

Das diesem Bericht zugrundeliegende Vorhaben wurde mit Mitteln des Bundesministeriums für Bildung, und Forschung unter dem Förderkennzeichen **01KI2128** gefördert. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt beim Autor.

Kurzbericht

Übertragungswege auf karibischen Inseln: Eine Analyse im Rahmen des One Health Ansatzes (ACRAS-R) - Antibiotikaresistenzgenome und Resistenzgene

Förderkennzeichen: 01KI2128

Zuwendungsempfängerin: Technische Universität Dresden

Vorhabenlaufzeit: 01.02.2021 – 30.04.2024

Einer der wichtigsten Fortschritte der modernen Gesellschaft ist die Entdeckung und industrielle Herstellung von Antibiotika zur Behandlung von ansonsten potenziell tödlichen bakteriellen Infektionen. Die Popularität solcher Medikamente in Verbindung mit den niedrigeren Produktionskosten und dem Wirtschaftswachstum selbst führte zu einem massiven Einsatz von Antibiotika, was letztlich die Selektion antimikrobiell resistenter Bakterienstämme begünstigte.

Da es immer weniger wirksame antimikrobielle Therapien gibt, werden alternative Methoden (d. h. aktive Überwachung und Aufklärung der Verbreitungswege von Antibiotikaresistenzen) immer wichtiger, um bakterielle Infektionen zu verhindern. In den letzten Jahren hat sich gezeigt, dass anthropogene Aktivitäten eng mit der Ausbreitung antimikrobiell resistenter Mikroorganismen nicht nur beim Menschen, sondern auch bei Tieren und in der Umwelt zusammenhängen. Trotzdem gibt es mehrere Wissenslücken, die umfassend untersucht werden müssen, um das Risiko im Zusammenhang mit der Antibiotikaresistenz (AMR) richtig einschätzen zu können, einschließlich der Rolle der Umweltbedingungen (Exposom, Matrixdiversität und deren intrinsische Mikrobiota).

In den letzten Jahren wurde immer deutlicher, dass der internationale Reiseverkehr eine Schlüsselrolle bei der Verbreitung von Krankheitserregern beim Menschen spielt, einschließlich antimikrobiell resistenter Bakterien.

Im Rahmen des ACRAS-R-Projekts haben wir folglich modernste genomische Analysen durchgeführt, um den Einfluss von Tourismus auf die Verbreitung von Antibiotikaresistenzen gezielt zu entschlüsseln. Untersuchungsort war die karibische Insel Guadeloupe, die sich durch die Insellage, geringe Fläche und Bevölkerungszahl sowie genauer statistischer Angaben zum Tourismus besonders für diese Fragestellung eignet. Die einzelnen Ziele des Projekts lauteten:

- Charakterisierung der AMR-Übertragungswege in einem Kontinuum von klinischen Einrichtungen bis hin zu einer natürlichen Aufnahmeumgebung.
- die Rolle des Tourismus bei der Verbreitung klinisch relevanter Bakterien in neuen Umgebungen vollständig zu charakterisieren.
- Analyse der Rolle des Tourismus als potenzielle Quelle neuartiger antimikrobieller Resistenzmutationen oder -gene.

Die Experimente lieferten wertvolle Einblicke in die Dynamik antimikrobiell resistenter bakterieller Erreger entlang von Kontinua und auch in deren Ausbreitung im Zusammenhang mit touristischen Aktivitäten. Wir stellten fest, dass Krankenhäuser und insbesondere deren Abwässer als Selektionsgründe für Krankheitserreger dienen, die mehrere antimikrobielle Resistenzgene (ARG) tragen, von denen ein Teil auch nach der Abwasserbehandlung in die aufnehmenden Gewässer gelangt, was bis zu einem gewissen Grad zum Vorhandensein von ARGs in vom Menschen beeinflussten Umgebungen beiträgt. Besonders bemerkenswert ist die Beobachtung, dass *E. coli* mit einer deutlich höheren ARG-Belastung durch touristische Aktivitäten in die einheimische Umwelt eingeschleppt werden könnten. Darüber hinaus wiesen virulente Pathotypen aus Proben, die mit touristischen Aktivitäten in Verbindung gebracht wurden, eine vergleichbare Häufigkeit auf wie diejenigen, die in Krankenhausproben gefunden wurden, was darauf hindeutet, dass Touristenströme ebenso zur Verbreitung besonders pathogener Isolate beitragen könnten wie infizierte Patienten aus Krankenhäusern. Interessanterweise stellten wir fest, dass der Tourismus für die Einschleppung von *E. coli*-Klonen verantwortlich war, die als „Hochrisiko“-Klone bekannt sind, und dass Proben aus dem

Tourismus eine größere klonale Ähnlichkeit mit Proben aus dem Krankenhaus aufwiesen. Schließlich stellten wir fest, dass die Größe des Pangenoms beider touristischer Proben signifikant höher war als die der nicht-touristischen Proben, was auf eine immense genetische Vielfalt der Isolate hinweist, die in Abhängigkeit vom Tourismus zirkulieren, was in direktem Zusammenhang mit einem größeren Mobilom und einem höheren Potenzial für den Erwerb von Genen und die Entstehung neuer Gene steht, zu denen auch ARGs gehören könnten.

Die GWAS-Analysen, bei denen der Tourismus berücksichtigt wurde, deckten Tausende signifikanter Genomvarianten auf, darunter bekannte Mutationen sowie neue Varianten, die möglicherweise mit der Resistenz gegen 14 Arten von Antibiotika in Verbindung stehen. Diese Ergebnisse verdeutlichen nicht nur die potenzielle Rolle des Tourismus bei der Verbreitung neuartiger ARGs, sondern liefern auch eine Liste potenziell neuartiger Mechanismen der antimikrobiellen Resistenz.

Die Ergebnisse des ACRAS-R-Projekts sind von großer Bedeutung, da sie einen soliden Beitrag zur Schließung der zuvor erwähnten Wissenslücken leisten. Insbesondere haben wir festgestellt, dass touristische Aktivitäten in der Tat eine wichtige Rolle bei der Ausbreitung relevanter Krankheitserreger spielen, mehr noch, wir kommen zu dem Schluss, dass das Risiko nicht nur mit der Verbreitung von AMR zusammenhängt, sondern eine vielschichtige Bedrohung darstellt. Das bedeutet, dass der Tourismus in Zukunft als ein wichtiger Faktor bei Risikobewertungsverfahren berücksichtigt werden muss. Die Anwendbarkeit der GWAS-Ergebnisse geht über das ACRAS-R-Projekt hinaus. Die vorhergesagten ARGs können nach experimenteller Bestätigung in ARG-Datenbanken zur Früherkennung neu auftretender Resistenzen gegen eine Vielzahl von Antibiotika eingesetzt werden.

Darüber hinaus konnten wir dank des experimentellen Designs des Projekts und der angewandten Methoden die Beschränkungen der bisherigen Literatur (geringe Anzahl von Proben, Analyse einer begrenzten Anzahl von ARGs oder verzerrte phänotypische Resistenzbewertung) überwinden, was die Ergebnisse für die wissenschaftliche Gemeinschaft sehr wertvoll macht.

Eingehende Darstellung

Übertragungswege auf karibischen Inseln: Eine Analyse im Rahmen des One Health Ansatzes (ACRAS-R) - Antibiotikaresistenzgenome und Resistenzgene

Förderkennzeichen: 01KI2128

Zuwendungsempfängerin: Technische Universität Dresden

Vorhabenlaufzeit: 01.02.2021 – 30.04.2024

Zusammenfassung

Antimikrobielle Resistenzen (AMR) werden von der Weltgesundheitsorganisation als ein wichtiges Gesundheitsproblem anerkannt. Die Epidemiologie der AMR ist komplex. Sie muss ganzheitlich betrachtet werden und Mensch, Tier sowie Umwelt berücksichtigen. Antibiotikaresistente Bakterien und Antibiotikaresistenzgene, die aus vom Menschen beeinflussten Quellen freigesetzt werden sowie der übermäßige bzw. missbräuchliche Einsatz von Antibiotika im menschlichen und veterinärmedizinischen Umfeld tragen entscheidend zur Problematik bei. Unklar sind die Verbreitungswege der Resistenzen und die Auswirkungen der Globalisierung, insbesondere von Reisen, auf die Verbreitung. Im ACRAS-R Projekt wurde ist eine integrierte Risikobewertung der Verbreitung von AMR in der Umwelt durch Tourismus durchgeführt. Untersuchungsort war die karibische Insel Guadeloupe, die sich durch die Insellage, geringe Fläche und Bevölkerungszahl sowie genauer statistischer Angaben zum Tourismus besonders für diese Fragestellung eignet.

Zu diesem Zweck wurden AMR-Übertragungswege vom Menschen hin zur Umwelt in touristisch beeinflussten und Gegenden mit geringem Tourismus untersucht (insbesondere die Genome resistenter *Escherichia coli*). Hierzu analysierten wir Proben von Flugzeugtoiletten, Kliniken, aus der Umwelt, dem Abwasser und aus Fäkalien von Wild- und Nutztieren. Erstmals wurden eine komplexe bioinformatische Analyse durchgeführt um zu erörtern, wie die Ausbreitungsdynamiken von resistenten *E.coli* in den verschiedenen AMR-Reservoirs im Rahmen der „One Health“-Perspektive verknüpft sind und wie der Tourismus sich auf diese Ausbreitung auswirkt. Zudem wurden neue bioinformatische Tools für die Analyse von Genomen und Vorhersage von Resistenzen entwickelt. Die Ergebnisse sollen künftig als Entscheidungshilfe und zur Entwicklung von Leitlinien dienen, um die Relevanz von Umweltreservoirs hinsichtlich der Entstehung, Übertragung und Verbreitung von AMR zu bewerten.

Das Forschungsvorhaben trägt somit im Sinne der Bekanntmachung dazu bei, antimikrobielle Resistenzen durch eine enge Zusammenarbeit verschiedener Disziplinen im Sinne des One-Health-Ansatzes zu erforschen, Kompetenzen über Grenzen hinweg zu bündeln und damit Fortschritte bei der Bekämpfung von und dem Umgang mit antibiotikaresistenten Infektionen zu erzielen.

