 5037 (IQR 2534,14-10793,57) SNPs/Jahr bei gleichem und einem medianen SNP-Abstand/Zeit von 3735,49 (IQR 2586,37-17243,21) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (exakter Mann-Whitney-U-Test:  $U = 87$ ,  $Z = -0,146$ ,  $p = 0,905$ ).



**Abb. 28:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 14, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.

Zusammenfassend konnte in Cluster 10 ein signifikanter Unterschied im SNP-Abstand/Zeit zwischen gleichem und ungleichem Probenmaterial nachgewiesen werden. In den übrigen Clustern waren die Unterschiede nicht signifikant, somit konnte die Hypothese 1b) nicht bestätigt. Eine zusammenfassende Übersicht der Ergebnisse des Mann-Whitney-U-Tests ist in Tab. 10 dargestellt.

**Tab. 10:** Werte des Mann-Whitney-U-Tests zum Vergleich des SNP-Abstands bei gleichem und ungleichem Probenmaterial der neun P3-Patientencluster

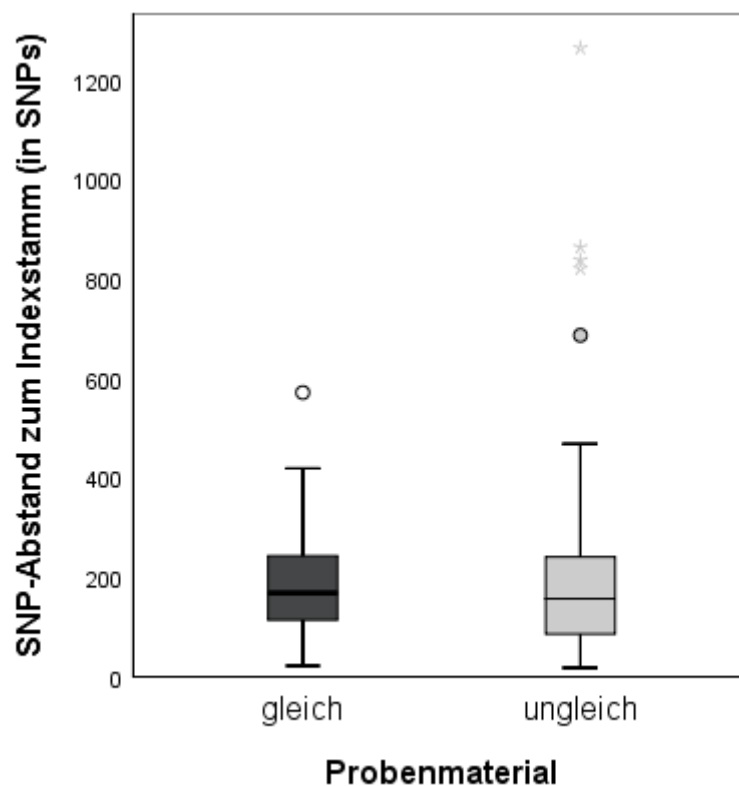
Cluster	SNP-Abstand/Zeit				Interpretation
	U	Z	Asymptomatische Signifikanz	Exakte Signifikanz	
1	24,000	-0,064	0,949	1,000	Nicht signifikant
2	199,000	-0,648	0,517	-	Nicht signifikant
4	14,000	-0,169	0,866	0,937	Nicht signifikant
6	2,000	-0,577	0,564	0,800	Nicht signifikant
7	52,000	-0,996	0,319	0,341	Nicht signifikant
8	196,000	-0,807	0,419	-	Nicht signifikant
10	0,000	-3,062	0,002	0,001	Signifikant
11	1112,000	-0,029	0,977	-	Nicht signifikant
14	87,000	-0,146	0,884	0,905	Nicht signifikant

### 3.4 Indexstamm-Verbindungen

In Anbetracht von Hypothese 2a) bzw. 2b) wurde nur der SNP-Abstand bzw. der SNP-Abstand/Zeit zwischen zwei Stämmen berechnet, wenn es sich bei einem der beiden Stämme um den Indexstamm des jeweiligen Clusters handelte, also den zeitlich erstdetektierten Stamm eines Clusters. Die Analyse wurde clusterübergreifend durchgeführt, also nicht nach Cluster aufgeschlüsselt, da die Datenmenge hierfür zu gering war. Es wurde der mediane SNP-Abstand zwischen gleichen und ungleichen Probenmaterialien miteinander verglichen.

### 3.4.1 SNP-Abstand ohne Berücksichtigung des Zeitfaktors

In der Untersuchung von Hypothese 2a) ergab sich bei einem medianen SNP-Abstand von 164 (IQR 121-206) SNPs bei gleichem und einem medianen SNP-Abstand von 140 (IQR 71,75-185) SNPs bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 938,5$ ,  $Z = -0,765$ ,  $p = 0,444$ ). Somit konnte die Hypothese 5a) nicht bestätigt werden. Graphisch dargestellt ist dies in Abb. 35.

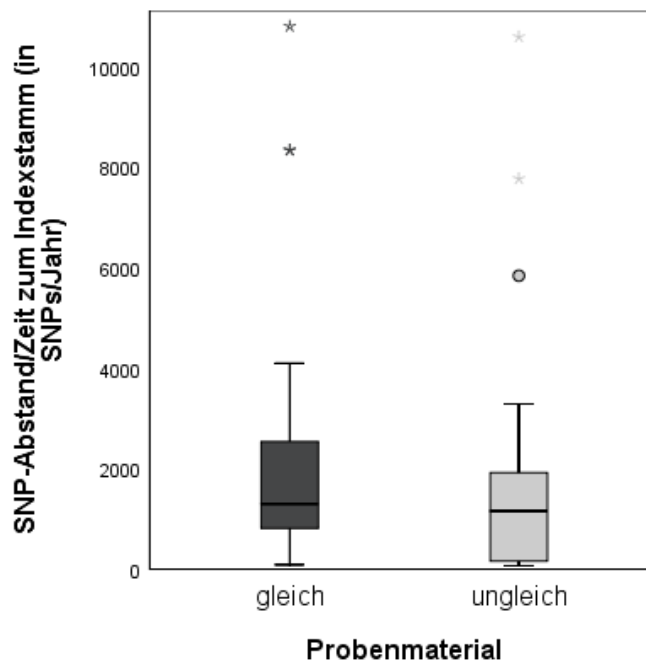


**Abb. 29:** SNP-Abstände (in SNPs) aller 17 Patientencluster zum jeweiligen Indexstamm, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.

### 3.4.2 SNP-Abstand mit Berücksichtigung des Zeitfaktors

In der Untersuchung von Hypothese 2b) ergab sich bei einem medianen SNP-Abstand/Zeit von 2153,5 (IQR 483,23-4449,67) SNPs/Jahr bei gleichem und

einem medianen SNP-Abstand/Zeit von 1478,59 (IQR 256,19-2836,35) SNPs/Jahr bei ungleichem Probenmaterial kein signifikanter Unterschied zwischen diesen beiden Gruppen (Mann-Whitney-U-Test:  $U = 846,0$ ,  $Z = -1,483$   $p = 0,138$ ). Somit konnte die Hypothese 5b) nicht bestätigt werden.



**Abb. 30:** SNP-Abstand/Zeit (in SNPs/Jahr) aller 17 Patientencluster zum jeweiligen Indexstamm, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstand. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.

### 3.4.3 Tabellarische Übersicht der deskriptiven Daten und des Mann-Whitney-U-Tests der Indexstamm-Verbindungen

Zur abschließenden Übersicht über die deskriptiven Daten und des Mann-Whitney-U-Tests unter ausschließlicher Berücksichtigung der Indexstamm-Verbindungen werden diese in Tab. 11 und Tab. 12 veranschaulicht.



**Tab. 11:** Deskriptive Daten des SNP-Abstands und des SNP-Abstands/Zeit bei gleichem und ungleichem Probenmaterial unter Berücksichtigung nur der Indexstamm-Verbindungen

		Indexstamm-Verbindungen				
Probenmaterial		N	M (95%KI)	Me (IQR)	Min	Max
SNP-Abstand (in SNPs)	Gleich	103	162,58 (147,46-177,71)	164,00 (121-206)	20	402
	Ungleich	226	139,16 (125,99-152,33)	140,00 (71,75-185)	15	649
SNP-Abstand/Zeit (in SNPs/Jahr)	Gleich	101	5443,77 (2859,13-8028,41)	2153,50 (483,23-4449,67)	88,24	11315,00
	Ungleich	218	4349,33 (2970,56-5728,10)	1478,59 (256,19-2836,35)	67,73	78961,67

Anzahl der Werte (N), Mittelwert (M), Median (Me), Minimal- (Min) und Maximalwert (Max)

**Tab. 12:** Werte des Mann-Whitney-U-Tests zum Vergleich des SNP-Abstands und des SNP-Abstands/Zeit bei gleichem und ungleichem Probenmaterial unter ausschließlicher Berücksichtigung der Indexstamm-Verbindungen

	Indexstamm-Verbindungen				
	U	Z	Asymptomatische Signifikanz	Exakte Signifikanz	Interpretation
SNP-Abstand	938,500	-0,765	0,444	-	Nicht signifikant
SNP-Abstand/Zeit	846,000	-1,483	0,138	-	Nicht signifikant

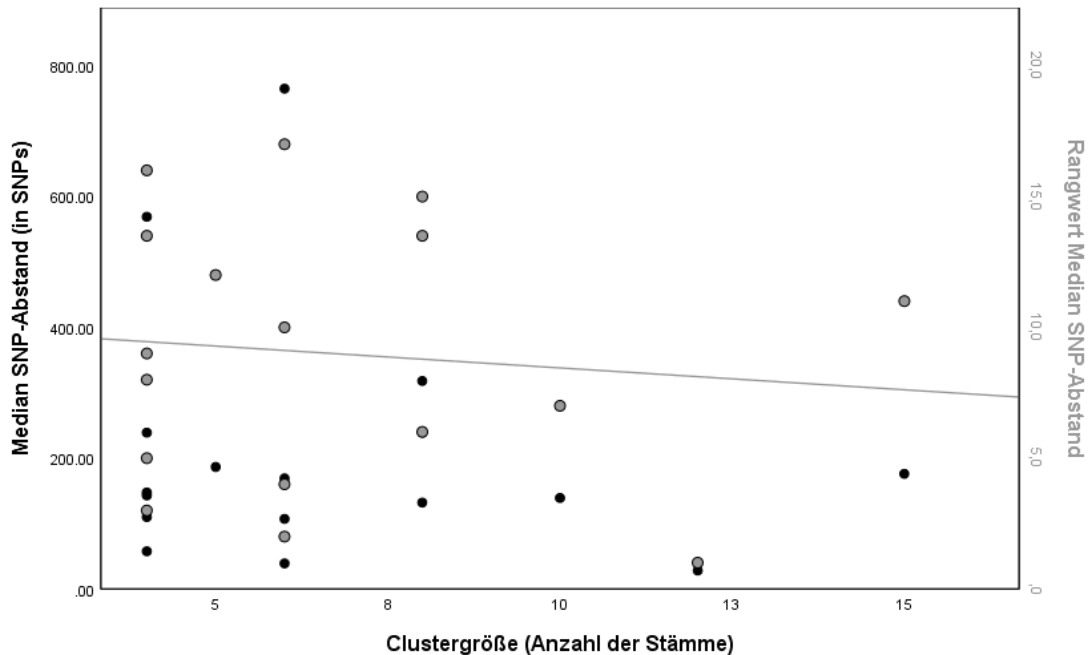
### 3.5 Korrelationsanalysen

Für die Korrelationsanalysen wurden die 17 Patientencluster verwendet. Unter der Annahme, dass der SNP-Abstand/Zeit einer konstanten Cluster-Mutationsrate unterliegt, wäre keine Korrelation zwischen SNP-Abstand/Zeit mit der Clustergröße, dem maximalen zeitlichen Abstand sowie mit der Anzahl der Probenmaterialien zu erwarten. Aus diesem Grund wurden die Korrelationsanalysen nur ohne Berücksichtigung des Zeitfaktors durchgeführt.

