

Mikrobielle Genomsequenzierung gewinnt an Bedeutung

Das Studienzentrum für Genomsequenzierung und -analyse im BfR stellt sich vor

Burkhard Malorny, Karsten Nöckler, Maria Borowiak, Laura Uelze, Andreas Hensel

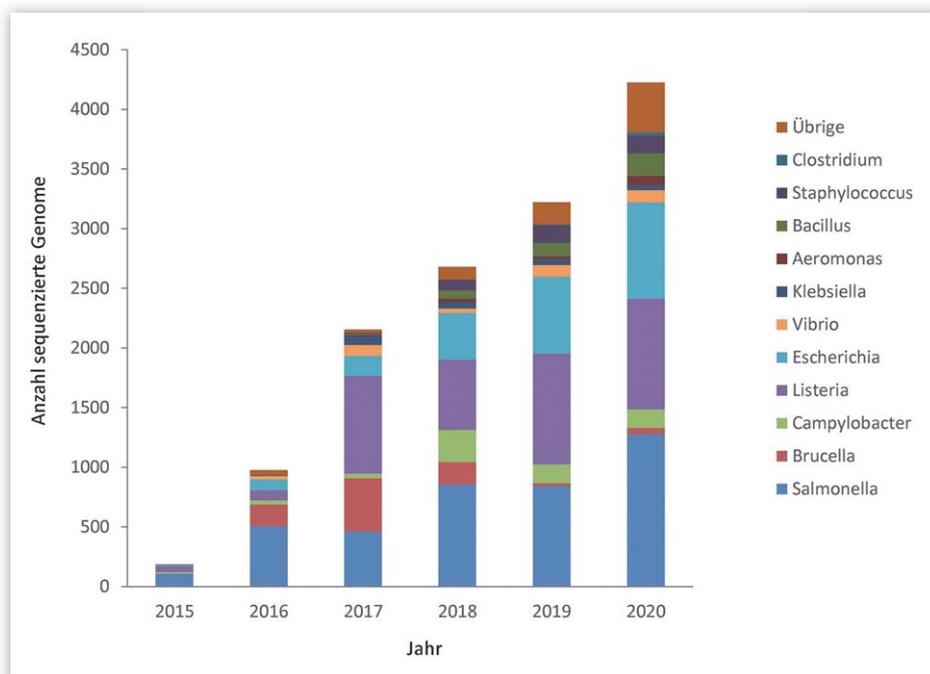


Abb. 1: Anzahl durchgeführter Genomsequenzierungen von bakteriellen Erregern im Studienzentrum für Genomsequenzierung und -analyse in Zusammenarbeit mit den Nationalen Referenz- und Konsiliarlaboratorien am BfR von 2015 bis 2020.

Das **Next-Generation-Sequencing (NGS)**¹ eröffnet der Mikrobiologie neue leistungsfähige Untersuchungsverfahren – und das nicht nur in der Forschung. Die Methode der Gesamtsequenzierung mikrobieller Genome findet mittlerweile auch im öffentlichen Gesundheitswesen, der Lebensmittelsicherheit und der Tiergesundheit Anwendung. Der Nutzen der molekularen Erregertypisierung reicht über die Identifizierung und Überwachung von Resistenz-, Virulenz- und Pathogenitätsfaktoren bis hin zur Analyse von komplexen Tier- und Umweltproben. Eine Herausforderung bei der Etablierung der Technologie in den Laboratorien ist die bioinformatische Verarbeitung der großen Datenmengen. Daher hat das Studienzentrum für Genomsequenzierung und -analyse am Bundesinstitut für Risikobewertung (BfR) die Aufgabe, NGS-Verfahren für die Nationalen Referenzlaboratorien in praktische Arbeitsabläufe umzusetzen und so Daten zu Erregereigenschaften und für die mikrobielle Risikobewertung bereitzustellen.