Ziel der Arbeit

Ziel des ACRAS-R-Projekts war es, die Auswirkungen des Tourismus auf die Verbreitung klinisch relevanter Krankheitserreger mit besonderem Schwerpunkt auf antimikrobiellen Resistenzgenen (ARG) zu analysieren. Zu diesem Zweck wurden auf der Insel Guadeloupe drei Kontinua definiert (Krankenhaus, Tourismus und Nicht-Tourismus), von denen an verschiedenen Stellen Proben entnommen wurden (u. a. Zu- und Abfluss von Kläranlagen, stationäre Patienten, Vorfluter und in der Nähe von Kläranlagen lebende Tiere); zusätzlich wurden Toilettenproben von Flugzeugen entnommen, die auf der Insel landeten. Insgesamt 1437 *E. coli* wurden vom lokalen Team mit chromogenen Medien isoliert, die Identität wurde mit biochemischen Tests überprüft und die Empfindlichkeitsmuster wurden mit Hilfe der Scheibendiffusionsmethode bestimmt. Die DNA jedes Isolats wurde mit dem PoweSoil Pro DNA-Extraktionskit (Qiagen) extrahiert und an das Institut für Hydrobiologie der TU-Dresden gesandt.

Methoden bezüglich der Verbreitung von E. coli, Resistenzen, Virulenz und Sequenztypen

DNA-Sequenzierung

Die DNA wurde am Dresden Concept Genome Center der TU Dresden sequenziert. Ein Illumina NovaSeq 6000 Sequenzer wurde verwendet, um 100 bp Paired End Reads zu generieren, wobei eine durchschnittliche 150-fache Abdeckung erreicht wurde.

Verarbeitung und Assemblierung der Rohsequenzen.

Die Qualität der Rohsequenzen wurde mit FastQC V 0.12.0 bewertet. Danach wurden die Reads mit Trimmomatic V0.39 beschnitten, um Adapter zu entfernen und qualitativ hochwertige Reads (Q>30) zu erhalten. Die Taxonomie wurde mit Kraken V2.1.3 und den getrimmten Reads als Input überprüft. Die Reads wurden mit Unicycler V0.4.8 im normalen Modus und unter Verwendung der automatischen Methode zur optimalen Auswahl der K-Mer-Länge assembliert. Schließlich wurden mehrere Runden des Pilon-basierten Polishing durchgeführt, um die Assemblierung zu optimieren.

Genvorhersage

Zum Nachweis interessanter Gene, d. h. erworbener Resistenzgene, Virulenzfaktoren und Replikons, wurden die assemblierten Genome mit der ResFinder-Datenbank V 4.5.0, der E. coli-spezifischen Virulenzfaktor-Datenbank V 0.1 und der PlasmidFinder-Datenbank Release 04/12/2023 unter Verwendung von BLAST+ V2.13 und einem Identitäts- und Abdeckungs-Cutoff von 90 % abgeglichen. Die Ergebnisse jedes Isolats wurden mit Hilfe von R zu einer einzigen Präsenz/Absenz-Matrix zusammengeführt.

Zuordnung der Sequenztypen

Die Sequenztypen wurden jedem einzelnen Isolat nach dem von Achtman vorgeschlagenen MLST-Schema unter Verwendung des mlst-Tools V2.23 zugewiesen. Exakte Übereinstimmungen wurden mit einem definierten ST versehen, während einzelne Locus-Varianten oder völlig neue STs als nicht bestimmt (ND) behandelt wurden. Isolate mit einer definierten ST wurden außerdem mit Hilfe des PubMLST-Webservers in die entsprechenden klonalen Komplexe eingeordnet.

Genomische Annotation und phylogenetische Beziehung

Die Genome wurden mit Prokka V1.14.5 annotiert, wobei nur Contigs von 200 bp oder mehr berücksichtigt wurden. Die erzeugten gff3-Dateien wurden als Eingabe für Roary V3.13 verwendet, um nach den Genen zu suchen, die in mindestens 98 % der Isolate vorhanden sind (Core), und nach denen, die weniger häufig vorkommen (Pan-Genom).

Statistische Datenanalyse

Die statistische Auswertung erfolgte in R, die Ordnungsanalyse wurde mit dem Vegan-Paket V 2.6-4 durchgeführt und die Diagramme wurden mit ggplot2 erstellt.

Methoden bezüglich Genome Wide Association Study und Phenotypischer Susceptibility

Analyse der Metadaten

Um die Variationen in der Probenanfälligkeit zu bewerten, verwendeten wir Violinplots, die mit ggplot2 v3.4.0 in R erstellt wurden, um die Verteilung der Resistenzwerte über verschiedene Proben hinweg zu visualisieren. Zu diesen Diagrammen wurden horizontale Linien hinzugefügt, um die oberen, unteren und mittleren EUCAST-Resistenzschwellenwerte für den Vergleich anzuzeigen.

Sequenzverarbeitung und Qualitätskontrolle

Die rohen Paired-End-Lese-dateien wurden zunächst mit FastQC (v0.11.9) und MultiQC (v1.14) auf die Lesequalität der Proben analysiert. Die Daten wurden dann mit Trimmomatic (v0.39) kuratiert, wobei Sequenzen, die den Qualitätsschwellenwerten nicht entsprachen, beschnitten oder entfernt und Verunreinigungen wie Adapter eliminiert wurden. Die Kontamination durch Nextera-Adapter, die nur in Proben aus Kampagne 3 festgestellt wurde, wurde speziell für diese Kohorte behandelt. Außerdem wurden die Daten aus den Kampagnen 1 und 2 gemeinsam verarbeitet. Nach der Sequenzkuratierung und der Entfernung von Adaptern wurden FastQC und MultiQC erneut verwendet, um die Effizienz des Qualitätskontrollschritts zu bewerten.

Mapping der Sequenzen auf die Referenzsequenz

Nach der Kuratierung der Probanden wurden die Sequenzen an das Referenzgenom von E. coli K12 MG1655 angeglichen. Das Referenzgenom wurde mit Samtools zur Erstellung einer Indexdatei und mit Picard Tools zur Erstellung eines Sequenzwörterbuchs vorbereitet. Das Alignment der Proben-Reads wurde mit dem BWA-MEM-Algorithmus durchgeführt, der SAM-Dateien für das anschließende Varianten-Calling generierte.

Vorbereitung der Proben für die Variantenbestimmung

Die SAM-Dateien wurden auf der Grundlage der in der FastA-Referenzdatei enthaltenen Koordinateninformationen sortiert. Dies geschah mit Hilfe der SortSam-Funktion des Picard-Tools. Die Zuweisung von Lesegruppen erfolgte wiederum mit Hilfe der Funktion AddOrReplaceReadGroups von Picard-Tools. Dies ermöglichte eine Unterscheidung zwischen den Probensätzen. Doppelte Reads, die entweder durch Kontamination oder durch die Sequenzierungsmethode entstanden waren, wurden markiert und mit dem Picard-Tool MarkDuplicates entfernt. Dadurch wurden Duplikate aus der nachgelagerten Analyse entfernt, und als Ergebnis wurden binäre SAM-Dateien (BAM) erzeugt. Die BAM-Dateien wurden dann mit dem QualiMap (v2.2.1) bamqc-Tool analysiert, um die Qualität der Ausrichtung der Reads auf das Referenzgenom MG1655 zu bestimmen.

Variantenbestimmung

Genomische Positionen wurden mit GATKs HaplotypeCaller identifiziert, um gVCFs zu erzeugen. Diese wurden mit GATK's GenomicsDBImport zu einer einzigen VCF-Datei zusammengeführt und anschließend mit GATK's GenotypeGVCFs genotypisiert. Für die harte Filterung wurden die VCF-Dateien mit GATK's SelectVariants nach Variantentyp (SNP oder INDEL) unterteilt und mit VariantFiltration qualitätskontrolliert. Varianten mit niedriger Genotypqualität ($GQ < 20$) und hoher Missingness ($>15\%$) wurden herausgefiltert. Multiallelische Stellen wurden mit BCFTools (v1.18) aufgeteilt, um die Analyse einzelner Allele zu ermöglichen. Schließlich wurden die Varianten mit SnpEff (v5.2) funktionell annotiert, um ihre Auswirkungen auf Aminosäuren zu bestimmen.

Struktur der Population

Zur Kontrolle der Populationsstruktur führten wir eine phylogenetische Analyse durch, die mit der Erstellung eines phylogenetischen Baums aus der VCF-Datei mit vcf-kit (v 0.2.9) begann. Die Distanzmatrizen zwischen den Proben wurden dann mit Pyseer (v 1.3.11) erstellt. Auf diese Abstände wurde eine multidimensionale Skalierung (MDS) angewandt, und Ausreißer aus jedem Kampagnendatensatz wurden entfernt. Mit der Pyseer-Funktion scree_plot_pyseer wurden Scree-Plots erstellt, um die optimale Anzahl von Dimensionen zu bestimmen, wobei neun Dimensionen als Kovariaten im GWAS-Modell beibehalten wurden. Zusätzliche Kovariaten, wie z. B. die damit verbundene Art des Tourismus und Informationen über das Kennzeichen, wurden auf der Grundlage der Metadatenanalyse ebenfalls einbezogen.

Genomweite Assoziationsstudie (GWAS)

Ein lineares Modell mit festen Effekten wurde mit Pyseer implementiert, um den Zusammenhang zwischen den Mutationen und den Resistenzwerten (Durchmesser der Hemmzone bei der Scheibendiffusionsmethode) zu bestimmen. Um falsch-positive Ergebnisse zu reduzieren, wurde mit demselben Tool eine Korrektur für Mehrfachtests durchgeführt, wobei der Schwellenwert für den p-Wert mit der Bonferroni-Methode angepasst wurde.

Biologische Analyse der Ergebnisse

Zur Analyse der biologischen Funktionen, die mit den Genen verbunden sind, die diese Varianten enthalten, haben wir die Datenbanken CARD(<https://card.mcmaster.ca/>), UniProt(<https://www.uniprot.org/>) und NCBI Gene(<https://www.ncbi.nlm.nih.gov/gene>) herangezogen. Auf diese Weise konnten wir feststellen, ob diese Gene eine Antibiotikaresistenz verleihen oder möglicherweise am Resistenzphänotyp beteiligt sind.

Ergebnisse und Diskussion bezüglich der Verbreitung von E. coli, Resistenzen, Virulenz und Sequenztypen

Von den 1437 eingegangenen DNA-Proben wurden 1136 Proben zurückbehalten und anhand von Qualitäts- und Taxonomiekriterien analysiert.