#### 3.5.1 Korrelation des SNP-Abstands mit der Clustergröße

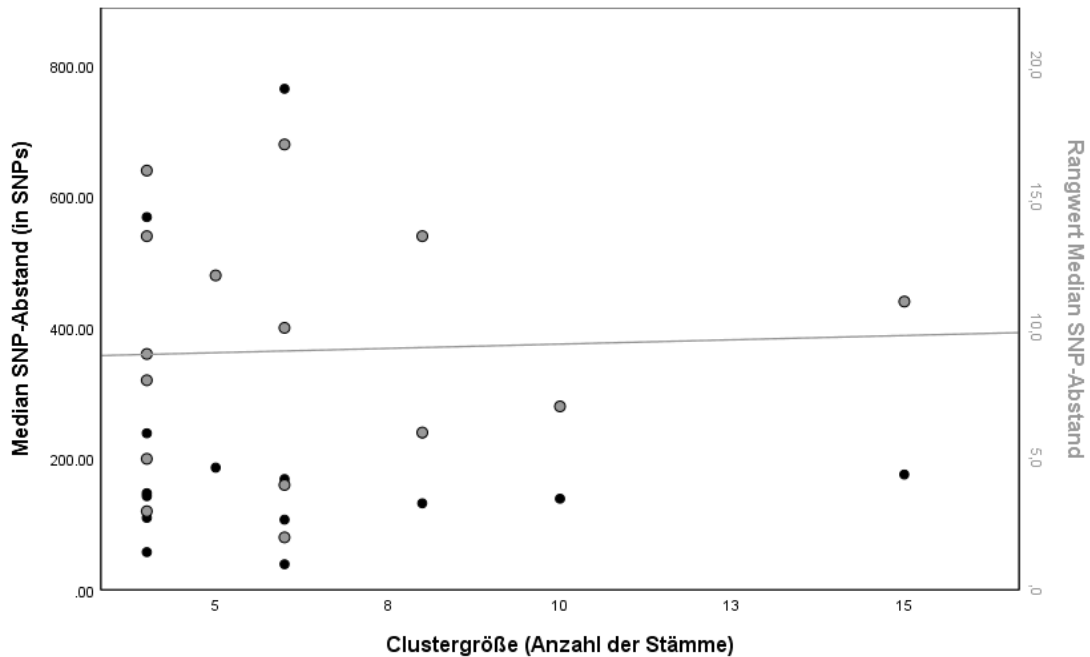
In Anbetracht von Hypothese 3 wurde der Zusammenhang zwischen der Anzahl an Stämmen in einem Patientencluster und dem medianen SNP-Abstand in

diesem Patientencluster untersucht. Es konnte hierbei keine signifikante Korrelation nachgewiesen werden (Spearman-Korrelationskoeffizient  $\rho(17) = -0,082$ ,  $p = 0,756$ ). Die Daten der bivariaten Korrelationsanalyse werden in Abb. 29 dargestellt.



**Abb. 31:** Darstellung der medianen SNP-Abstände (in SNPs) in schwarz und der zugehörigen Rangwerte des medianen SNP-Abstands in grau in Bezug zur Clustergröße (Anzahl der Stämme in einem Cluster) der 17 Patientencluster. Eine lineare Anpassungslinie der Rangwerte ist als zentrale Linie dargestellt.

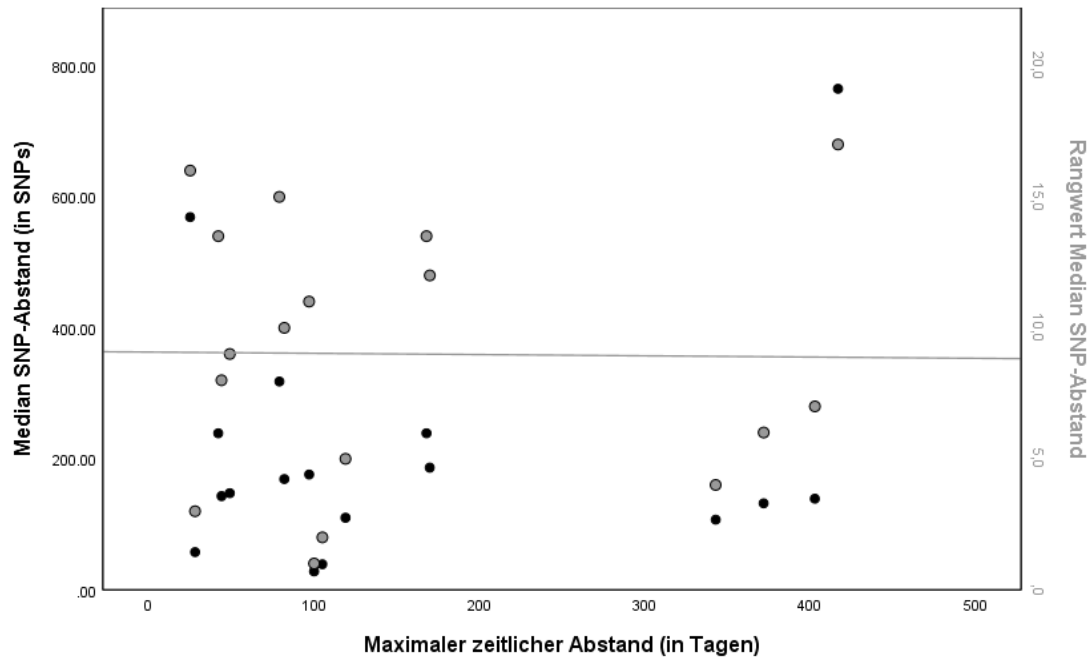
Dieselbe Berechnung wurde nach Ausschluss der Cluster 2 und 17 durchgeführt, da in diesen zwei Clustern der maximale SNP-Abstand mehr als das Dreifache des Medians der SNP-Abstände betrug und somit der Verdacht bestand, dass in diesen Clustern sogenannte Hypermulator-Stämme vorkamen, also Stämme mit einer signifikant erhöhten Mutationsrate (z.B. durch Schäden des DNA-Reparatursystems). Auch dann konnte jedoch kein signifikanter Zusammenhang festgestellt werden (Spearman-Korrelationskoeffizient  $\rho(15) = 0,007$ ,  $p = 0,979$ ). Die Daten der bivariaten Korrelationsanalyse werden nach Ausschluss von Cluster 2 und 17 in Abb. 30 dargestellt.



**Abb. 32:** Darstellung der medianen SNP-Abstände (in SNPs) in schwarz und der zugehörigen Rangwerte des medianen SNP-Abstands in grau in Bezug zur Clustergröße (Anzahl der Stämme in einem Cluster) von 15 der 17 Patientencluster nach Ausschluss von Cluster 2 (Patienten-ID 365) und Cluster 17 (Patienten-ID 149). Eine lineare Anpassungslinie der Rangwerte ist als zentrale Linie dargestellt.

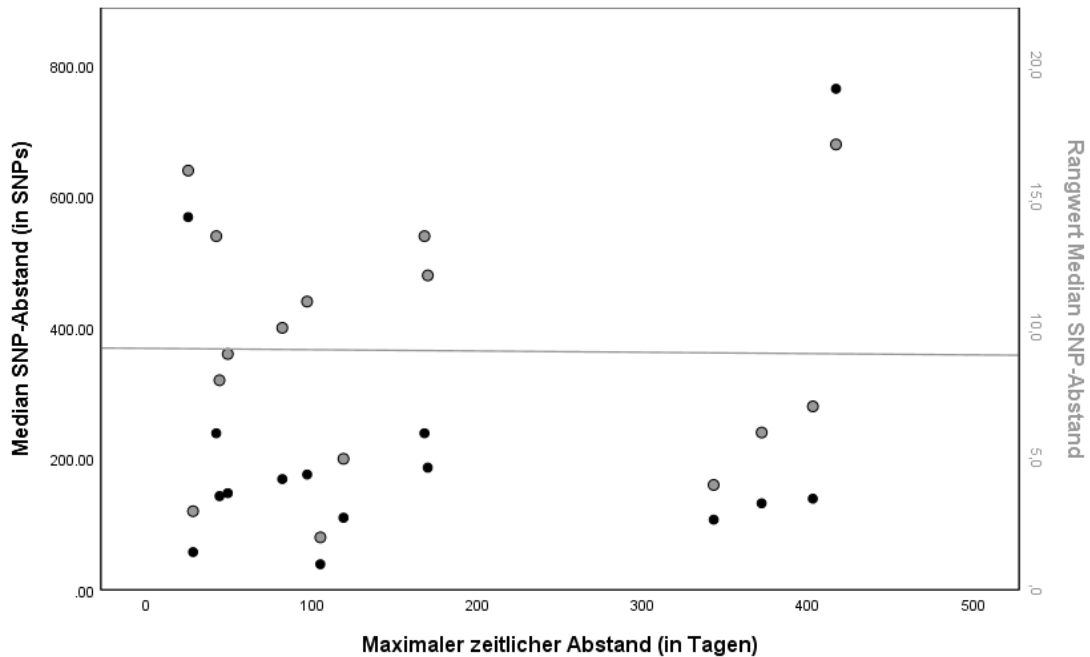
### 3.5.2 Korrelation des SNP-Abstands mit dem maximalen zeitlichen Abstand

In Anbetracht von Hypothese 4 wurde der Zusammenhang zwischen dem maximalen zeitlichen Abstand, also der Differenz der zeitlich frühesten und der zeitlich spätesten Isolation eines Stammes (in Tagen) innerhalb eines Patientenclusters, und dem medianen SNP-Abstand in diesem Patientencluster untersucht. Es konnte hierbei keine signifikante Korrelation nachgewiesen werden (Spearman-Korrelationskoeffizient  $\rho(17) = -0,099$ ,  $p = 0,704$ ). Die Daten der bivariaten Korrelationsanalyse werden in Abb. 31 dargestellt.



**Abb. 33:** Darstellung der medianen SNP-Abstände (in SNPs) in schwarz und der zugehörigen Rangwerte des medianen SNP-Abstands in grau in Bezug zum maximalen zeitlichen Abstand (in Tagen) der 17 Patientencluster. Eine lineare Anpassungslinie der Rangwerte ist als zentrale Linie dargestellt.

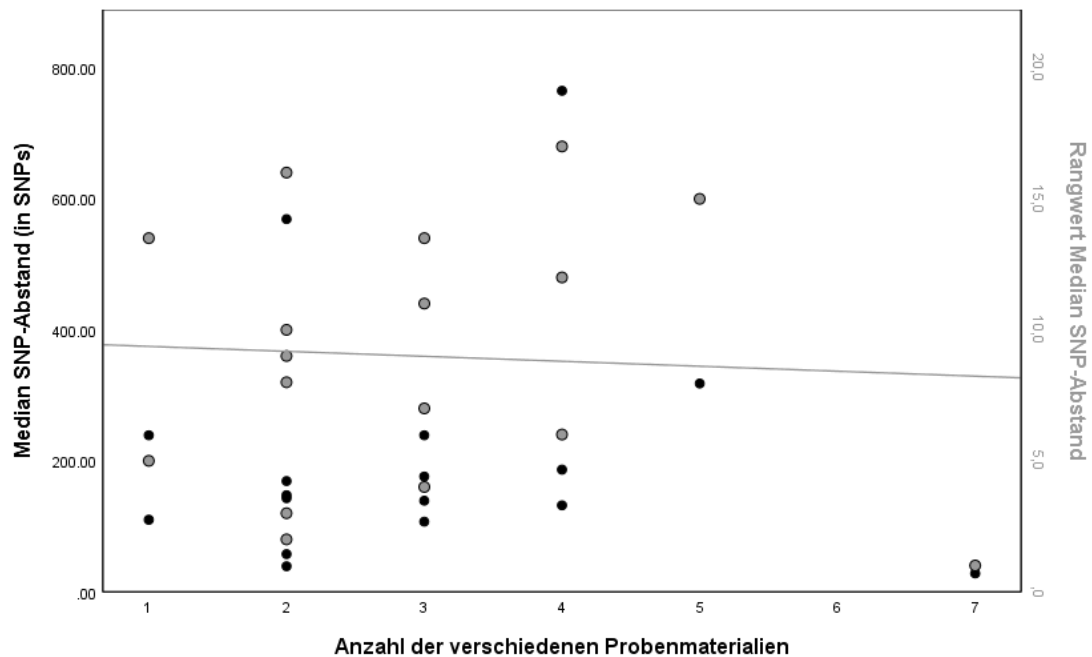
Dieselbe Berechnung wurde nach Ausschluss der Cluster 2 (Patienten-ID 365) und 17 (Patienten-ID 149) durchgeführt, da in diesen zwei Clustern der maximale SNP-Abstand mehr als das Dreifache des Medians der SNP-Abstände war und somit der Verdacht bestand, dass in diesen Clustern sogenannte Hypermulator-Stämme vorkamen, also Stämme mit einer signifikant erhöhten Mutationsrate (z.B. durch Schäden des DNA-Reparatursystems). Auch hier konnte kein signifikanter Zusammenhang festgestellt werden (Spearman-Korrelationskoeffizient  $\rho(15) = -0,045$ ,  $p = 0,874$ ). Die Daten der bivariaten Korrelationsanalyse werden nach Ausschluss von Cluster 2 und 17 in Abb. 32 dargestellt.



**Abb. 34:** Darstellung der medianen SNP-Abstände (in SNPs) in schwarz und der zugehörigen Rangwerte des medianen SNP-Abstands in grau in Bezug zum maximalen zeitlichen Abstand (in Tagen) von 15 der 17 Patientencluster nach Ausschluss von Cluster 2 (Patienten-ID 365) und Cluster 17 (Patienten-ID 149). Eine lineare Anpassungslinie der Rangwerte ist als zentrale Linie dargestellt.

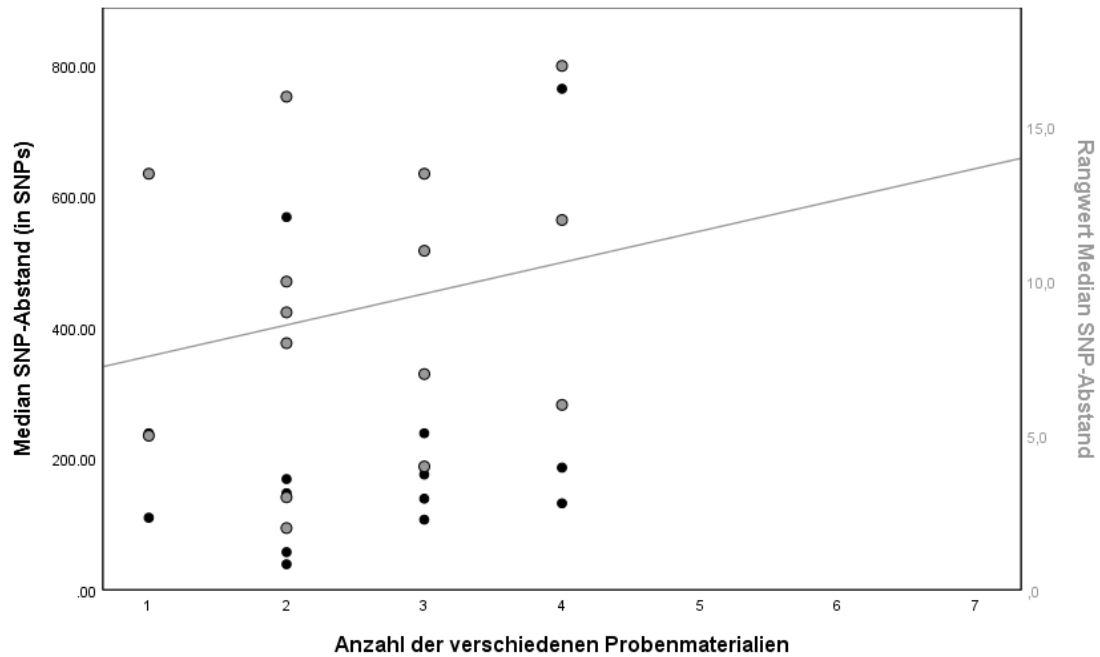
### 3.5.3 Korrelation des SNP-Abstands mit der Anzahl der Probenmaterialien

In Anbetracht von Hypothese 5 wurde der Zusammenhang zwischen der Anzahl an verschiedenen Probenmaterialien eines Clusters (Rektalabstrich, Rachenabstrich, etc.) und dem medianen SNP-Abstand in diesem Patientencluster untersucht. Es konnte hierbei keine signifikante Korrelation nachgewiesen werden (Spearman-Korrelationskoeffizient  $\rho(17) = 0,087$ ,  $p = 0,739$ ). Die Daten der bivariaten Korrelationsanalyse werden in Abb. 33 dargestellt. Die Berechnung der Anpassungslinie ist nur zur Darstellung einer Tendenz dargestellt und stimmt in diesem Fall nicht mit dem berechneten Korrelationskoeffizienten überein.



**Abb. 35:** Darstellung der medianen SNP-Abstände (in SNPs) in schwarz und der zugehörigen Rangwerte des medianen SNP-Abstands in grau in Bezug zur Anzahl der verschiedenen Probenmaterialien der 17 Patientencluster. Eine lineare Anpassungslinie der Rangwerte ist als zentrale Linie dargestellt.

Dieselbe Berechnung wurde nach Ausschluss der Cluster 2 (Patienten-ID 365) und 17 (Patienten-ID 149) durchgeführt, da in diesen zwei Clustern der maximale SNP-Abstand mehr als das Dreifache des Medians der SNP-Abstände war und somit der Verdacht bestand, dass in diesen Clustern sogenannte Hypermulator-Stämme vorkamen, also Stämme mit einer signifikant erhöhten Mutationsrate (z.B. durch Schäden des DNA-Reparatursystems). Auch hier konnte kein signifikanter Zusammenhang festgestellt werden (Spearman-Korrelationskoeffizient  $\rho(15) = 0,206$ ,  $p = 0,461$ ). Die Daten der bivariaten Korrelationsanalyse werden nach Ausschluss von Cluster 2 und 17 in Abb. 34 dargestellt.



**Abb. 36:** Darstellung der medianen SNP-Abstände (in SNPs) in schwarz und der zugehörigen Rangwerte des medianen SNP-Abstands in grau in Bezug zur Anzahl der verschiedenen Probenmaterialien von 15 der 17 Patientencluster nach Ausschluss von Cluster 2 (Patienten-ID 365) und Cluster 17 (Patienten-ID 149). Eine lineare Anpassungslinie der Rangwerte ist als zentrale Linie dargestellt.

## 4 Diskussion

In der hier vorliegenden Studie zur Untersuchung der Transmission von nosokomialen *P. aeruginosa*-Stämmen wurden 690 *P. aeruginosa*-Stämme, die in unterschiedlichen Probenmaterialien von 410 Patienten über einen Zeitraum von 54 Monaten nachgewiesen wurden, durch NGS sequenziert und einer phylogenetischen sowie nachfolgend deskriptiven statistischen Analyse unterzogen. Im Gegensatz zu vorangegangenen Studien zur Erforschung der genetischen Evolution von *P. aeruginosa*-Stämmen wurden in dieser Studie eine deutlich größere Anzahl an Stämmen eines vielfältigeren Patientenkollektivs untersucht. [29], [37], [38], [64]–[67] Diese Voraussetzungen ermöglichten die Erstellung eines umfangreichen sowie repräsentativen Datensatzes. Aus diesem konnten anhand definierter Einschlusskriterien 114 Stämme von 17 Patienten in 17 phylogenetische Cluster eingeteilt und für die Untersuchungen dieser Arbeit zur Mikroevolution von *P. aeruginosa* verwendet werden. Maßgebend für die Messung der Mikroevolution war die genetische Distanz gemessen am SNP-Abstand des Core-Genoms zwischen den *P. aeruginosa*-Stämmen innerhalb eines Patienten.

Im Rahmen der Studie wurden fünf Hypothesen zu Einflussfaktoren des SNP-Abstands auf ihre Richtigkeit untersucht. Aus diesen Ergebnissen konnten Rückschlüsse auf den Nutzen eines einheitlichen SNP-Grenzwertes zur Unterscheidung, ob ein Stamm durch Transmission erworben wurde oder sich durch Mutationen innerhalb des Patienten entwickelt hat, gezogen werden.