Auswertung der antimikrobiellen Resistenzgene (ARGs)

In den Proben wurden insgesamt 111 einzigartige Gene gefunden. In Übereinstimmung mit den anfänglichen Hypothesen dieses Projekts wurde bei den Isolaten aus der Flugzeugsammlung ein signifikant höherer ARG-Gehalt festgestellt als bei allen anderen Gruppen (Kruskal-Wallis-Test, Bonferroni-korrigierter p-Wert < 0,001 für jeden Vergleich). Erwartungsgemäß wiesen die Isolate mit den geringsten ARG-Werten in der vom Tourismus wenig beeinflussten Region auf (Abbildung 2).

Unter der Hypothese, dass der Tourismus die Verbreitung von ARG fördert, ist eine Diskrepanz zwischen dem ARG-Gehalt von Isolaten aus Flugzeugen und solchen aus dem touristischen Gebiet zu beobachten. Dies könnte dadurch erklärt werden, dass i) die beprobte Bevölkerung in der Touristenregion nur einen Bruchteil der Menschen ausmacht, die auf die Insel kommen, ii) die Proben aus den Touristengebieten auch Proben aus der Kläranlage enthalten, die nicht ausschließlich das Wasser aus den Touristengebieten behandelt, was zu einem möglichen Verdünnungseffekt führt. Dies unterstreicht auch, dass Flugzeuge als besserer Proxy für die Bewertung der Auswirkungen des Tourismus dienen könnten als die Orte auf dem Festland an sich.

Andererseits deutet das Fehlen einer Signifikanz zwischen dem krankenhausbazogenen Kontinuum und der nicht-touristischen Region auf eine niedrige Resistenzbasis in Guadeloupe im Allgemeinen hin, selbst in nosokomialen Umgebungen, was die Rolle der Flugzeuge, die auf die Insel kommen, noch mehr unterstreicht.

Bewertung der Virulenzgene (VGs)

Es wurden insgesamt 282 verschiedene Virulenzgene nachgewiesen. Anders als erwartet war die Übertragung im krankenhausbazogenen Kontinuum trotz der anfänglichen Beobachtungen nicht höher, stattdessen war die beobachtete Verteilung in allen Gruppen ähnlich, ohne signifikante Unterschiede. (Abbildung 2)

In Anbetracht der Tatsache, dass nicht jede E. coli VF die gleichen Auswirkungen hat. Die getesteten Isolate wurden in einen von neun Pathotypen nach Nascimento et al.¹ eingeteilt. Die Anzahl der jedem Pathotyp zugeordneten Isolate wurde auf die Anzahl der in jeder Gruppe vorhandenen Isolate normiert, um mögliche Auswirkungen auf die Stichprobengröße zu berücksichtigen. (Abbildung 1). Erwartungsgemäß war der Anteil der extraintestinalen pathogenen E. coli (EXPEC) und der uropathogenen E. coli (UPEC) in den Proben aus dem krankenhausbazogenen Kontinuum höher; der Anteil dieser beiden Pathotypen war in den übrigen analysierten Gruppen ähnlich. Andererseits waren zwei Pathotypen interessant verteilt: enteroaggregative E. coli (EAEC) und atypische enteropathogene E. coli (aEPEC) wiesen einen höheren Anteil an Proben aus Flugzeugen und touristischen Gebieten auf als an Proben aus Krankenhäusern und einen deutlich höheren Anteil als an Proben aus nichttouristischen Gebieten. Bei den enterotoxischen E. coli (ETEC) wurde in der Region mit geringem Tourismus ein wesentlich höherer Anteil festgestellt als in der touristischen Region; trotz dieses unerwarteten Ergebnisses ist es erwähnenswert, dass die Anzahl der Isolate, die zu diesem Pathotyp gehören, gering war.

Es wurden Isolate mit gemeinsamen Pathotypen festgestellt, von denen EAEC-EXPEC-UPEC mit 11 Isolaten aus Flugzeugen, 7 aus der Tourismusregion, 2 aus dem Krankenhausbereich und 3 aus dem nichttouristischen Bereich am häufigsten vorkamen. Zwei Isolate aus dem Krankenhauskontinuum wiesen die Kombination EXPEC-UPEC auf, während zwei Isolate aus dem touristischen Kontinuum EAEC-EPEC-EXPEC-UPEC-Merkmale aufwiesen. Diese Ergebnisse machen deutlich, dass, auch wenn der Gesamt-VF-Gehalt pro Isolat in jedem Kontinuum ungefähr gleich ist, tourismusbezogene Aktivitäten zur Einführung wichtiger E. coli-Pathotypen, aber auch relevanter hybrider Pathotypen beitragen können.

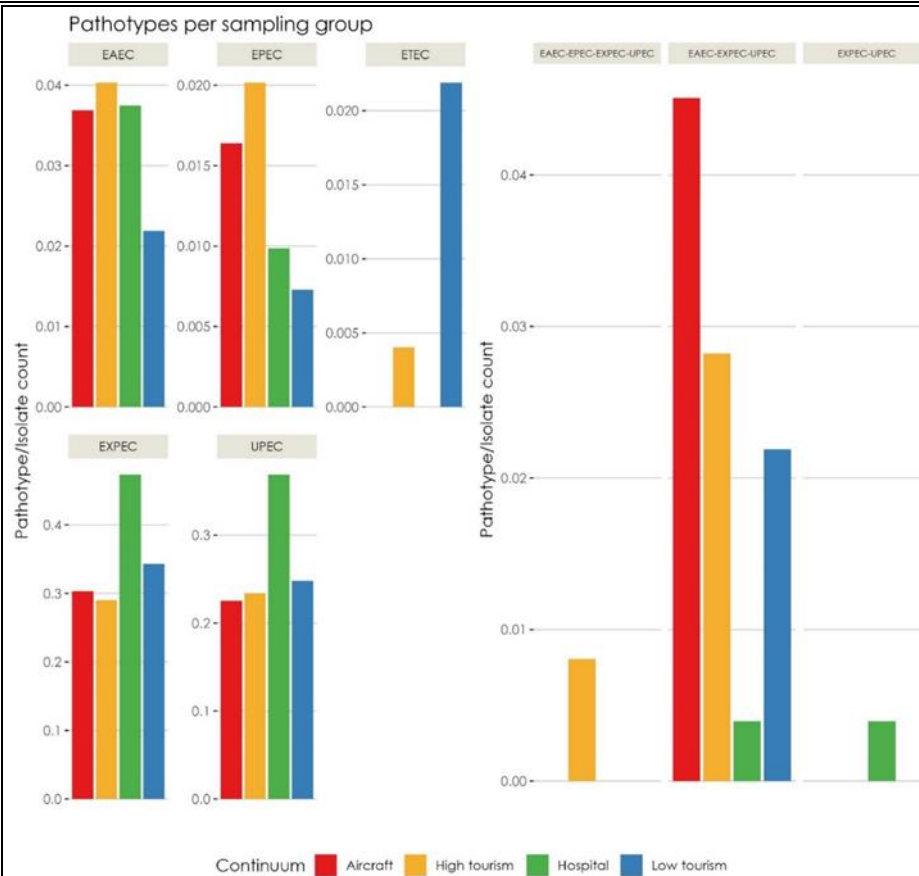


Abbildung 1: Verteilung der Pathotypen auf die untersuchten Gruppen.

Auswertung der Replikone

Insgesamt wurden 45 verschiedene Replikone identifiziert, wobei eine ähnliche Anzahl von Replikonen/Isolaten in den verschiedenen Gruppen beobachtet wurde (Mittelwert = 3/Isolate - Abbildung 2). Eine Erklärung hierfür könnte in dem breiten Spektrum der Plasmidfunktionen zu finden sein, bei denen nicht nur Virulenz und antimikrobielle Resistenz, sondern auch sehr unterschiedliche Stoffwechselwege und damit eine größere Heterogenität berücksichtigt werden können.

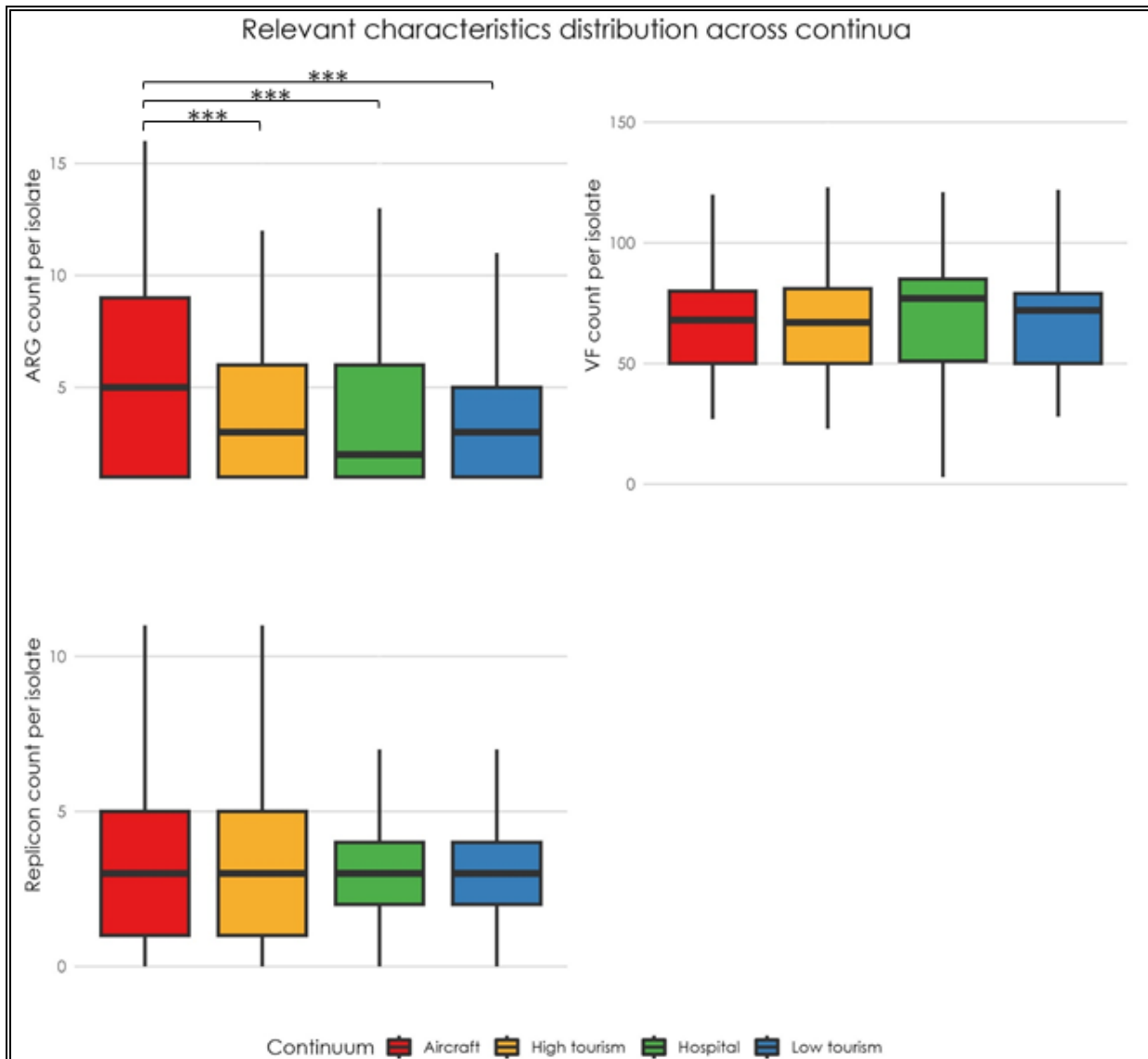


Abbildung 2: ARGs, VFs und Replikonverteilung pro Isolat in den vier untersuchten Gruppen. ***= p<0.001

Analyse der Populationsstruktur

Die Populationsstruktur wurde durch eine phylogenetische Analyse bewertet, insbesondere durch die Analyse der Sequenztyp-Verteilung (ST) und eine pangenomische Analyse.