Lediglich in einem der 17 Cluster gab es einen signifikanten Unterschied des SNP-Abstands zwischen Stämmen aus gleichen gegenüber jenen aus verschiedenen Probenmaterialien. Der SNP-Abstand fiel entgegen der Erwartungen in diesem Cluster jedoch zwischen Stämmen des gleichen Probenmaterials signifikant größer aus als zwischen jenen unterschiedlichen Probenmaterials. Eine mögliche Erklärung hierfür ist, dass im Gastrointestinaltrakt im Gegensatz zum weniger bakteriell besiedeltem Milieu des Urethral-Traktes allgemein eine schnellere genetische Diversifizierung der



Stämme festgestellt werden könnte. Dies ist auf den möglichen horizontalen Gentransfer mit anderen Darmbakterien und der Konkurrenz um Ressourcen, sowie aufgrund der dort vorherrschenden höheren Antibiotikakonzentrationen zurückzuführen. [68]–[70] Diese Überlegung würde den in diesem Cluster festgestellten größeren SNP-Abstand zwischen den Stämmen aus dem gleichen Probenmaterial erklären, da diese alle aus Rektalabstrichen stammten. Das Ergebnis ist aber mit Vorsicht zu betrachten, da die Werte für die SNP-Abstände zwischen unterschiedlichen Probenmaterialien nur aus dem Vergleich mit einem Stamm (aus dem Mittelstrahlurin) gewonnen wurden.

Somit konnte trotz signifikanten Ergebnissen in diesem Cluster die Hypothese, dass *P. aeruginosa*-Stämme sich an verschiedenen Stellen des menschlichen Körpers isoliert entwickelt haben könnten und damit einen größeren SNP-Abstand hätten als jene, die im gleichen Probenmaterial nachgewiesen wurden, nicht bestätigt werden. Hierfür könnte ein regelmäßiger Austausch der Stämme, zwischen verschiedenen Körperstellen, die in räumlicher Nähe zueinander lagen, ursächlich gewesen sein. Dies wurde schon für *Escherichia coli*-Stämme zwischen dem Gastrointestinal- und Urethral-Trakt nachgewiesen. [71] Auch ein Austausch aufgrund der gleichen Passage, wie dies beispielsweise bei Rachen- und Rektalabstrichen der Fall gewesen sein könnte, ist denkbar.

Während in anderen Studien nur wenige SNPs zwischen *P. aeruginosa*-Stämmen innerhalb eines Jahres akkumulierten, fielen die hier festgestellten SNP-Abstände zwischen Stämmen aus dem gleichen Probenmaterial in Anbetracht der oft kurzen Zeitspanne allgemein deutlich größer aus. [37], [49], [72] Hierfür kommen mehrere Gründe in Frage. Zum einen ist die Möglichkeit der Koexistenz verschiedener Stämme unterschiedlicher Abstammungslinien innerhalb eines Patienten zu bedenken. [67] Eine Mischinfektion mit verschiedenen *P. aeruginosa*-Stämmen zu Beginn oder im Verlauf einer Infektion anstelle einer Entwicklung durch Evolution innerhalb eines Patienten ist vor allem bei großen SNP-Abständen zweier Stämme aus demselben Probenmaterial nicht auszuschließen. [46], [47], [73] Insbesondere während dieser Mischinfektionen war außerdem die Möglichkeit von horizontalem Gentransfer und die dadurch bedingte genomische Plastizität von *P. aeruginosa*

ein weiterer Faktor, der zur Größe des SNP-Abstands beigetragen haben könnte. [37], [66], [67], [74]–[78] Nicht zuletzt könnte auch das in chronisch infizierten Patienten gehäufte Vorkommen von hypermutierenden Stämmen relevant gewesen sein. [49], [72], [79]–[81] War ein solcher Hypermutator-Stamm vertreten, könnte dieser den unerwartet hohen SNP-Abstand zwischen Stämmen des gleichen Probenmaterials hervorgerufen haben. Allgemein waren die Stämme trotz ihrer räumlichen Nähe einer unterschiedlichen Mikroumgebung ausgesetzt. Diese könnte für einen individuell unterschiedlichen Selektionsdruck gesorgt haben, was die divergente Entwicklung der Stämme bis hin zur jahrelangen Koexistenz von Subpopulationen einer Abstammungslinie an der selben Körperstelle zur Folge gehabt haben könnte. [38], [64], [66], [80]–[83]

Mit großer Wahrscheinlichkeit musste auch aus eben diesen Gründen die Hypothese, dass bei einer größeren Anzahl an unterschiedlichen Probenmaterialien im Cluster, ein größerer medianer SNP-Abstand festzustellen sein müsste, verworfen werden. Da in Stämmen aus unterschiedlichen Probenmaterialien kein größerer SNP-Abstand als zwischen Stämmen aus gleichen Probenmaterialien nachgewiesen werden konnte, war demnach auch kein größerer SNP-Abstand in Clustern mit einer größeren Anzahl an unterschiedlichen Probenmaterialien zu erwarten.

Da die Annahme, dass Stämme mit einem geringen zeitlichen Abstand und großem SNP-Abstand direkt auseinander hervorgingen in Anbetracht der genannten Argumente kritisch zu hinterfragen ist, wurden zur Testung der Hypothese nur Indexstamm-Verbindungen und damit die SNP-Abstände zum vermeintlichen Ursprungsstamm berücksichtigt. Doch auch diese Untersuchung erbrachte aufgrund der schon erwähnten möglichen Gründe keinen signifikanten Unterschied zwischen den SNP-Abständen zweier Stämme aus dem gleichen oder ungleichen Probenmaterial.

Als Schlussfolgerung dieser Untersuchungen ist somit festzuhalten, dass kein Hinweis auf die isolierte Entwicklung eines *P. aeruginosa*-Stammes an einer Körperstelle innerhalb eines Patienten bestand. Dieses Ergebnis steht im Kontrast zu einer früheren Studie, die anhand der Genetik und Phylogenie von

*P. aeruginosa*-Subpopulationen auf deren Verteilung in räumlichen Kompartimenten innerhalb eines Patienten schlussfolgerte. [65] Somit sind zum einen Rückschlüsse vom SNP-Abstand zweier *P. aeruginosa*-Stämme eines Patienten auf deren Abstammungsverhältnis aufgrund der genannten Einflussfaktoren nur mit eingeschränkter Aussagekraft möglich. Zum anderen scheint aufgrund der nachgewiesenen großen SNP-Abstände innerhalb eines Patienten die genetische Mikroevolution von *P. aeruginosa*-Stämme mit einer höheren Geschwindigkeit voranzuschreiten als initial angenommen.

Die Untersuchung unterliegt aber auch diversen Limitationen. Um die Möglichkeit des horizontalen Gentransfers ausschließen zu können, hätte ein Ausschluss von Rekombinationen mittels Gubbins-Analyse [84] durchgeführt werden können. Dies wurde jedoch aufgrund des hier vorliegenden Core-Genoms aus 690 Stämmen und der somit geringen Wahrscheinlichkeit, dass plasmidische Regionen in diesem Core-Genom enthalten waren, unterlassen. Des Weiteren variierten die Art und Anzahl der gleichen und ungleichen Probenmaterialien zwischen den Clustern teils stark, sodass auf die eingeschränkte Möglichkeit der Verallgemeinerung und Vergleichbarkeit der Cluster an dieser Stelle explizit hingewiesen werden muss. Auch der teils sehr kurze Beobachtungszeitraum limitiert die statistische Aussagekraft der Ergebnisse. Zusätzlich ist für die Untersuchung der Indexstamm-Verbindungen anzumerken, dass es sich zum einen bei dem zeitlich erstdetektierten Stamm nicht sicher um den Ursprungsstamm der Abstammungslinie handelte. Zum anderen handelte es sich in der Untersuchung dieser Hypothese um einen clusterübergreifenden Vergleich der SNP-Abstände, da für eine clusterindividuelle Beurteilung die Anzahl an Stämmen zu gering war. Da die Cluster sich in ihrer Zusammensetzung und den medianen SNP-Abständen deutlich unterschieden, ist eine statistische Untersuchung auf Basis dieser Werte nur unter Vorbehalt aussagekräftig.

Allgemein hätten für eine bessere Gegenüberstellung der SNP-Abstände von Stämmen eines Clusters aus gleichen und ungleichen Materialien mehr Probenmaterialien, die zum Rektalabstrich weder eine räumliche Nähe noch die gleiche Passage aufwiesen, routinemäßig und über einen längeren Zeitraum

abgenommen werden müssen. Darüber hinaus lässt die Untersuchung offen, ob die Art des Probenmaterials Einfluss auf die Größe des SNP-Abstand gehabt haben könnte. Weitere Untersuchungen zu Zusammenhängen zwischen der Art des Probenmaterials und dem SNP-Abstand wären hierfür notwendig.

Auch die Hypothese, dass Cluster mit einer größeren Anzahl an Stämmen aufgrund der durch die Messzahl bedingten Heterogenität einen größeren SNP-Abstand aufweisen müssten, wurde in der Korrelationsanalyse nicht bestätigt. Zum einen könnten die schon erwähnten Hypermutator-Stämme, vor allem in kleineren Clustern, den SNP-Abstand größer erscheinen gelassen haben als in den übrigen Clustern. Zum anderen hängt die Evolution von mehreren Faktoren wie der Mutationsrate des Stammes, der Populationsgröße und dem Überlebensvorteil, welcher durch die Mutationen erzielt wurde, ab. [78] Somit ist die Geschwindigkeit und damit die Entwicklung der genetischen Distanz von Cluster zu Cluster unterschiedlich. Dies wird auch durch die große Spannweite der medianen SNP-Abstände der Cluster bekräftigt.

Schlussendlich zeigt diese Korrelationsanalyse nicht den Zusammenhang zwischen der Heterogenität innerhalb eines Clusters und der Clustergröße, sondern weist vielmehr auf den Cluster-individuellen heterogenen Prozess der Evolution verschiedener *P. aeruginosa*-Stämme innerhalb eines Patienten hin. Man muss jedoch als Limitation anmerken, dass hohe SNP-Abstände in einem großen Cluster nicht so ausgeprägt ins Gewicht fielen wie in Clustern mit einer kleineren Anzahl an Stämmen. Eine Untersuchung von Clustern mit der gleichen Anzahl an Stämmen hätte durchgeführt werden müssen.

Weiterhin wurde auch die Hypothese, dass Cluster mit einem größeren maximalen zeitlichen Abstand einen größeren medianen SNP-Abstand aufweisen müssten, verworfen. Die Diversifizierung von *P. aeruginosa*-Stämmen ist ein Prozess, der sehr schnell und in unterschiedlicher Geschwindigkeit voranschreiten kann, sodass auch ein, wie in dieser Studie vorhandener, großer SNP-Abstand in den Clustern mit einem kurzen Beobachtungszeitraum so zu begründen wäre. [85] Dies erklärt auch den

fehlenden Zusammenhang zwischen dem zeitlichen Abstand und der SNP-Größe. Allgemein muss man jedoch anmerken, dass die recht hohen SNP-Abstände nicht mit den in anderen Studien festgestellten niedrigen Mutationsraten vereinbar waren. [37], [49], [72] Demnach könnten auch hier Hypermutator-Stämme oder Stämme anderer Abstammungslinien eine Rolle in Clustern mit einem kleinen maximalen zeitlichen Abstand und gleichzeitig unerwartet hohem medianen SNP-Abstand gespielt haben.

Somit scheint weniger der zeitliche Faktor als vorrangig der Einfluss der individuellen Selektionsdrücke und die Geschwindigkeit der dadurch bedingten evolutionären Prozesse entscheidend für die Entwicklung der SNP-Abstände gewesen zu sein.

Als Limitation der Untersuchung muss aber erwähnt werden, dass die Stämme eines Clusters nicht in regelmäßigen Zeitintervallen nachgewiesen wurden. Dies hatte zur Folge, dass Cluster mit einer größeren Anzahl an zeitlich eng beieinander liegenden Stämmen trotz eines großen maximalen zeitlichen Abstands im Cluster einen unerwartet kleinen medianen SNP-Abstand aufwiesen. Allgemein sind die drei Untersuchungen der Zusammenhänge mit einer Anzahl von 17 Patientenclustern durchgeführt worden. Dies ist eine geringe Fallzahl mit eingeschränkter statistischer Aussagekraft. Für eine höhere statistische Power wären Korrelationsanalysen mit einer größeren Anzahl an Clustern nötig.

In Zusammenschau der Ergebnisse wurde das Ziel der Studie – die Bewertung des Nutzens eines einheitlichen SNP-Grenzwertes zur Unterscheidung zwischen Mutation und Transmission – diskutiert. Als Fazit dieser Studie ließen die unerwartet hohen und vor allem in den Patientenclustern diversen SNP-Abstände zwischen *P. aeruginosa*-Stämmen aber keine Bestimmung eines solchen SNP-Grenzwertes zu. Während einige Studien mit *Mycobacterium tuberculosis* oder *Salmonella enterica* einen Grenzwert des SNP-Abstands verwendeten und die recht konstante natürliche Variabilität dieser Bakteriengattungen eine relativ gute Rückverfolgungsmöglichkeit in Ausbruchsanalysen zuließ, war dies in unserer Studie mit *P. aeruginosa* nicht

der Fall. [55], [86]–[88] Dies ist der großen Diversität der SNP-Abstände, sowohl innerhalb des jeweiligen Clusters als auch zwischen den 17 Clustern, geschuldet. Schon vorangegangene Studien hoben die Diversität von *P. aeruginosa*-Populationen innerhalb eines Individuums hervor, welche sich nun auch in den SNP-Abständen zwischen den Stämmen dieser Studie widerspiegelt. [64]–[67] Dieses Ergebnis steht im Einklang mit einer Mehrzahl an vergangenen Studien, die zeigten, dass die alleinige Berücksichtigung des SNP-Abstands zur Rekonstruktion von Ausbruchsgeschehen zu fehlerhaften Schlussfolgerungen bezüglich des Transmissionsgeschehens führen kann. Grund hierfür ist, wie schon erwähnt, die hohe Variabilität in den Genomen der sequenzierten Stämme innerhalb eines Patienten. [53], [89]–[93] Diese Variabilität kommt zum einen durch die unbekannte Anzahl an evolutionären Ereignissen und Einflussfaktoren, als auch durch Fehler im Sequenzierungs- und Aufbereitungsprozess zustande. [86] Da die hier vorliegende Studie außerdem ein breiteres Patientenkollektiv als vorangegangene Studien einschloss, könnten die patientenindividuellen Unterschiede zusätzlich Einfluss auf die hier nachgewiesenen größeren SNP-Abstände gehabt haben.

Aus diesem Grund ist neben des SNP-Abstands eine Mitberücksichtigung von epidemiologischen Daten von essenzieller Bedeutung. [53], [89], [90], [94] Dies ist jedoch aufgrund der Komplexität der evolutionären Prozesse von *P. aeruginosa* und dem dadurch erschwerten Verständnis der Cluster-individuellen Heterogenität zum aktuellen Zeitpunkt weiterhin nur begrenzt möglich. [53], [65]

Zusammenfassend scheint der Prozess der Evolution und die Diversität von *P. aeruginosa*-Stämmen zu heterogen zu sein, um ihn allein auf den SNP-Abstand reduzieren zu können. Die Aussagekraft eines einheitlichen SNP-Grenzwertes zum Nachweis von Transmissionen ist somit sehr kritisch zu hinterfragen. Auf die Festlegung eines SNP-Grenzwertes wird daher in dieser Arbeit verzichtet. Nichtsdestotrotz liefern die aus NGS gewonnenen Daten des SNP-Abstands der hier vorliegenden Arbeit Hinweise auf eine deutlich rascher voranschreitende Mikroevolution von *P. aeruginosa* innerhalb eines Patienten als anhand vorangegangener Studien anzunehmen war und bestätigen erneut

dessen Fähigkeit, das Gesundheitssystem vor wiederkehrende Herausforderungen bezüglich des Verständnisses seiner Komplexität sowie seiner Bekämpfung und Eradikation im klinischen Setting zu stellen.

Als allgemeine Limitation der Studie bleibt zu erwähnen, dass trotz des initial großen Stamm- und Patientenkollektivs aufgrund der definierten Einschlusskriterien die Untersuchung einer im Vergleich nur kleinen Anzahl an Clustern möglich war. Diese Einschlusskriterien waren jedoch notwendig, um eine aussagekräftige Vergleichbarkeit der SNP-Abstände zu gewährleisten. Ein längerer longitudinaler Beobachtungszeitraum der Patienten und einer daraus resultierenden größeren Anzahl an nachgewiesenen *P. aeruginosa*-Stämmen wären nötig, um die Aussagekraft der Studie zu erhöhen. Da diese Studie den SNP-Abstand als alleiniges Vergleichsinstrument nutzte und dieser in Zusammenschau der Ergebnisse nicht ausreichte, sollte ein Einbezug von Faktoren, die die Entwicklung des SNP-Abstands beeinflussen, Gegenstand zukünftiger Forschungen zur Rekonstruktion von Ausbruchsgeschehen von *P. aeruginosa* sein.

## 5 Zusammenfassung

*Pseudomonas aeruginosa* (*P. aeruginosa*) stellt im klinischen Alltag eine große klinische Herausforderung dar. Ansteigende Morbiditäts- und Mortalitätsraten von intensivpflichtigen und vorerkrankten Patienten sind vor allem seiner Anpassungsfähigkeit sowie zunehmenden Antibiotikaresistenzraten geschuldet. Das Ziel dieser Arbeit war die Erforschung von Verwandtschaftsgraden von nosokomialen *P. aeruginosa*-Stämmen auf Genomebene und die anschließende Bewertung des Nutzens eines sinnvollen einheitlichen Grenzwertes einer Anzahl an sogenannten Einzelnukleotid-Polymorphismen (SNPs, engl. Single Nucleotide Polymorphism) zweier Stämme, welche mit der genetischen Distanz der jeweiligen Stämme korreliert. Der Grenzwert sollte die Wahrscheinlichkeit einer Transmission festlegen können, um nachfolgend eine schnelle Identifikation von Transmissionsquellen und -routen zu ermöglichen. Dabei wurde auf Grundlage der durch Next-Generation Sequencing (NGS) ermittelten Sequenzierungsdaten diese Anzahl der Core-Genom-SNPs als Vergleichsinstrument herangezogen und der Einfluss verschiedener Faktoren auf die Größe des SNP-Abstands untersucht.

Über einen Zeitraum von 54 Monaten wurden dabei 690 *P. aeruginosa*-Stämme von 410 Patienten sequenziert und Core-Genom-SNPs bestimmt. Von diesen wurden nach einer phylogenetischen Analyse 114 Stämme von 17 Patienten in 17 Clustern genauer betrachtet und für nachfolgende statistische Untersuchungen verwendet. Dabei konnte kein größerer SNP-Abstand zwischen Stämmen aus unterschiedlichen Probenmaterialien als jener zweier Stämme aus dem gleichen Probenmaterial nachgewiesen werden. Ferner spielte weder die Clustergröße noch der zeitliche Abstand oder die Anzahl an unterschiedlichen Probenmaterialien eines Clusters eine Rolle für den SNP-Abstand zwischen zwei Isolaten. Stattdessen zeigte sich allgemein eine überraschend große Diversität in den SNP-Abständen innerhalb der Patientencluster.



Schlussfolgernd lässt sich festhalten, dass die Bestimmung eines immer gültigen Grenzwertes des SNP-Abstands zur phylogenetisch sicheren Unterscheidung zweier *P. aeruginosa*-Stämme hinsichtlich Transmissionsereignissen sehr kritisch zu hinterfragen und nur unter Einschränkungen sinnvoll ist, da verschiedenste Faktoren wie die Möglichkeit des Stammaustausches, das Vorkommen von Mischinfektionen, horizontalem Gentransfer und Hypermutator-Stämmen sowie die unterschiedlichen Selektionsdrücke innerhalb eines Patienten Einfluss auf den SNP-Abstand zweier Stämme nehmen. Dies hatte eine große Diversität der Cluster und einen heterogenen, schwer zu beurteilenden Evolutionsprozess der *P. aeruginosa*-Stämme zur Folge. Im Einklang mit vorangegangenen Studien zeigte die hier vorliegende Arbeit, dass der SNP-Abstand als alleiniges Werkzeug somit zur Analyse von Ausbruchsgeschehen nicht geeignet ist. Anhand der hier erhobenen Daten und Ergebnisse wird zusätzlich das im Gegensatz zu vorangegangenen Studien raschere und größtenteils noch unverstandene Voranschreiten der Mikroevolution von *P. aeruginosa* verdeutlicht und inspiriert für zukünftige Studien in diesem Gebiet. Als Ziel zukünftiger Studien wäre ein multifaktorieller Ansatz zur Rekonstruktion von Ausbruchsgeschehen erstrebenswert, der auch epidemiologische Daten und andere Einflussfaktoren des SNP-Abstands miteinbezieht.

## 6 Tabellenverzeichnis

<b>Tab. 1:</b> Anzahl der Probenmaterialien mit Nachweis von <i>P. aeruginosa</i> mit absoluter und relativer Häufigkeit, aufgeschlüsselt nach Zeitraum.....	23
<b>Tab. 2:</b> Übersicht über die den Clustern entsprechenden Patienten-IDs .....	27
<b>Tab. 3:</b> Demographische Daten und Clustergröße der Patientencluster .....	32
<b>Tab. 4:</b> Absolute und relative Häufigkeit der verschiedenen Probenmaterialien der Patientencluster .....	33
<b>Tab. 5:</b> Übersicht über SNP-Abstand, Clustergröße, Anzahl der verschiedenen Probenmaterialien und über den maximalen zeitlichen Abstand der 17 Patientencluster. Die neun P3-Patientencluster sind fett gedruckt, grau hinterlegt und mit einem *,* gekennzeichnet. ....	35
<b>Tab. 6:</b> Übersicht über den SNP-Abstand unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) der 17 Patientencluster. Die neun P3-Patientencluster sind fett gedruckt und mit einem *,* gekennzeichnet und grau hinterlegt.....	38
<b>Tab. 7:</b> Übersicht über den SNP-Abstand (in SNPs) der neun P3-Patientencluster, aufgeschlüsselt nach Cluster sowie gleichem und ungleichem Probenmaterial .....	41
<b>Tab. 8:</b> Werte des Mann-Whitney-U-Tests zum Vergleich des SNP-Abstands bei gleichem und ungleichem Probenmaterial der neun P3-Patientencluster ..	51
<b>Tab. 9:</b> Übersicht über den SNP-Abstand unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) der neun P3-Patientencluster, aufgeschlüsselt nach gleichem und ungleichem Probenmaterial. ....	53
<b>Tab. 10:</b> Werte des Mann-Whitney-U-Tests zum Vergleich des SNP-Abstands bei gleichem und ungleichem Probenmaterial der neun P3-Patientencluster ..	63
<b>Tab. 11:</b> Deskriptive Daten des SNP-Abstands und des SNP-Abstands/Zeit bei gleichem und ungleichem Probenmaterial unter Berücksichtigung nur der Indexstamm-Verbindungen .....	66
<b>Tab. 12:</b> Werte des Mann-Whitney-U-Tests zum Vergleich des SNP-Abstands und des SNP-Abstands/Zeit bei gleichem und ungleichem Probenmaterial unter ausschließlicher Berücksichtigung der Indexstamm-Verbindungen .....	66
<b>Tab. 13:</b> Mutationsraten von <i>P. aeruginosa</i> aus unterschiedlichen Studien im Vergleich .....	<b>Fehler! Textmarke nicht definiert.</b>