Sequenztyp-Analyse

Jedem E. coli wurde anhand des von Achtman et al. vorgeschlagenen MLST-Schemas und der auf dem PubMLST-Server hinterlegten Allele der entsprechende ST zugeordnet. Unter den analysierten Probenahmegruppen erwies sich das krankenhausbezogene Kontinuum mit 0,26 eindeutigen Sequenztypen/Isolat (95 % CI: 0,22 - 0,29) als am wenigsten divers, was aufgrund der Annahme, dass nur eine begrenzte Anzahl von E. coli STs tatsächlich mit menschlichen Krankheiten in Verbindung steht, eine erwartete Beobachtung war. Andererseits war die vielfältigste Gruppe diejenige, die nicht mit dem Tourismus zusammenhing, mit einem Verhältnis von ST zu Proben von 0,41 (95% CI: 0,33 - 0,49), gefolgt von der Gruppe mit hohem Tourismusanteil (0,35, 95% CI: 0,29 - 0,42) und Proben aus Flugzeugen (0,34, 95% CI: 0,28 - 0,40) (Abbildung 3). In den drei letztgenannten Fällen erklärt sich die höhere Diversität dadurch, dass nicht nur potenzielle Krankheitserreger, sondern auch eine große Vielfalt an kommensalen E. coli nachgewiesen werden konnte.

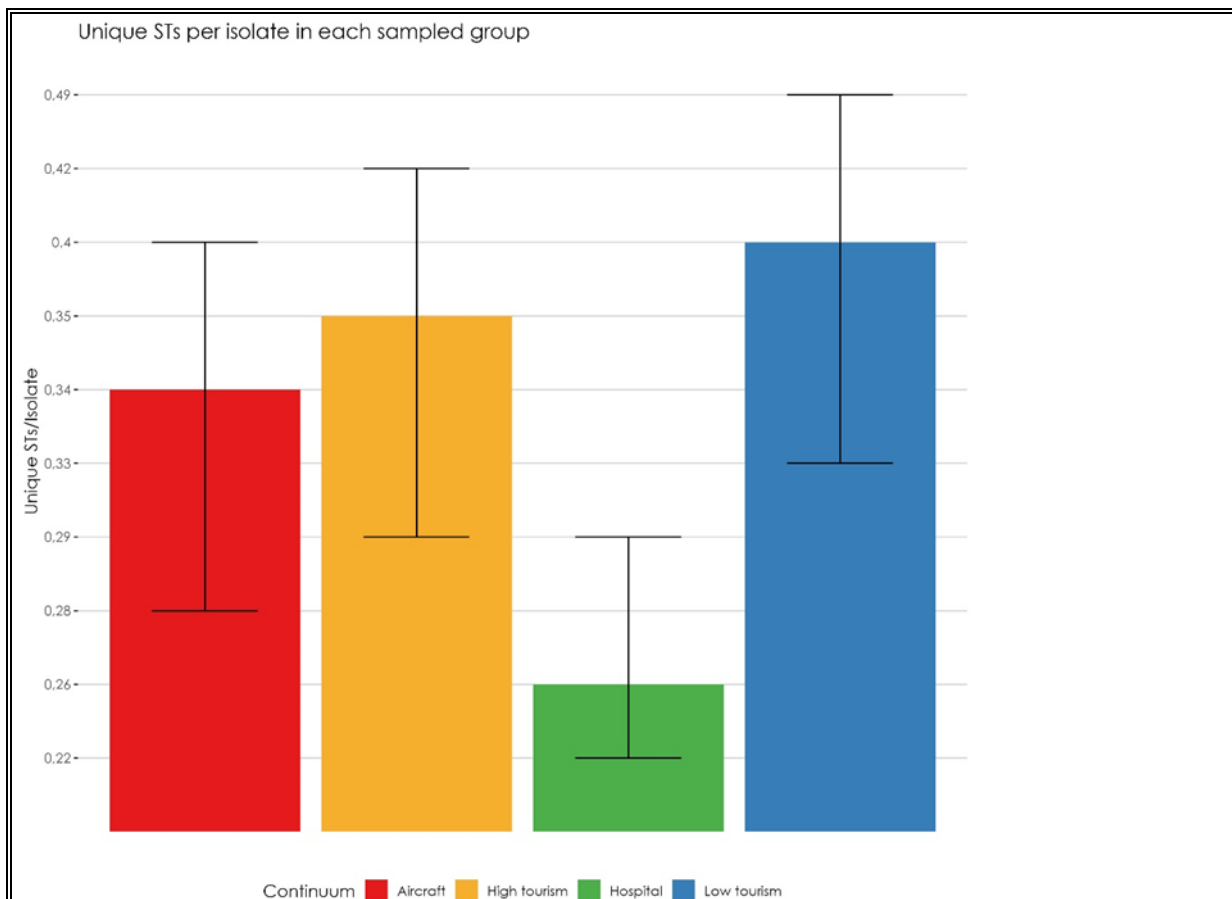


Abbildung 3: Anteil einzigartiger Sequenztypen in jeder Stichprobengruppe. Die Balken stellen das 95%-Konfidenzintervall dar.

Es ist wichtig hervorzuheben, dass sich der ST-Gehalt der untersuchten Gruppen deutlich unterscheidet. Nur 25 % bzw. 33 % der einzigartigen STs in den Proben aus dem Flugzeug und der touristischen Region waren mit denen aus der nicht-touristischen Region identisch, was auf einen wichtigen Beitrag neuer Sequenztypen zur „einheimischen“ ST-Population schließen lässt. 51 % bzw. 42 % der einzigartigen STs, die in den Flugzeugen und in der touristischen Region gefunden wurden, waren auch in den Krankenhausproben vorhanden.

Insgesamt wurden 236 eindeutige STs nachgewiesen, deren Häufigkeit jedoch ungleich war. 6 STs (69, 38, 131, 10, 73 und 95) waren am häufigsten vertreten und machten 36 % der nachgewiesenen Isolate aus. Es ist erwähnenswert, dass diese Isolate zu den wichtigsten und am weitesten verbreiteten pathogenen E. coli gehören. Vier von ihnen (ST69, ST131, ST73 und ST95) gelten als besonders erfolgreiche Klone, weshalb ihr Vorhandensein in jeder Probengruppe erwartet wurde, wobei die relative Häufigkeit in den krankenhausbefundenen Proben möglicherweise höher war. Interessanterweise traf dies nur auf die STs 73 und 95 zu, die auch in den krankenhausbefundenen Proben in geringen Anteilen vorhanden waren. Die Sequenztypen 69 und 131 waren in höheren Anteilen vorhanden, wenn man eine der tourismusbezogenen Probengruppen mit den krankenhausbefundenen Proben oder den nichttouristischen Proben verglich. Dieser Unterschied war sogar noch größer, wenn die Flugzeugproben und die Proben aus der touristischen Region zusammen betrachtet wurden.

Der Sequenztyp 38 zeichnete sich dadurch aus, dass er als einziger ST am häufigsten in der Region mit geringem Tourismus gefunden wurde, dicht gefolgt von der touristischen Region. Kürzlich wurden in der Europäischen Union erhöhte Erholungsraten und sogar Ausbrüche mit diesem ST dokumentiert, was ST38 zu einem potenziell grenzüberschreitend auftretenden Klon macht. Auch wenn weitere Bestätigungen erforderlich sind, könnte der ähnliche Anteil von ST38-Isolaten in touristischen und nichttouristischen Gebieten die Auswirkungen des Tourismus auf die Einschleppung und Ausbreitung spezifischer Krankheitserreger verdeutlichen, wenn man bedenkt, dass eine große Zahl von Touristen auf die Inseln der EU kommt.

Pangenom-Analyse

Der Vergleich des genetischen Gesamtinhalts ermöglichte uns die Berechnung eines Pangenoms für jede der Probenahmegruppen. In Übereinstimmung mit der Literatur war die Größe des Kerngenoms (Gene, die von $\geq 95\%$ der Isolate geteilt werden) in jeder Gruppe konserviert und hatte einen Mittelwert von 2901 (95% CI $\pm 13,20$). Eine genauere Betrachtung ergab signifikante Unterschiede zwischen den einzelnen Gruppen. Abbildung 4 zeigt die Variation der Pangenomgröße (Kern- und Nicht-Kerngene) mit dem Hinzufügen neuer Isolate. Auch wenn die Aufnahme neuer Genome erwartungsgemäß zu einer Vergrößerung des Pangenoms führte, war die Anzahl der neu hinzugefügten Gene bei Proben aus dem Krankenhausbereich deutlich höher. Die Proben aus dem Tourismusbereich wiesen zwar eine geringe Variation zwischen ihnen auf, zeigten aber im Vergleich zur nicht-touristischen Region einen bemerkenswert höheren Gehalt an neuen Genen. Diese Ergebnisse zeigen die hohe genetische Vielfalt von *E. coli*-Isolaten in tourismusbezogenen Proben und unterstreichen deren Rolle als Quelle genetischer Variation.