## 7 Abbildungsverzeichnis

<b>Abb. 1:</b> Kriterien der ID-Vergabe .....	16
<b>Abb. 2:</b> Phylogenetisches Cluster in FigTree.....	19
<b>Abb. 3:</b> Arbeitsablauf mit Anzahl der eingeschlossenen Patienten und <i>P. aeruginosa</i> -Stämme.....	25
<b>Abb. 4:</b> Altersverteilung der zugrunde liegenden Studienpopulation der 17 Patientencluster, aufgeschlüsselt nach Altersgruppe.....	28
<b>Abb. 5:</b> Altersverteilung der zugrunde liegenden Studienpopulation der neun P3-Patientencluster, aufgeschlüsselt nach Altersgruppe .....	29
<b>Abb. 6:</b> Clustergröße (Anzahl der Stämme pro Cluster) der 17 Patientencluster, aufgeschlüsselt nach Cluster.....	30
<b>Abb. 7:</b> Probenmaterialien der 17 Patientencluster mit absoluter und relativer Häufigkeit. Probenmaterialien, die anteilmäßig < 5 % vertreten sind, sind unter „Andere“ zusammengefasst (Ulcusabstrich, Blutkultur, Abstrich, Leistenabstrich, Wundabstrich, Abstrich Hoden, Abstrich Mundschleimhaut, Zentraler Venenkatheter, Katheterurin, Abstrich Ulcus cruris).....	31
<b>Abb. 8:</b> Probenmaterialien der neun P3-Patientencluster mit absoluter und relativer Häufigkeit. Probenmaterialien, die anteilmäßig < 5 % vertreten sind, sind unter „Andere“ zusammengefasst (Stuhl, Blutkultur, Zentraler Venenkatheter, Katheterurin). .....	32
<b>Abb. 9:</b> Verteilung des mittleren SNP-Abstands (in SNPs) innerhalb der 17 Patientencluster, aufgeschlüsselt nach Cluster. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	36
<b>Abb. 10:</b> Verteilung des mittleren SNP-Abstands unter Berücksichtigung des Zeitfaktors (in SNPs/Jahr) innerhalb der 17 Patientencluster, aufgeschlüsselt nach Cluster. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Die Ausreißer sind zum Zweck der besseren Skalierung nicht vollständig dargestellt. ....	39
<b>Abb. 11:</b> SNP-Abstände von Cluster 1, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	42
<b>Abb. 12:</b> SNP-Abstände von Cluster 2, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	43
<b>Abb. 13:</b> SNP-Abstände von Cluster 4, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	44
<b>Abb. 14:</b> SNP-Abstände von Cluster 6, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	45

<b>Abb. 15:</b> SNP-Abstände von Cluster 7, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	46
<b>Abb. 16:</b> SNP-Abstände von Cluster 8, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	47
<b>Abb. 17:</b> SNP-Abstände von Cluster 10, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	48
<b>Abb. 18:</b> SNP-Abstände von Cluster 11, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	49
<b>Abb. 19:</b> SNP-Abstände von Cluster 14, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands.....	50
<b>Abb. 20:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 1, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands..	54
<b>Abb. 21:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 2, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.....	55
<b>Abb. 22:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 4, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands..	56
<b>Abb. 23:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 6, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands..	57
<b>Abb. 24:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 7, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.....	58
<b>Abb. 25:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 8, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.....	59
<b>Abb. 26:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 10, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands..	60
<b>Abb. 27:</b> SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 11, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt.....	61

- Abb. 28:** SNP-Abstände/Zeit (in SNPs/Jahr) von Cluster 14, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt..... 62
- Abb. 29:** Darstellung des medianen SNP-Abstands in Bezug zur Clustergröße (Anzahl der Stämme in einem Cluster) der 17 Patientencluster. Eine lineare Anpassungslinie ist als zentrale Linie dargestellt. .... 67
- Abb. 30:** Darstellung des medianen SNP-Abstands in Bezug zur Anzahl der verschiedenen Probenmaterialien eines Clusters der 17 Patientencluster. Nicht dargestellt sind Cluster 2 (Patienten-ID 365) und Cluster 17 (Patienten-ID 149). Eine lineare Anpassungslinie ist als zentrale Linie dargestellt. Ausreißer sind nicht dargestellt. .... 68
- Abb. 31:** Darstellung des medianen SNP-Abstands in Bezug zum maximalen zeitlichen Abstand der 17 Patientencluster. Eine lineare Anpassungslinie ist als zentrale Linie dargestellt. Ausreißer sind nicht dargestellt. .... 69
- Abb. 32:** Darstellung des medianen SNP-Abstands in Bezug zur Anzahl der verschiedenen Probenmaterialien eines Clusters der 17 vorläufigen Patientencluster. Nicht dargestellt sind Cluster 2 (Patienten-ID 365) und Cluster 17 (Patienten-ID 149). Eine lineare Anpassungslinie ist als zentrale Linie dargestellt. Ausreißer sind nicht dargestellt..... 70
- Abb. 33:** Darstellung des medianen SNP-Abstands in Bezug zur Anzahl der verschiedenen Probenmaterialien eines Clusters der 17 vorläufigen Patientencluster. Eine lineare Anpassungslinie ist als zentrale Linie dargestellt. .... 71
- Abb. 34:** Darstellung des medianen SNP-Abstands in Bezug zur Anzahl der verschiedenen Probenmaterialien eines Clusters der 17 vorläufigen Patientencluster. Nicht dargestellt sind Cluster 2 (Patienten-ID 365) und Cluster 17 (Patienten-ID 149). Eine lineare Anpassungslinie ist als zentrale Linie dargestellt..... 72
- Abb. 35:** SNP-Abstände (in SNPs) aller 17 Patientencluster zum jeweiligen Indexstamm, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands..... 64
- Abb. 36:** SNP-Abstand/Zeit (in SNPs/Jahr) aller 17 Patientencluster zum jeweiligen Indexstamm, aufgeschlüsselt nach SNP-Abständen zwischen gleichem und ungleichem Probenmaterial. Die Längen der Whisker entsprechen dem 1,5-fachen des Interquartilsabstands. Aus Skalierungsgründen sind nicht alle Ausreißer dargestellt..... 65

## 8 Literaturverzeichnis

- [1] M. C. Sonmezer *et al.*, "Evaluation of Risk Factors for Antibiotic Resistance in Patients with Nosocomial Infections Caused by *Pseudomonas aeruginosa*," *Can. J. Infect. Dis. Med. Microbiol.*, vol. 2016, 2016.
- [2] J. Xu, X. Duan, H. Wu, and Q. Zhou, "Surveillance and correlation of antimicrobial usage and resistance of *Pseudomonas aeruginosa*: A hospital population-based study," *PLoS One*, vol. 8, no. 11, 2013.
- [3] and K. L. Lawrence R. Mulcahy, Vincent M. Isabella, "Pseudomonas aeruginosa biofilms in disease," *Microb. Ecol.*, vol. 68, no. 1, pp. 1–12, 2015.
- [4] P. Ruiz-garbajosa and R. Canton, "Update in Bacteriology Epidemiology of antibiotic resistance in *Pseudomonas aeruginosa* . Implications for empiric and definitive therapy," *Updat. Bacteriol.*, vol. 30, pp. 8–12, 2017.
- [5] A. Lizioli *et al.*, "Prevalence of nosocomial infections in Italy: Result from the Lombardy survey in 2000," *J. Hosp. Infect.*, vol. 54, no. 2, pp. 141–148, 2003.
- [6] J. M. Kim *et al.*, "Multicenter surveillance study for nosocomial infections in major hospitals in Korea," *Am. J. Infect. Control*, vol. 28, no. 6, pp. 454–458, 2000.
- [7] S. Hoang *et al.*, "Risk factors for colonization and infection by *Pseudomonas aeruginosa* in patients hospitalized in intensive care units in France," *PLoS One*, vol. 13, no. 3, pp. 1–12, 2018.
- [8] B. H. Iglewski, "Chapter 27 *Pseudomonas*," *Med. Microbiol. 4th Ed.*, pp. 1–9, 1996.
- [9] R. P. A. Pereira, J. Peplies, D. Mushi, I. Brettar, and M. G. Höfle, "Pseudomonas-Specific NGS assay provides insight into abundance and dynamics of pseudomonas species including *P. aeruginosa* in a cooling tower," *Front. Microbiol.*, vol. 9, no. AUG, pp. 1–15, 2018.
- [10] J. Klockgether, N. Cramer, L. Wiehlmann, C. F. Davenport, and B. Tümmler, "Pseudomonas aeruginosa genomic structure and diversity," *Front. Microbiol.*, vol. 2, no. JULY, pp. 1–18, 2011.
- [11] A. Oliver, X. Mulet, C. López-Causapé, and C. Juan, "The increasing threat of *Pseudomonas aeruginosa* high-risk clones," *Drug Resist. Updat.*, vol. 21–22, pp. 41–59, 2015.
- [12] L. Wiehlmann *et al.*, "Population structure of *Pseudomonas aeruginosa*," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 19, pp. 8101–8106, 2007.

- [13] M. W. Silby, C. Winstanley, S. A. C. Godfrey, S. B. Levy, and R. W. Jackson, "Pseudomonas genomes: Diverse and adaptable," *FEMS Microbiol. Rev.*, vol. 35, no. 4, pp. 652–680, 2011.
- [14] C. K. Stover *et al.*, "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen," *Nature*, vol. 406, no. 6799, pp. 959–964, 2000.
- [15] D. G. Lee *et al.*, "Genomic analysis reveals that *Pseudomonas aeruginosa* virulence is combinatorial," *Genome Biol.*, vol. 7, no. 10, 2006.
- [16] J.-L. Ramos, J. B. Goldberg, and A. Filloux, *Pseudomonas*, vol. 7, no. 9, 2015.
- [17] A. R. Hauser, "So Many Virulence Factors, So Little Time," *Crit Care Med.*, vol. 39, no. 9, pp. 2193–2194, 2012.
- [18] A. R. Hauser, "The type III secretion system of *Pseudomonas aeruginosa*: Infection by injection," *Nat. Rev. Microbiol.*, vol. 7, no. 9, pp. 654–665, 2009.
- [19] J. S. Lam, V. L. Taylor, S. T. Islam, Y. Hao, and D. Kocíncová, "Genetic and functional diversity of *Pseudomonas aeruginosa* lipopolysaccharide," *Front. Microbiol.*, vol. 2, no. JUNE, pp. 1–25, 2011.
- [20] J. A. Driscoll, S. L. Brody, and M. H. Kollef, "The epidemiology, pathogenesis and treatment of *Pseudomonas aeruginosa* infections," *Drugs*, vol. 67, no. 3, pp. 351–368, 2007.
- [21] M. E. Skindersoe *et al.*, "*Pseudomonas aeruginosa* quorum-sensing signal molecules interfere with dendritic cell-induced T-cell proliferation," *FEMS Immunol. Med. Microbiol.*, vol. 55, no. 3, pp. 335–345, 2009.
- [22] O. Ciofu, T. Tolker-Nielsen, P. Ø. Jensen, H. Wang, and N. Høiby, "Antimicrobial resistance, respiratory tract infections and role of biofilms in lung infections in cystic fibrosis patients," *Adv. Drug Deliv. Rev.*, vol. 85, pp. 7–23, 2015.
- [23] M. Muller and T. C. Sorrell, "Production of leukotriene B4 and 5-hydroxyeicosatetraenoic acid by human neutrophils is inhibited by *Pseudomonas aeruginosa* phenazine derivatives," *Infect. Immun.*, vol. 59, no. 9, pp. 3316–3318, 1991.
- [24] G. M. Denning, S. S. Iyer, K. J. Reszka, Y. O'Malley, G. T. Rasmussen, and B. E. Britigan, "Phenazine-1-carboxylic acid, a secondary metabolite of *Pseudomonas aeruginosa*, alters expression of immunomodulatory proteins by human airway epithelial cells," *Am. J. Physiol. - Lung Cell. Mol. Physiol.*, vol. 285, no. 3 29-3, pp. 584–592, 2003.
- [25] M. Rybtke, L. D. Hultqvist, M. Givskov, and T. Tolker-Nielsen, "*Pseudomonas aeruginosa* Biofilm Infections: Community Structure,

- Antimicrobial Tolerance and Immune Response,” *J. Mol. Biol.*, vol. 427, no. 23, pp. 3628–3645, 2015.
- [26] T. F. C. Mah and G. A. O’Toole, “Mechanisms of biofilm resistance to antimicrobial agents,” *Trends Microbiol.*, vol. 9, no. 1, pp. 34–39, 2001.
- [27] G. Cabot *et al.*, “Overexpression of AmpC and efflux pumps in *Pseudomonas aeruginosa* isolates from bloodstream infections: Prevalence and impact on resistance in a Spanish multicenter study,” *Antimicrob. Agents Chemother.*, vol. 55, no. 5, pp. 1906–1911, 2011.
- [28] C. Kiewitz, B. Tu, K. Forschergruppe, and M. H. Hannover, “Sequence Diversity of *Pseudomonas aeruginosa*: Impact on Population Structure and Genome Evolution,” vol. 182, no. 11, pp. 3125–3135, 2000.
- [29] A. L. Corona-Nakamura *et al.*, “Epidemiologic study of *Pseudomonas aeruginosa* in critical patients and reservoirs,” *Arch. Med. Res.*, vol. 32, no. 3, pp. 238–242, 2001.
- [30] A. C. Gales, R. N. Jones, J. Turnidge, R. Rennie, and R. Ramphal, “Characterization of *Pseudomonas aeruginosa* Isolates: Occurrence Rates, Antimicrobial Susceptibility Patterns, and Molecular Typing in the Global SENTRY Antimicrobial Surveillance Program, 1997–1999,” *Clin. Infect. Dis.*, vol. 32, no. s2, pp. S146–S155, 2001.
- [31] G. Yetkin, B. Otlu, A. Cicek, C. Kuzucu, and R. Durmaz, “Clinical, microbiologic, and epidemiologic characteristics of *Pseudomonas aeruginosa* infections in a University Hospital, Malatya, Turkey,” *Am. J. Infect. Control*, vol. 34, no. 4, pp. 188–192, 2006.
- [32] M. Kielhofner, R. L. Atmar, R. J. Hamill, and D. M. Musher, “Life-Threatening *Pseudomonas aeruginosa* Infections in Patients with Human Immunodeficiency Virus Infection,” *Clin. Infect. Dis.*, vol. 14, no. 2, pp. 103–111, 1992.
- [33] R. E. Wood, “*Pseudomonas*: the compromised host.,” *Hosp. Pract.*, vol. 11, no. 8, pp. 91–100, 1976.
- [34] T. Miyoshi-Akiyama *et al.*, “Emergence and Spread of Epidemic Multidrug-Resistant *Pseudomonas aeruginosa*,” *Genome Biol. Evol.*, vol. 9, no. 12, pp. 3238–3245, 2017.
- [35] M. Willmann *et al.*, “Clinical and treatment-related risk factors for nosocomial colonisation with extensively drug-resistant *Pseudomonas aeruginosa* in a haematological patient population: A matched case control study,” *BMC Infect. Dis.*, vol. 14, no. 1, pp. 1–8, 2014.
- [36] E. Johansson, C. Welinder-Olsson, and M. Gilljam, “Genotyping of *Pseudomonas aeruginosa* isolates from lung transplant recipients and aquatic environment-detected in-hospital transmission,” *Apmis*, vol. 122,



no. 2, pp. 85–91, 2014.

- [37] M. Willmann *et al.*, “Analysis of a long-term outbreak of XDR *Pseudomonas aeruginosa*: A molecular epidemiological study,” *J. Antimicrob. Chemother.*, vol. 70, no. 5, pp. 1322–1330, 2014.
- [38] U. Muthukumarasamy *et al.*, “Single-nucleotide polymorphism-based genetic diversity analysis of clinical *Pseudomonas aeruginosa* isolates,” *Genome Biol. Evol.*, vol. 12, no. 4, pp. 396–406, 2020.
- [39] L. Freschi *et al.*, “The *Pseudomonas aeruginosa* Pan-Genome Provides New Insights on Its Population Structure, Horizontal Gene Transfer, and Pathogenicity,” *Genome Biol. Evol.*, vol. 11, no. 1, pp. 109–120, 2019.
- [40] J. Mosquera-Rendón, A. M. Rada-Bravo, S. Cárdenas-Brito, M. Corredor, E. Restrepo-Pineda, and A. Benítez-Páez, “Pangenome-wide and molecular evolution analyses of the *Pseudomonas aeruginosa* species,” *BMC Genomics*, vol. 17, no. 1, pp. 1–14, 2016.
- [41] E. A. Ozer, J. P. Allen, and A. R. Hauser, “Characterization of the core and accessory genomes of *Pseudomonas aeruginosa* using bioinformatic tools Spine and AGEnt,” *BMC Genomics*, vol. 15, no. 1, pp. 1–17, 2014.
- [42] D. H. Spencer *et al.*, “Whole-genome sequence variation among multiple isolates of *Pseudomonas aeruginosa*,” *J. Bacteriol.*, vol. 185, no. 4, pp. 1316–1325, 2003.
- [43] A. C. Schürch, S. Arredondo-Alonso, R. J. L. Willems, and R. V. Goering, “Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene-based approaches,” *Clin. Microbiol. Infect.*, vol. 24, no. 4, pp. 350–354, 2018.
- [44] J. R. Dettman, J. L. Sztepanacz, and R. Kassen, “The properties of spontaneous mutations in the opportunistic pathogen *Pseudomonas aeruginosa*,” *BMC Genomics*, vol. 17, no. 1, pp. 1–14, 2016.
- [45] S. Kumar, T. W. Banks, and S. Cloutier, “SNP discovery through next-generation sequencing and its applications,” *Int. J. Plant Genomics*, vol. 2012, 2012.
- [46] Bethesda, “SNP FAQ Archive,” *National Center for Biotechnology Information (US)*, 2014. [Online]. Available: [https://www.ncbi.nlm.nih.gov/books/NBK3848/pdf/Bookshelf\\_NBK3848.pdf](https://www.ncbi.nlm.nih.gov/books/NBK3848/pdf/Bookshelf_NBK3848.pdf). [Accessed: 16-Jul-2020].
- [47] J. Stimson, J. L. Gardy, B. Mathema, V. Crudu, T. Cohen, and C. Colijn, “Beyond the SNP threshold: identifying outbreak clusters using inferred transmissions,” *bioRxiv*, p. 319707, 2018.
- [48] S. R. Harris *et al.*, “Evolution of MRSA during hospital transmission and

- intercontinental spread," *Science* (80-. ), vol. 327, no. 5964, pp. 469–474, 2010.
- [49] N. Cramer, L. Wiehlmann, and B. Tümmler, "Clonal epidemiology of *Pseudomonas aeruginosa* in cystic fibrosis," *Int. J. Med. Microbiol.*, vol. 300, no. 8, pp. 526–533, 2010.
- [50] B. E. Slatko, A. F. Gardner, and F. M. Ausubel, "Overview of Next Generation Sequencing technologies (and bioinformatics) in cancer," *Mol. Biol.*, vol. 122, no. 1, pp. 1–15, 2018.
- [51] D. Qin, "Next-generation sequencing and its clinical application," *Cancer Biol. Med.*, vol. 16, no. 1, pp. 4–10, 2019.
- [52] A. J. Sabat *et al.*, "Overview of molecular typing methods for outbreak detection and epidemiological surveillance.," *Eur. Commun. Dis. Bull.*, vol. 18, no. 4, p. 20380, 2013.
- [53] C. J. Worby, H. H. Chang, W. P. Hanage, and M. Lipsitch, "The distribution of pairwise genetic distances: A tool for investigating disease transmission," *Genetics*, vol. 198, no. 4, pp. 1395–1404, 2014.
- [54] A. J. Taylor *et al.*, "Characterization of foodborne outbreaks of salmonella enterica serovar enteritidis with whole-genome sequencing single nucleotide polymorphism-based analysis for surveillance and outbreak detection," *J. Clin. Microbiol.*, vol. 53, no. 10, pp. 3334–3340, 2015.
- [55] M. Payne *et al.*, "Enhancing genomics-based outbreak detection of endemic *Salmonella enterica* serovar Typhimurium using dynamic thresholds," *Microb. Genomics*, 2019.
- [56] A. J. Page *et al.*, "Roary: Rapid large-scale prokaryote pan genome analysis," *Bioinformatics*, vol. 31, no. 22, pp. 3691–3693, 2015.
- [57] A. Tritt, J. A. Eisen, M. T. Facciotti, and A. E. Darling, "An Integrated Pipeline for de Novo Assembly of Microbial Genomes," *PLoS One*, vol. 7, no. 9, 2012.
- [58] A. Bankevich *et al.*, "SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing," *J. Comput. Biol.*, vol. 19, no. 5, pp. 455–477, 2012.
- [59] T. Seemann, "Prokka: Rapid prokaryotic genome annotation," *Bioinformatics*, vol. 30, no. 14, pp. 2068–2069, 2014.
- [60] C. K. Stover *et al.*, "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen," *Nature*, vol. 406, no. 6799, pp. 959–964, 2000.
- [61] D. S. Sarovich and E. P. Price, "SPANDx: A genomics pipeline for comparative analysis of large haploid whole genome re-sequencing datasets," *BMC Res. Notes*, vol. 7, no. 1, pp. 1–9, 2014.