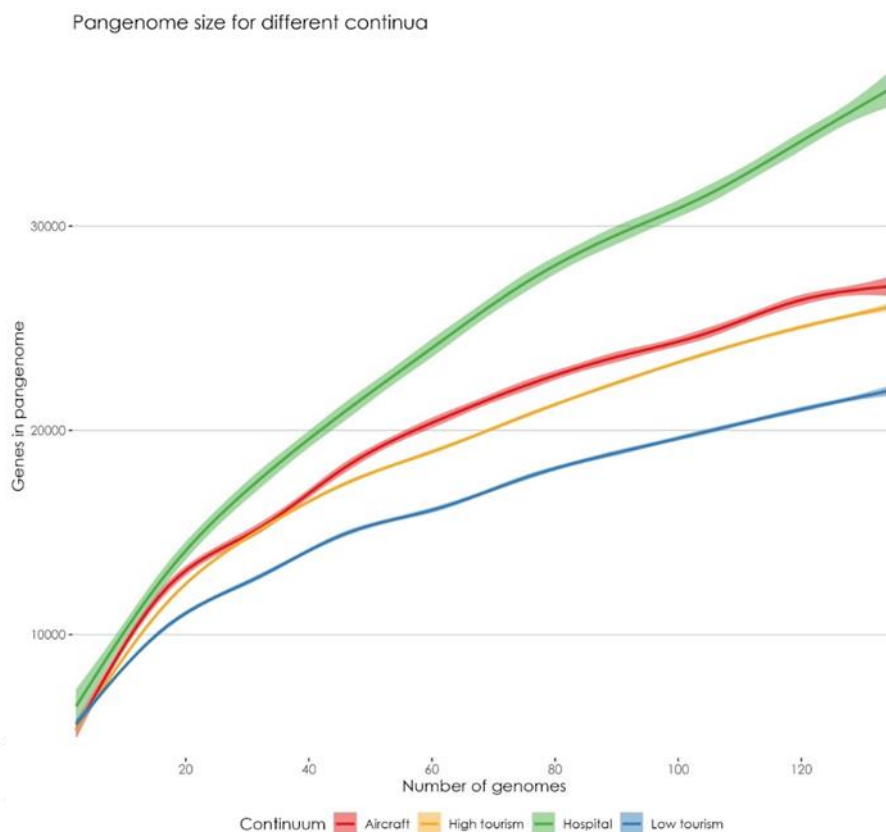


Abbildung 4: Veränderung der Pangenom-Größe mit der Hinzufügung neuer Genome. Die Kurven wurden mithilfe eines verallgemeinerten additiven Modells angepasst; die heller schattierten Bereiche stellen ein 99 %-Konfidenzintervall dar. Um mögliche Auswirkungen auf den Stichprobenumfang zu berücksichtigen, wurden die Kurven auf den kleinsten Stichprobenumfang ($n=137$) reduziert.

Die Proben aus der lokalen Bevölkerung (nicht-touristisches Kontinuum) spiegeln größtenteils die bereits berichteten phäno- und genotypischen Merkmale von *E. coli* wider, einschließlich der vorherrschenden Sequenztypen und Pathotypen²⁻⁴. Darüber hinaus stellten wir erwartungsgemäß fest, dass Krankenhäuser und insbesondere deren Abwässer als Selektionsgründe für Krankheitserreger dienen, die mehrere antimikrobielle Resistenzgene tragen und von denen ein Teil auch nach der Abwasserbehandlung in die aufnehmenden Gewässer gelangt. Dieses Phänomen ist weithin beschrieben worden und trägt in gewissem Maße zum Vorhandensein von ARGs in vom Menschen beeinflussten Umgebungen bei. Besonders bemerkenswert ist, dass die vorgestellten Ergebnisse zeigen, dass *E. coli* mit einer deutlich höheren ARG-Belastung durch touristische Aktivitäten in die einheimische Umwelt eingebracht werden könnten. In diesem Zusammenhang haben wir auch festgestellt, dass touristische Standorte nützliche Informationen liefern können, die jedoch für die tatsächliche Situation nicht repräsentativ sind, so dass die Sammlung mit dem Flugzeug eine bessere Alternative für solche Studien darstellt.

Hochvirulente Stämme schienen häufig bei Krankenhauspatienten aufzutreten, während sie im Abwasser nicht selektiert zu werden schienen, so dass dieses eine geringere Quelle für solche Erreger in der städtischen Umwelt darstellt. Andererseits wiesen virulente Pathotypen aus Proben, die mit touristischen Aktivitäten in Verbindung gebracht wurden, eine vergleichbare Häufigkeit auf wie diejenigen, die in Krankenhausproben gefunden wurden, was darauf hindeutet, dass Touristenströme in ähnlicher Weise zur Verbreitung besonders pathogener Isolate beitragen könnten wie infizierte Patienten aus Krankenhäusern.

Noch interessanter ist die potenzielle genetische Diversifizierung im Zusammenhang mit dem Tourismus. In dieser Studie wurde eine hohe Diversität der Sequenztypen beobachtet, die jedoch zu etwa 30 % mit der der einheimischen Bevölkerung übereinstimmte, was die Rolle des Tourismus bei der Ausbreitung und Einschleppung von nicht endemischen STs verdeutlicht. Besonders deutlich wird dies am Beispiel von ST38 und ST10. Dadurch erhöht sich die Chance der Einführung und Verbreitung von eventuell erfolgreichen Klonen. Darüber hinaus ist die vergleichbare Pangenomgröße der beiden touristischen Proben hervorzuheben, die signifikant höher war als die der nicht-touristischen Proben. Ein derart großes Pangenom deutet auf die immense genetische Vielfalt der Isolate hin, die in Abhängigkeit vom Tourismus zirkulieren, was in direktem Zusammenhang mit einem größeren Mobilom und einem höheren Potenzial für den Erwerb von Genen und die Entstehung neuer Gene steht, zu denen auch ARGs gehören könnten.

Ergebnisse und Diskussion bezüglich Genome Wide Association Study und Phänotypischer Susceptibility

Analyse der Metadaten

Die Analyse der phänotypischen Daten/Metadaten unserer Proben war entscheidend für die Bestimmung von Faktoren, die einen Einfluss auf die Empfindlichkeit der Proben gegenüber den verschiedenen Antibiotika haben. Dies war nicht nur wichtig, um zugrundeliegende Verzerrungen zu ermitteln, sondern auch für die Bestimmung potenzieller Kovariaten für die Assoziationsstudien, die wir im Anschluss an diesen Schritt durchführen würden. Wir untersuchten die Auswirkungen verschiedener Faktoren auf die Anfälligkeit der Proben und entschieden uns schließlich dafür, Tourismusgruppen und Sequenzierplatten als Kovariaten in unsere nachgelagerte Analyse aufzunehmen. Als Beispiel werden die Auswirkungen des Tourismus im Folgenden näher erläutert.

Der Einfluss des Tourismus auf die Anfälligkeit der Proben

Die Studie konzentrierte sich auf die Bewertung des Einflusses von Tourismus/menschlichen Aktivitäten auf die Antibiotikaresistenz. Eine phänotypische/Metadaten-Analyse kategorisierte die Proben auf der Grundlage ihrer touristischen Exposition in „hoher Tourismus“, „geringer Tourismus“, „Krankenhaus“ oder „Flugzeug“. „Hoher Tourismus“ bezeichnete Orte mit hoher touristischer Aktivität, während „geringer Tourismus“ sich auf Gebiete mit geringer touristischer Aktivität bezog. „Krankenhaus“ stand für Proben aus klinischen Umgebungen und der dazugehörigen Abwasseraufbereitung, und „Flugzeug“ bezeichnete Proben, die aus Flugzeugtoiletten entnommen wurden. Es wurde kein signifikanter Unterschied zwischen den Gruppen „hoher Tourismus“ und „geringer Tourismus“ festgestellt. Die Proben aus Krankenhäusern wiesen jedoch in einigen Fällen eine unterschiedliche Verteilung auf (siehe Abbildung 5).

Proben aus Regionen mit unterschiedlich starker touristischer Aktivität wiesen nicht nur keine Unterschiede in der Anfälligkeit auf, sondern zeigten auch einheitliche genomische Profile, wie in Abbildung 6 dargestellt.

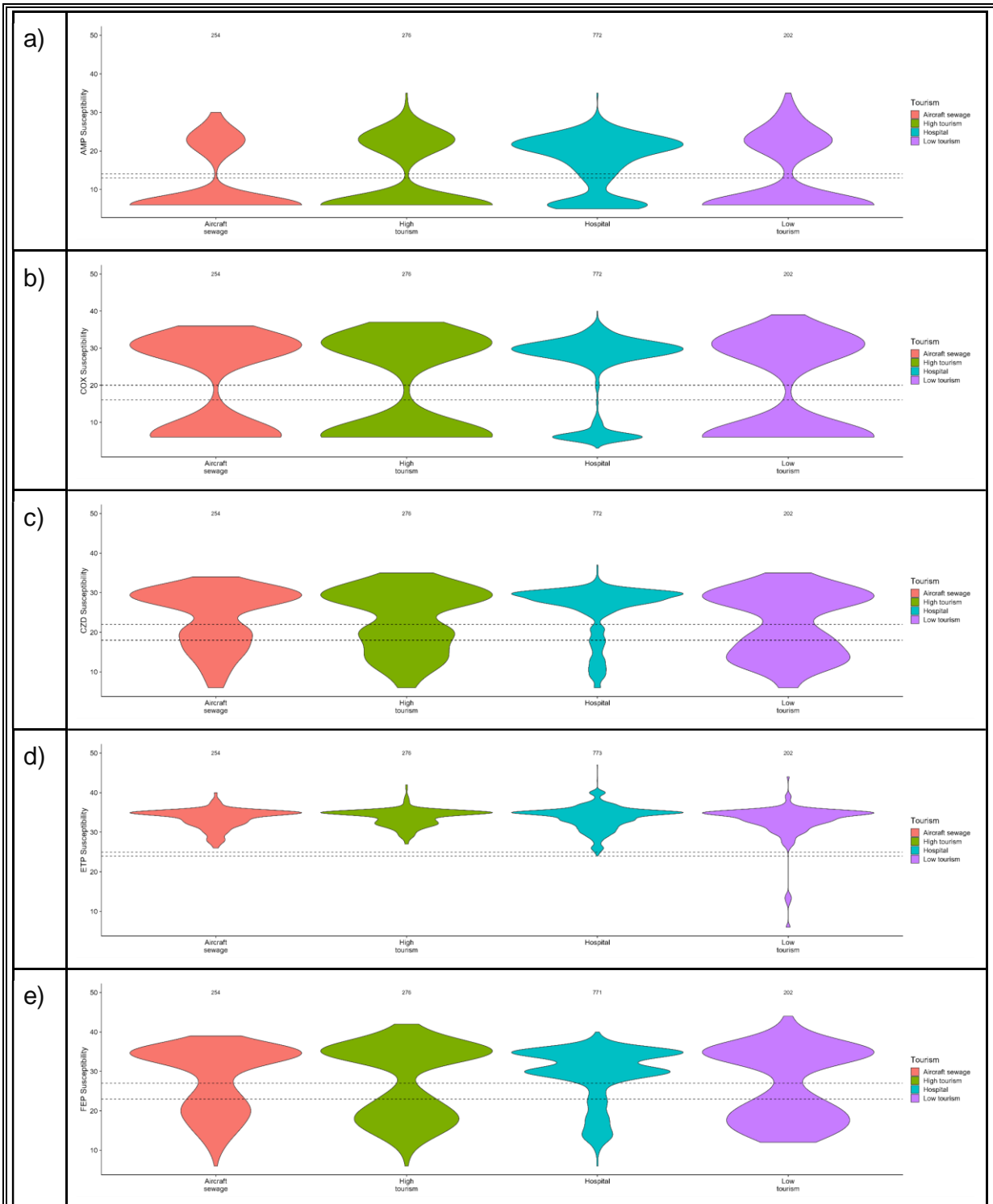


Abbildung 5. Geigenplots, die die Empfindlichkeitsverteilung unserer Proben gegenüber verschiedenen Antibiotika (a bis i) in verschiedenen Regionen, so genannten Tourismusgruppen, zeigen. Die gestrichelten Linien zeigen die oberen und unteren Empfindlichkeitsschwellen an. Es wurden keine signifikanten Unterschiede in der Anfälligkeit der Proben zwischen Regionen mit hohem und niedrigem Tourismus festgestellt. Die Proben aus dem Krankenhaus wiesen jedoch in einigen Fällen unterschiedliche Verteilungen auf.

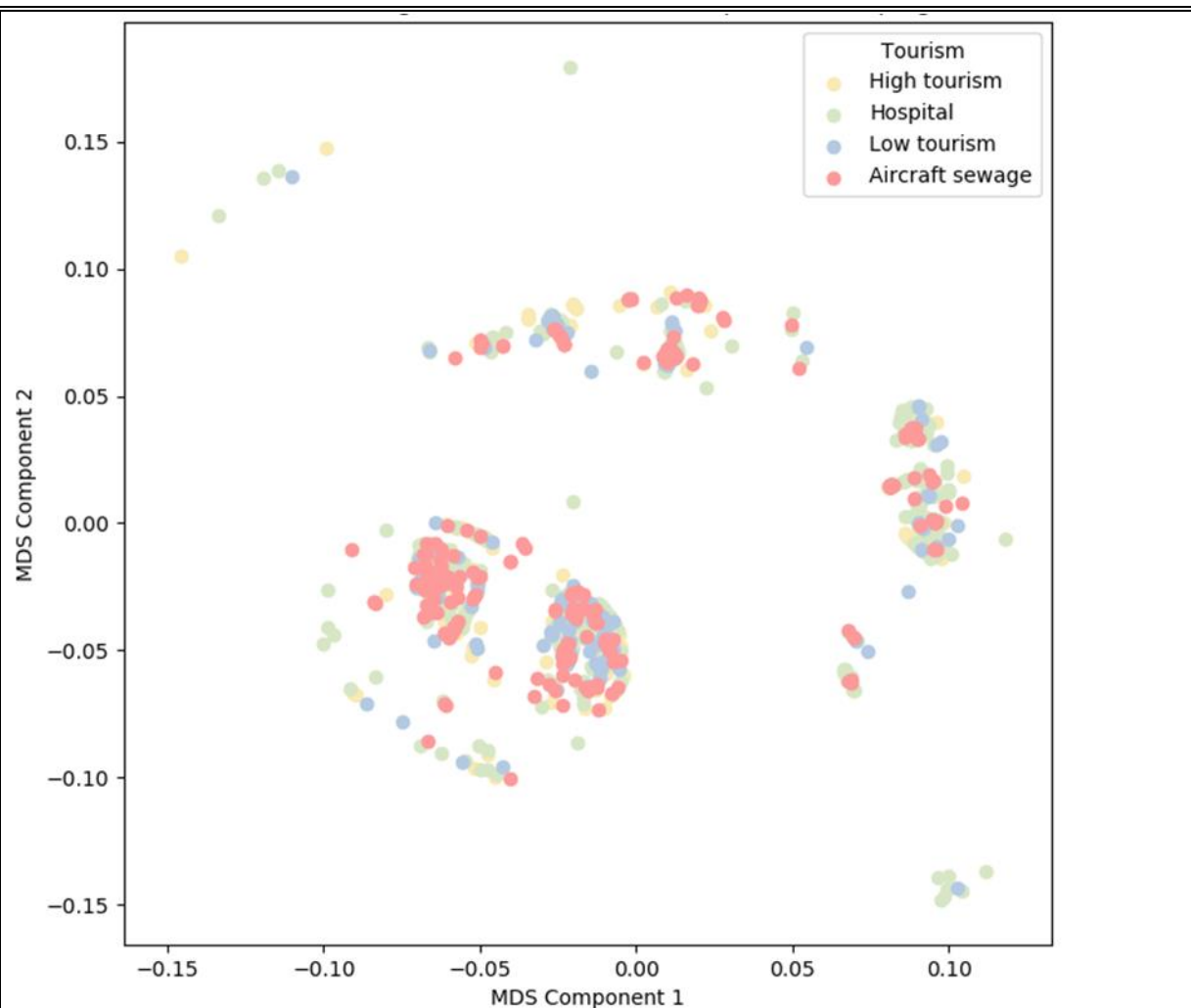


Abbildung 6. MDS-Diagramm basierend auf den genomischen Abständen der Proben. Es gibt keine erkennbare Clusterung der Proben auf der Grundlage des Grades der touristischen Aktivität in den Regionen, unabhängig davon, ob es sich um Gebiete mit hohem oder niedrigem Tourismusaufkommen, Flugzeugabwässer oder Krankenhäuser handelt.

Von Raw-Reads zu hochwertigen Reads

Die Auswirkungen der verschiedenen Qualitätsstufen auf die Rohdaten sind in Tabelle 1 dargestellt. Zur Bewertung der Probenqualität wurden verschiedene Metriken analysiert, wobei der Schwerpunkt auf der „Per base Sequence Quality“ lag, die den phred-Score für jede Basenposition innerhalb eines Proben-Reads angibt. Ein Score von 30 bedeutet eine Basenfehlerrate von 0,1 % und gewährleistet eine Genauigkeit von 99,9 % bei der Basenerkennung. Alle Proben aller Kampagnen bestanden diese Metrik, was auf eine Basenqualität von über 30 in den ursprünglichen Rohsequenzen hinweist.

Obwohl alle Proben bei der Analyse des Gehalts an Basen pro Sequenz durchfielen, wurde dies auf der Grundlage früherer Untersuchungen als akzeptabel angesehen. Dieser Misserfolg wurde auf das mögliche Vorhandensein von Adaptersequenzen am Anfang jeder Lesesequenz zurückgeführt. Trotz einer Warnung über den GC-Gehalt der Sequenzen in allen Proben-Reads wurde bei der nachgeschalteten Analyse auf mögliche Verunreinigungen eingegangen. Es wurde eine Sequenzduplikation festgestellt, die auf eine mögliche PCR-Amplifikationsverzerrung hinwies, die jedoch später mit dem Picard-Tool MarkDuplicates behoben wurde. Das Vorhandensein von Nextera-Adaptersequenzen in Kampagne 3-Reads wurde erfolgreich mit dem Trimmomatic-Tool entfernt, wodurch die Qualität des Alignments verbessert und ungenaue Varianten-Calls in den nachfolgenden Schritten minimiert wurden.

Tabelle 1. Zusammenfassende Tabelle der mit den Tools FastQC und MultiQC durchgeführten Datenqualitätsprüfungen. Die Proben wurden vor und nach dem Trimm-/Kurationsschritt analysiert.

Analysis	Condition	Campaign 1&2 samples		Campaign 3 samples		Campaign 4 Samples	
		Before trimming	After trimming	Before trimming	After trimming	Before trimming	After trimming
Per Base Sequence Quality	Passed	1454	1454	704	704	714	714
	Warning	0	0	0	0	0	0
	Failed	0	0	0	0	0	0
Per Base Sequence Content	Passed	0	0	0	0	714	714
	Warning	0	0	0	0	0	0
	Failed	1454	1454	704	704	0	0
Per Sequence GC content	Passed	0	0	0	0	0	0
	Warning	1448	1448	702	702	704	711
	Failed	6	6	2	2	10	2
Sequence Length Distribution	Passed	0	0	704	704	0	0
	Warning	1454	1454	0	0	714	714
	Failed	0	0	0	0	0	0
Sequence Duplication Levels	Passed	5	5	0	704	0	714
	Warning	679	679	291	0	118	0
	Failed	770	770	413	0	596	0
Overrepresented Sequences	Passed	1454	1454	704	704	709	704
	Warning	0	0	0	0	5	10
	Failed	0	0	0	0	0	0
Adapter Sequence Content	Passed	1454	1454	35	704	714	714
	Warning	0	0	661	0	0	0
	Failed	0	0	8	0	0	0

Von Rohvarianten zu hochwertigen Varianten

Nach Anwendung von Qualitätskontrollschritten wurde etwa ein Viertel der Varianten beibehalten (siehe Tabelle 2). Beim ersten Aufruf von Varianten durch HaplotypeCaller betrug die Gesamtzahl der Varianten 913.379. Diese verringerte sich nach jedem Filterungsschritt, mit Ausnahme des Normalisierungsschritts. Dies führte zu einer Aufspaltung der VCF-Dateidatensätze, wobei zuvor

mehrere alternative Allele für eine einzelne Variantenposition als ein einziger Datensatz aufgezeichnet wurden, während nun jedes alternative Allel in einer einzelnen Zeile dargestellt wurde. Die größte proportionale Veränderung bei den Varianten wurde nach der Entfernung von Varianten festgestellt, die den Cutoff für die Minor-Allel-Häufigkeit (MAF) nicht erfüllten.