- [62] A. Stamatakis, "RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies," *Bioinformatics*, vol. 30, no. 9, pp. 1312–1313, 2014.
- [63] A. Rambaut, "Figtree," 2009. [Online]. Available: <https://github.com/rambaut/figtree/releases/tag/v1.4.3>. [Accessed: 16-Jul-2020].
- [64] S. Feliziani *et al.*, "Coexistence and Within-Host Evolution of Diversified Lineages of Hypermutable *Pseudomonas aeruginosa* in Long-term Cystic Fibrosis Infections," *PLoS Genet.*, vol. 10, no. 10, 2014.
- [65] T. Markussen *et al.*, "Environmental heterogeneity drives within-host diversification and evolution of *Pseudomonas aeruginosa*," *MBio*, vol. 5, no. 5, pp. 1–10, 2014.
- [66] J. C. S. Chung *et al.*, "Genomic variation among contemporary *Pseudomonas aeruginosa* isolates from chronically infected cystic fibrosis patients," *J. Bacteriol.*, vol. 194, no. 18, pp. 4857–4866, 2012.
- [67] D. Williams *et al.*, "Divergent, coexisting *Pseudomonas aeruginosa* lineages in chronic cystic fibrosis lung infections," *Am. J. Respir. Crit. Care Med.*, vol. 191, no. 7, pp. 775–785, 2015.
- [68] G. Cabot *et al.*, "Evolution of *Pseudomonas aeruginosa* antimicrobial resistance and fitness under low and high mutation supply rates," *Antimicrob. Agents Chemother.*, vol. 60, no. January, p. AAC.02676-15, 2016.
- [69] R. S. Ramiro, A. Sousa, N. Fraza, and I. Gordo, "ScienceDirect Evolution of commensal bacteria in the intestinal tract of mice," pp. 114–121, 2017.
- [70] P. D. Scanlan, "Microbial evolution and ecological opportunity in the gut environment," *Proc. R. Soc. B Biol. Sci.*, vol. 286, no. 1915, pp. 1–9, 2019.
- [71] S. Yamamoto, T. Tsukamoto, A. Terai, H. Kurazono, and Y. Takeda, "GENETIC EVIDENCE SUPPORTING THE FECAL-PERINEAL-URETHRAL HYPOTHESIS IN CYSTITIS CAUSED BY ESCHERICHIA COLI," pp. 1127–1129, 1997.
- [72] R. L. Marvig, H. K. Johansen, S. Molin, and L. Jelsbak, "Genome Analysis of a Transmissible Lineage of *Pseudomonas aeruginosa* Reveals Pathoadaptive Mutations and Distinct Evolutionary Paths of Hypermutators," *PLoS Genet.*, vol. 9, no. 9, 2013.
- [73] C. Chaguza *et al.*, "Within-host microevolution of *Streptococcus pneumoniae* is rapid and adaptive during natural colonisation," *Nat. Commun.*, vol. 11, no. 1, pp. 1–14, 2020.
- [74] K. Mathee *et al.*, "Dynamics of *Pseudomonas aeruginosa* genome

- evolution,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 105, no. 8, pp. 3100–3105, 2008.
- [75] L. Cullen and S. McClean, “Bacterial adaptation during chronic respiratory infections,” *Pathogens*, vol. 4, no. 1, pp. 66–89, 2015.
- [76] S. E. Darch *et al.*, “Recombination is a key driver of genomic and phenotypic diversity in a *Pseudomonas aeruginosa* population during cystic fibrosis infection,” *Sci. Rep.*, vol. 5, pp. 1–12, 2015.
- [77] J. R. Dettman, N. Rodrigue, S. D. Aaron, and R. Kassen, “Evolutionary genomics of epidemic and nonepidemic strains of *Pseudomonas aeruginosa*,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 110, no. 52, pp. 21065–21070, 2013.
- [78] X. Didelot, A. S. Walker, T. E. Peto, D. W. Crook, and D. J. Wilson, “Within-host evolution of bacterial pathogens,” *Nat. Rev. Microbiol.*, vol. 14, no. 3, pp. 150–162, 2016.
- [79] A. Mena *et al.*, “Genetic adaptation of *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients is catalyzed by hypermutation,” *J. Bacteriol.*, vol. 190, no. 24, pp. 7910–7917, 2008.
- [80] C. Winstanley, S. O’Brien, and M. A. Brockhurst, “*Pseudomonas aeruginosa* Evolutionary Adaptation and Diversification in Cystic Fibrosis Chronic Lung Infections,” *Trends Microbiol.*, vol. 24, no. 5, pp. 327–337, 2016.
- [81] L. Yang *et al.*, “Evolutionary dynamics of bacteria in a human host environment,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 108, no. 18, pp. 7481–7486, 2011.
- [82] C. N. Wilder, G. Allada, and M. Schuster, “Instantaneous within-patient diversity of *Pseudomonas aeruginosa* quorum-sensing populations from cystic fibrosis lung infections,” *Infect. Immun.*, vol. 77, no. 12, pp. 5631–5639, 2009.
- [83] J. D. Caballero *et al.*, “Selective sweeps and parallel pathoadaptation drive *pseudomonas aeruginosa* evolution in the cystic fibrosis lung,” *MBio*, vol. 6, no. 5, pp. 1–14, 2015.
- [84] N. J. Croucher *et al.*, “Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins,” *Nucleic Acids Res.*, vol. 43, no. 3, p. e15, 2015.
- [85] S. T. Clark *et al.*, “Phenotypic diversity within a *Pseudomonas aeruginosa* population infecting an adult with cystic fibrosis,” *Sci. Rep.*, vol. 5, no. February, pp. 1–10, 2015.
- [86] M. W. Allard *et al.*, “High resolution clustering of *Salmonella enterica* serovar Montevideo strains using a next-generation sequencing

approach,” *BMC Genomics*, vol. 13, no. 1, 2012.

- [87] T. M. Walker *et al.*, “Assessment of *Mycobacterium tuberculosis* transmission in Oxfordshire, UK, 2007-12, with whole pathogen genome sequences: An observational study,” *Lancet Respir. Med.*, vol. 2, no. 4, pp. 285–292, 2014.
- [88] P. E. M. Fine *et al.*, “Large-scale whole genome sequencing of *M. tuberculosis* provides insights into transmission in a high prevalence area,” *Elife*, vol. 2015, no. 4, pp. 1–17, 2015.
- [89] A. W. Pightling, J. B. Pettengill, Y. Luo, J. D. Baugher, H. Rand, and E. Strain, “Interpreting whole-genome sequence analyses of foodborne bacteria for regulatory applications and outbreak investigations,” *Front. Microbiol.*, vol. 9, no. JUL, pp. 1–13, 2018.
- [90] C. J. Worby, M. Lipsitch, and W. P. Hanage, “Within-Host Bacterial Diversity Hinders Accurate Reconstruction of Transmission Networks from Genomic Distance Data,” *PLoS Comput. Biol.*, vol. 10, no. 3, 2014.
- [91] J. Bryant, C. Chewapreecha, and S. D. Bentley, “Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences,” *Future Microbiol.*, vol. 7, no. 11, pp. 1283–1296, 2012.
- [92] S. D. Bentley and J. Parkhill, “Genomic perspectives on the evolution and spread of bacterial pathogens,” *Proc. R. Soc. B Biol. Sci.*, vol. 282, no. 1821, 2015.
- [93] H. A. Hatherell, C. Colijn, H. R. Stagg, C. Jackson, J. R. Winter, and I. Abubakar, “Interpreting whole genome sequencing for investigating tuberculosis transmission: A systematic review,” *BMC Med.*, vol. 14, no. 1, pp. 1–13, 2016.
- [94] S. Wesley Long, S. B. Beres, R. J. Olsen, and J. M. Musser, “Absence of Patient-to-Patient intrahospital transmission of *staphylococcus aureus* as determined by Whole-Genome sequencing,” *MBio*, vol. 5, no. 5, pp. 1–10, 2014.

## 9 Erklärung zum Eigenanteil

Die Durchführung der Arbeit fand am Institut für Medizinische Mikrobiologie und Hygiene, Elfriede-Aulhorn-Str. 6, 72076 Tübingen, unter der Leitung und Konzeption von Prof. Dr. med. Matthias Willmann, statt.

Die Akquisition der *P. aeruginosa*-Stämme und Sequenzierungsdaten erfolgte durch die Forschungsgruppe von Prof. Dr. med. Matthias Willmann und eine externe Auftragsfirma (CeMeT, Tübingen. Germany).

Die weitere Aufarbeitung und Auswertung der Daten sowie die statistischen Analysen erfolgten nach Beratung mit Herrn Prof. Dr. med. Matthias Willmann eigenständig durch mich.

Ich versichere hiermit, das gesamte Manuskript selbst verfasst zu haben und keine weiteren Quellen, als die hier angegebenen, verwendet zu haben.

Kornwestheim, den 16.09.2021



Madeleine Desch

## 10 Danksagung

Allen voran danke ich Frau Prof. Dr. med. Julia-Stefanie Frick für die Möglichkeit, meine Dissertation am Institut für Medizinische Mikrobiologie und Hygiene der Universität Tübingen durchführen zu dürfen.

Darüber hinaus gebührt ein großes Dankeschön meinem Doktorvater und Betreuer, Herrn Prof. Dr. med. Matthias Willmann, der mir das Verfassen meiner Dissertation unter seiner Betreuung über das Thema der Umwelt- und Transmissionsfaktorenanalyse von nosokomialen *P. aeruginosa*-Stämmen ermöglichte. Betonen möchte ich an dieser Stelle die außerordentlich gute Betreuung, die sich durch den angenehmen Kontakt, die schnelle Erreichbarkeit und regelmäßigen Treffen und die Hilfestellungen und Anregungen bei Problemen und Fragen auszeichnete. Insbesondere für die Geduld, die mir beim Verfassen der Arbeit entgegengebracht wurde, bin ich sehr dankbar.

Des Weiteren danke ich allen Personen, die an der Konzeption und Durchführung der Studie beteiligt waren und mir dadurch erst die Grundlage für meine statistischen Auswertungen geschaffen haben und bedaure gleichzeitig zutiefst, dass sie hier nicht alle namentlich erwähnt werden können.

Ein großes Dankeschön gebührt darüber hinaus meinen Kommilitonen und Freunden Dominik Goldschmid und Lena Kugler, die mir mit Rat und Tat zur Seite standen und mich sowohl mit inhaltlichen Korrektur- und Lösungsvorschlägen als auch mit ihrer Motivation stets unterstützt haben.

Ein besonderer Dank gilt außerdem meinen Eltern Christine und Amand Desch, die mir in jeder Lebenslage bedingungslos zur Seite stehen.