Tabelle 2. Veränderung der Variantenzahl während der Qualitätskontrollschritte für alle Proben über alle Kampagnen hinweg. Von den ursprünglich 859.231 Varianten wurden nach den Filterungsschritten 248.602 Varianten beibehalten.

Step	No. of SNPs	No. of Indels	Total no.	Proportional change in variants
Initial calling of variants	859,231	54,148	913,379	-
Hard filtering	805,083	32,902	837,985	-8.25
Filtering by GQ and missingness	796,055	31,161	827,216	-1.28
Splitting alternative alleles	882,993	38,693	921,686	+11.4
Filtering by MAF	-	-	248,602	-73.03

Von hochwertigen Varianten zu hochsignifikanten Varianten

Genomische Varianten, die mit den Resistenzwerten der 14 Antibiotika in Verbindung stehen, wurden identifiziert und nach ihrer Signifikanz gefiltert. Zur Identifizierung solcher signifikanter Varianten wurde der p-Wert nach der Bonferonni-Korrektur herangezogen. Die Varianten wurden dann nach ihrer Signifikanz sortiert; die 10 wichtigsten Varianten, die mit jedem Antibiotikum assoziiert sind, sind in Tabelle 3 zu sehen .

Unter Verwendung der Gendatenbanken CARD, UniProt und NCBI wurden die zellulären Funktionen der Gene, die diese Varianten beherbergen, aufgeklärt, kategorisch in vier Hauptgruppen eingeteilt und entsprechend farblich kodiert. Die Farbkodierung der Varianten erfolgte auf der Grundlage ihrer Assoziation: gelb für eine Antibiotikaresistenz, die spezifisch für ein bestimmtes Antibiotikum ist, orange für eine allgemeine Antibiotikaresistenz, die jedoch nicht spezifisch für dieses Antibiotikum ist, grün für Varianten, die zuvor von der Schroeder-Gruppe identifiziert wurden, und blau für Varianten, die mit der Antibiotikaresistenz in Verbindung stehen. Bemerkenswerterweise ergab unsere Analyse Treffer in allen vier Kategorien und brachte neue Genomvarianten ans Licht, die weiter erforscht werden müssen.

Tabelle 3. Top 10 der signifikanten Varianten von Proben mit Tourismus und Platte, Kovariaten. Die Varianten wurden farbcodiert, je nachdem, ob sie mit Antibiotikaresistenz assoziiert waren (orange), mit einer für ein bestimmtes Antibiotikum charakteristischen Resistenz assoziiert waren (gelb), in einer früheren Arbeit der Schroeder-Gruppe gefunden wurden (grün) oder mit der Antibiotikaresistenzreaktion zusammenhängen (blau).

	AMC	AMP	TIC	FOX	COX	CZD	FEP	ETP	AKN	GMN	NAL	LVX	SXT	FOS
YmfQ P148A	gyrA D87N	gyrA D87N	yegT S158	gyrA S83L	malY A308	malY A308	ycjW T68	uvrA UPSTREAM	fadK A442D	gyrA S83L	parC S80I	gyrA S83L	fabR A225 T	
AfuC UPSTREAM	gyrA S83L	gyrA S83L	aceB Q56	malY A308g	csgG UPSTREAM	clsB A366	ynal UPSTREAM	gspA V7	pdxY L29	parC S80I	gyrA D87N	fecB UPSTREAM	dxr L297	
YmfJ UPSTREAM	fadI I309	parC S80I	yegT F147	fadI I309	clsB A366	ampG g320	fucP UPSTREAM	NrfA A58	gitX V432	gyrA D87N	gyrA S83L	parC S80I	proX Q115	
TfaE UPSTREAM	nimT L354	nimT L354	aceB I58	gyrA D87N	amp G320G	fadI I309	pdeR UPSTREAM	AlsR I64	ung S192	pitB N61	yghO UPSTREAM	gyrA D87N	tdcA A37	
McrA V59I	malY A308	yihS L281	ycdJ R408C	clsB A366	csgG UPSTREAM	ygiQ L491	fucl A281T	NrfA A362	alx V119	speC UPSTREAM	nuoA UPSTREAM	fecB UPSTREAM	rhaM L52	
YmfQ T115S	yihS L281	fadI I309	gatC UPSTREAM	cysI F249	cysI F249	gyrA D87N	sapF H216	YjcO A174	glnE G265	speC UPSTREAM	yghQ L339	amtB L405	ygiQ S185	
McrA L11	parC S80I	ftsK L198	nanR T11	parC S80I	fadI I309	gyrA S83L	frvB K82	YniB S156	ttdT N215	speC UPSTREAM	yqjF UPSTREAM	ybaV S28P	fhlA R599	
Nfi L152	ampG G320	fecB UPSTREAM	yegT L80P	cysI S258	gpr S343	parC S80I	hslJ UPSTREAM	EcpE L82	ttdT F302	speC UPSTREAM	yqjF UPSTREAM	ybaV V29A	gldA G192	
SpeC UPSTREAM	nei P35S	malY A308	talB T266	exuR L98	mug UPSTREAM	gpr S343	tam A71	adeD S484	yqjG I49T	yghQ L34	yqjF UPSTREAM	rrrQ T5617	glyA G312	
Pine G158	artQ S137	yafZ UPSTREAM	ykgH L198I	cysI P224	thiP S436N	csgG UPSTREAM	glpA C71	adeD A485	yqjG T177P	speC G25A	yghR D2	fecB UPSTREAM	glyA S311	

Schlussfolgerung

Wir haben über 1000 E. coli-Proben sorgfältig analysiert und festgestellt, dass der Tourismus keinen signifikanten Einfluss auf die Anfälligkeit der Proben hat. Unsere GWAS-Analyse, bei der der Tourismus berücksichtigt wurde, deckte jedoch Tausende signifikanter Genomvarianten auf, darunter bekannte

Mutationen in gyrA und parC sowie neue Varianten, die möglicherweise mit Antibiotikaresistenz in Verbindung stehen. Daraus ergibt sich, dass diese neuen Varianten nach einer möglichen experimentellen Validierung in der Lage sind, bestehende Antibiotikaresistenz-Datenbanken wie CARD oder Resfinder zu verbessern oder vielleicht zur Entwicklung von Antibiotika beizutragen, die auf solche Mutationen abzielen.

Wichtigste Positionen des zahlenmäßigen Nachweises und Notwendigkeit der geleisteten Arbeiten

Kostenübersicht ACRAS-R			
Projekt:		ACRAS-R	
Förderkennzeichen:		16GW0355	
Projektleiter/-in:		Prof. Dr. habil. Thomas Berendonk	
Berichtszeitraum:		01.01.2023 - 30.04.2024	
Personalausgaben			254.775,85 €
	Zeitraum		Kosten
	0812 Beschäftigte E12-E15		236.522,44 €
	0817 Beschäftigte E1-E11		10.002,11 €
	0822 Beschäftigungsentgelte		8.251,30 €
Sächliche Verwaltungsausgaben			136.570,27 €
			Kosten
	0835 Aufträge		12.500,00 €
	0843 Sonst. Verwaltungsausgaben		105.441,45 €
	0846 Dienstreisen		18.628,82 €
Gesamt			
	Gesamtkosten (direkt)		391.346,12 €

Abbildung 7. Projektausgaben

Der bewilligte Ausgabenplan wurde im Wesentlichen eingehalten (Abbildung 7). Angesichts der komplexen Aufgaben im Projekt und waren die geleisteten Projektarbeiten notwendig und angemessen. Dies zeigt sich auch in den erfolgreichen Projektergebnissen (s.o.).

Veröffentlichungen und Verwertung

Projektbezogene Publikationen mit Peer Review

Fang P, Elena AX, Kunath MA, Berendonk TU, Klümper U. Reduced selection for antibiotic resistance in community context is maintained despite pressure by additional antibiotics. ISME Commun. 2023 May 31;3(1):52. doi: 10.1038/s43705-023-00262-4. PMID: 37258727; PMCID: PMC10232432.

Kneis, D., Lemay-St-Denis, C., Cellier-Goetghebeur, S. et al. Trimethoprim resistance in surface and wastewater is mediated by contrasting variants of the *dfxB* gene. ISME J 17, 1455–1466 (2023). <https://doi.org/10.1038/s41396-023-01460-7>

Malekian, N., Sainath, S., Al-Fatlawi, A., & Schroeder, M. (2023). Word-based GWAS harnesses the rich potential of genomic data for *E. coli* quinolone resistance. *Frontiers in Microbiology*, 14, 1276332.

Malekian, N., Agrawal, A. A., Berendonk, T. U., Al-Fatlawi, A., & Schroeder, M. (2022). A genome-wide scan of wastewater *E. coli* for genes under positive selection: focusing on mechanisms of antibiotic resistance. *Scientific Reports*, 12(1), 8037.

Malekian, N., Al-Fatlawi, A., Berendonk, T. U., & Schroeder, M. (2021). Mutations in *bdca* and *vals* correlate with quinolone resistance in wastewater *Escherichia coli*. *International journal of molecular sciences*, 22(11), 6063.

Projektbezogene Publikationen als Preprint veröffentlicht

Fang, P., Konyali, D., Fischer, E., Mayer, R.P., Huang, J., Elena, A.X., Orzechowski, G.H., Tony-Odigie, A., Kneis, D., Dalpke, A. and Krebs, P., 2023. Cigarette smoking promotes the spread of antimicrobial resistance in the human lung and the environment. *bioRxiv*, pp.2023-08.

Elena, Alan X., Neza Orel, Peiju Fang, Gerhard J. Herndl, Thomas U. Berendonk, Tinkara Tinta, and Uli Klümper. "Jellyfish blooms-an overlooked hotspot and potential vector for the transmission of antimicrobial resistance in marine environments." *bioRxiv* (2024): 2024-07.

Geplante Publikationen in Vorbereitung

Elena, A., Batantou, D., Klümper, U., Kneis, D., Breurec, S., Berendonk, T. - Tourism: A multifaceted threat in the dissemination of relevant bacterial pathogens. A two-year study on the Island of Guadeloupe.

Elena, A., Batantou, D., Klümper, U., Breurec, S., Berendonk, T. - Touristic activities play a major role in the dissemination of successful *Escherichia coli* clones harbouring blaCTX-M-15.

Elena, A., Klümper, U., Malekian, N., Schröder, M., Berendonk, T. - Epidemiologic and genetic characterisation of shiga-toxin producing *Escherichia coli* isolates recovered from animals grazing near a hospital wastewater treatment plant.

Elena, A., Klümper, U., Malekian, N., Schröder, M., Berendonk, T. - Genetic characteristics of *Escherichia coli* isolates carrying blaNDM recovered from mangroves receiving urban treated wastewater.

Präsentation der Projektergebnisse auf Konferenzen, etc.

DGL congress 2024 - Tourism-related activities as a potential source of *Escherichia coli* isolates with epidemiological significance on the Island of Guadeloupe. Elena Alan Xavier, Batantou Degrace, Klümper Uli, Breurec Sebastien, Berendonk Thomas Ulrich.

FEMS congress 2023 - Are we in trouble if we travel? – An Epidemiological analysis on the impact of tourism on the spread of relevant *Escherichia coli* isolates. Elena Alan Xavier, Batantou Degrace, Pimenta Melanie, Klümper Uli, Breurec Sebastien, Dagot Christophe, Berendonk Thomas Ulrich.

Konferenzteilnahmen der Schröder Gruppe:

- ISMB, 2021, online (poster)

- GLBIO, 2021, online (poster and talk)
- ECCB 2022, Sitges (poster)
- GCB 2022, Halle (poster)
- Maltaomics summer school 2023, Valletta
- NHR conference, 2023 Berlin
- Long Night of Science event, annually at Dresden
- SCADS AI meetings

Wissenschaftliche Verwertbarkeit, Impact und Anschlussfähigkeit

Die Versuchsplanung, insbesondere der Umfang der Probenahme, und die umfassende Analyse (Vorhersage mehrerer ARGs, Virulenzfaktoren, Sequenztypen und pangenomische Analyse) heben dieses Projekt von der bisherigen Forschung ab. Die Tiefe der hier erzielten Ergebnisse wird von den bereits verfügbaren Daten nicht übertroffen, was bedeutet, dass die aus diesem Projekt gewonnenen Ergebnisse wichtige Einblicke in die Beziehung zwischen Tourismus und bakteriellen Krankheitserregern liefern. Die wichtigsten Ergebnisse wurden bereits in einem Manuskript zusammengefasst, das derzeit von den Co-Autoren überarbeitet wird und als Open-Access-Publikation bei einer renommierten Fachzeitschrift (z. B. Journal of Travel Medicine) eingereicht werden soll.

Die sequenzierten Genome werden nach Abschluss des Projekts oder zusammen mit der Veröffentlichung des Manuskripts zur Verfügung gestellt und mit den entsprechenden Metadaten in offene Datenbanken (NCBI, ENA) hochgeladen. Die Hinzufügung von 1136 E. coli-Sequenzen wird einen bedeutenden Beitrag zu den genannten Datenbanken darstellen, die Wirkung unseres Projekts erweitern und sicherlich eine wertvolle Grundlage für künftige wissenschaftliche Fortschritte nicht nur in diesem Bereich, sondern in der E. coli-Genomik insgesamt bilden.

Wir haben eine Liste von genomischen Elementen identifiziert, die in drei öffentlich zugänglichen Veröffentlichungen ausführlich beschrieben sind (s. Eingehende Darstellung). Dies gibt anderen Wissenschaftlern und Forschern die Möglichkeit, unsere Ergebnisse zu validieren, zum Beispiel durch experimentelle Methoden. Nach erfolgreicher Validierung können diese Ergebnisse genutzt werden, um aktuelle Datenbanken zur Antibiotikaresistenz zu erweitern. Zudem können die neuen bioinformatischen Methoden in der zukünftigen Forschung genutzt werden und sind mit Open Access publiziert. So wurden die erhaltenen Stämme und entwickelten bioinformatischen Methoden beispielsweise schon in einer unserer Publikationen und zwei Preprints die nicht direkt ACRAS-R assoziiert waren genutzt.

Einige der gewonnenen Sequenzen werden bereits seit Beginn des ACRAS-R-Projekts im Rahmen des Whole Genome Sequencing Workshops des Instituts für Hydrobiologie verwendet, bei dem Bachelor- und Masterstudenten praktische Erfahrungen mit der bakteriellen Sequenzierung und Analyse sammeln können. Im Rahmen des Projekts haben wir zudem Masterstudenten wie Amay A. Agrawal, Conor Sexton und Jingbo He ausgebildet. Durch ihre Masterarbeit/ihr Projekt verfügen diese Studenten nun über hervorragende Fähigkeiten und Kenntnisse in der Verarbeitung von Sequenzierungsdaten und der Entwicklung von Analysepipelines, die sie auf künftige Aufgaben in der Branche vorbereiten.

Zudem ist bereits ein neuer Projektantrag basierend auf dem hier gewonnenen Datensatz eingereicht worden, um eine Anschlussfinanzierung zu gewährleisten.

Einschlägige Veröffentlichungen außerhalb des Projekts während der Projektlaufzeit

Während der Projektlaufzeit sind Veröffentlichungen erschienen, die einen Zusammenhang zwischen der Verbreitung antimikrobieller Resistenzen und dem Tourismus herstellen wollen. Einige von ihnen haben jedoch einen meta-analytischen Ansatz und behandeln den Begriff „Reisen“ auf eine breite Art und Weise, ohne beispielsweise eine klare Unterscheidung zwischen touristischen und Flüchtlingsbewegungen zu treffen (Bokhary et al., Petersen et al.)^{5,6}. Andere Autoren betrachteten die antimikrobielle Resistenz ausschließlich aus einer phänotypischen Perspektive (Patjas et al.)⁷, oder sie beschränkten die gesuchten ARG auf einige wenige mit klinischen Auswirkungen auf den Menschen (Sridhar et al.)⁸. In den wenigen Fällen, in denen die Autoren eine relevante Anzahl von ARGs untersuchten, war der Stichprobenumfang relativ gering (<100 Isolate) (Kolokotsa et al.)⁹. Aus diesem Grund sind die Ergebnisse dieses Projekts von großer Relevanz und bieten das Potenzial für neue Erkenntnisse in diesem Bereich.

Während des Projektzeitraums haben bakterielle genomweite Assoziationsstudien (GWAS) weiterhin die genetischen Grundlagen von Antibiotikaresistenz und Virulenz bei verschiedenen Bakterienarten aufgedeckt. Fortschritte bei den Sequenzierungstechnologien und Berechnungsmethoden haben die Möglichkeiten zur Durchführung dieser Studien trotz der inhärenten Komplexität bakterieller Genome erheblich verbessert¹⁰. Bei binären Merkmalen wurden logistische Regressionsmodelle verwendet, um die Assoziationen zwischen genetischen Varianten und phänotypischen Ergebnissen zu bewerten, während lineare Regressionsmodelle auf quantitative Merkmale wie minimale Hemmkonzentrationen (MICs) oder Scheibendiffusionsdurchmesser angewendet wurden. Um die Auswirkungen der Populationsstruktur - verursacht durch den horizontalen Gentransfer in bakteriellen Populationen - abzuschwächen, wurden Techniken wie die Hauptkomponentenanalyse (PCA)¹⁰⁻¹² und die multidimensionale Skalierung (MDS)^{10,13} eingesetzt. Diese Methoden helfen dabei, die Populationsstruktur und Verwandtschaftsbeziehungen zu berücksichtigen, indem die durch PCA oder MDS identifizierten Hauptpopulationsgruppen als Kovariaten verwendet werden. Es gibt jedoch noch weitere Herausforderungen, darunter die Notwendigkeit speziellerer Analyseinstrumente und standardisierter Methoden, um die Reproduzierbarkeit und Genauigkeit bakterieller GWAS zu verbessern.

Referenzen

1. Nascimento, J. A. S. et al. Molecular Epidemiology and Presence of Hybrid Pathogenic *Escherichia coli* among Isolates from Community-Acquired Urinary Tract Infection. *Microorganisms* 10, 302 (2022).
2. Mo, S. S. et al. *Escherichia coli* multilocus sequence type 38 from humans and broiler production represent distinct monophyletic groups. *Front. Microbiol.* 14, 1173287 (2023).
3. Falgenhauer, L. et al. Cross-border emergence of clonal lineages of ST38 *Escherichia coli* producing the OXA-48-like carbapenemase OXA-244 in Germany and Switzerland. *Int. J. Antimicrob. Agents* 56, 106157 (2020).
4. Patiño-Navarrete, R. et al. Specificities and Commonalities of Carbapenemase-Producing *Escherichia coli* Isolated in France from 2012 to 2015. *mSystems* 7, e01169-21 (2022).
5. Bokhary, H., Pangesti, K. N. A., Rashid, H., Abd El Ghany, M. & Hill-Cawthorne, G. A. Travel-Related Antimicrobial Resistance: A Systematic Review. *Trop. Med. Infect. Dis.* 6, 11 (2021).
6. Petersen, E., Shan Lee, S., Blumberg, L. & Levison, M. E. Antimicrobial resistance – A global problem in need of global solutions. *IJID Reg.* 9, 102–103 (2023).
7. Patjas, A., Martelius, A., Ollgren, J. & Kantele, A. International travel increases risk of urinary tract infection caused by extended-spectrum beta-lactamase-producing Enterobacterales —three-arm case-control study. *J. Travel Med.* 31, taad155 (2024).
8. Sridhar, S., Turbett, S. E., Harris, J. B. & LaRocque, R. C. Antimicrobial-resistant bacteria in international travelers. *Curr. Opin. Infect. Dis.* 34, 423–431 (2021).
9. Kolokotsa, A., Leotsinidis, M., Kalavrouziotis, I. & Sazakli, E. Effects of tourist flows on antibiotic resistance in wastewater of a Greek island. *J. Appl. Microbiol.* 130, 516–527 (2021).
10. Mosquera-Rendón, Jeanneth, et al. "Genome-Wide Association Studies (GWAS) Approaches for the Detection of Genetic Variants Associated with Antibiotic Resistance: A Systematic Review." *Microorganisms* 11.12 (2023): 2866.
11. Boeck, Lucas, et al. "Mycobacterium abscessus pathogenesis identified by phenogenomic analyses." *Nature Microbiology* 7.9 (2022): 1431-1441.
12. Saber, Morteza M., and B. Jesse Shapiro. "Benchmarking bacterial genome-wide association study methods using simulated genomes and phenotypes." *Microbial genomics* 6.3 (2020): e000337.
13. McDougall, Scott, et al. "Mechanisms of β -lactam resistance of *Streptococcus uberis* isolated from bovine mastitis cases." *Veterinary Microbiology* 242 (2020): 108592