Die Notwendigkeit, ursächliche Erreger für Infektionskrankheiten in Mensch und Tier schnell zu detektieren und zu charakterisieren, die Infektionsquelle zu identifizieren und ihre Ausbreitung zu verhindern, führte in den letzten Jahrzehnten zur Entwicklung einer Vielzahl von immunologischen, biochemischen und molekularen Untersuchungsverfahren. Diese besitzen jedoch nicht das Auflösungsvermögen, das z. B. für die Zuordnung auf Stammebene und die Charakterisierung von Erregern notwendig ist, da nur ein Bruchteil der genetischen Information verfügbar wird. Das NGS hat durch die Möglichkeit, die Basenabfolge eines DNA-Moleküls oder ganzen Genoms eines Erregers zu bestimmen, das Potenzial, klassische Methoden zu ersetzen bzw. zu erweitern [1]. Es können z. B. spezifische Eigenschaften eines Erregers, wie Virulenz- und Resistenzdeterminanten, in einem Ansatz untersucht werden. Die Anwendung von NGS-Verfahren ist durch die stetige Verringerung der Kosten und die Entwicklung von einfach bedienbaren Sequenziergeräten in

den letzten Jahren auch für Behörden im Bereich der öffentlichen Gesundheit, Veterinärmedizin und Lebensmittelsicherheit sehr attraktiv geworden [2,3,4]. NGS-Verfahren z. B. für die genetische Charakterisierung von bakteriellen Isolaten, die in den regelmäßig durchgeführten nationalen Zoonosemonitoring- und Bekämpfungsprogrammen gewonnen werden, eröffnet den Nationalen Referenzlaboratorien (NRLs) des BfR die Möglichkeit, neue Erregervarianten zu erkennen, ihr pathogenes Potenzial abzuschätzen und ihre Ausbreitung zu überwachen.

Lösungen für zielgerichtete Fragestellungen

Das NGS ermöglicht es, effiziente, erregernunabhängige Arbeitsabläufe im Labor zu entwickeln. Das im BfR angesiedelte Studienzentrum für Genomsequenzierung und -analyse vereint die Expertise zu diversen mikrobiellen NGS-Verfahren sowohl für die experimentelle Erstellung als auch für die bioinformatische Auswertung von Genomsequenzen. Hierfür wurden Protokolle und Arbeitsabläufe im Labor entwickelt, um mikrobielle Nucleinsäuren in Reinform (genomische Isolat-DNA) oder aus komplexen DNA-Gemischen (Lebensmittel-, Tier-, Umweltproben) zu sequenzieren. Des Weiteren wurden bioinformatische Pipelines entwickelt, die eine schnelle und genaue Auswertung der Genomsequenzen vornehmen.

Für die genannten Aufgaben steht dem Studienzentrum eine Vielzahl von Sequenzierungsplattformen zur Verfügung. Es betreibt Geräte, die sowohl für eine Hochdurchsatzsequenzierung von Isolaten als auch von komplexen Metagenomproben eingesetzt werden können (Sequenzierer der zweiten Generation). Für die Erstellung von geschlossenen Referenzgenomsequenzen einschließlich mobiler genetischer Elemente (Plasmide) werden darüber hinaus Geräte benutzt, die lange Leselängen erzeugen können (Sequenzierer der dritten Generation). Im Jahr 2020 wurden durch die NRLs des BfR mehr als 4000 bakterielle Isolate verschiedenster Erreger für weitergehende Typisierungen sequenziert (Abb. 1) und über 50 Genome vollständig aufgeklärt.

Die im Studienzentrum entwickelten bioinformatischen Auswertungspipelines² ermöglichen

¹ **Next-Generation-Sequencing:** Hochdurchsatzverfahren zum Bestimmen der Nucleotidsequenz eines Genoms oder eines Teils davon.

² **Pipeline/bioinformatische Auswertungspipeline:** Programme, Skripte oder Softwaremodule, die aufeinander aufbauen. Das Ergebnis eines Programms wird als Input für eine nachfolgende Datenverarbeitung verwendet.

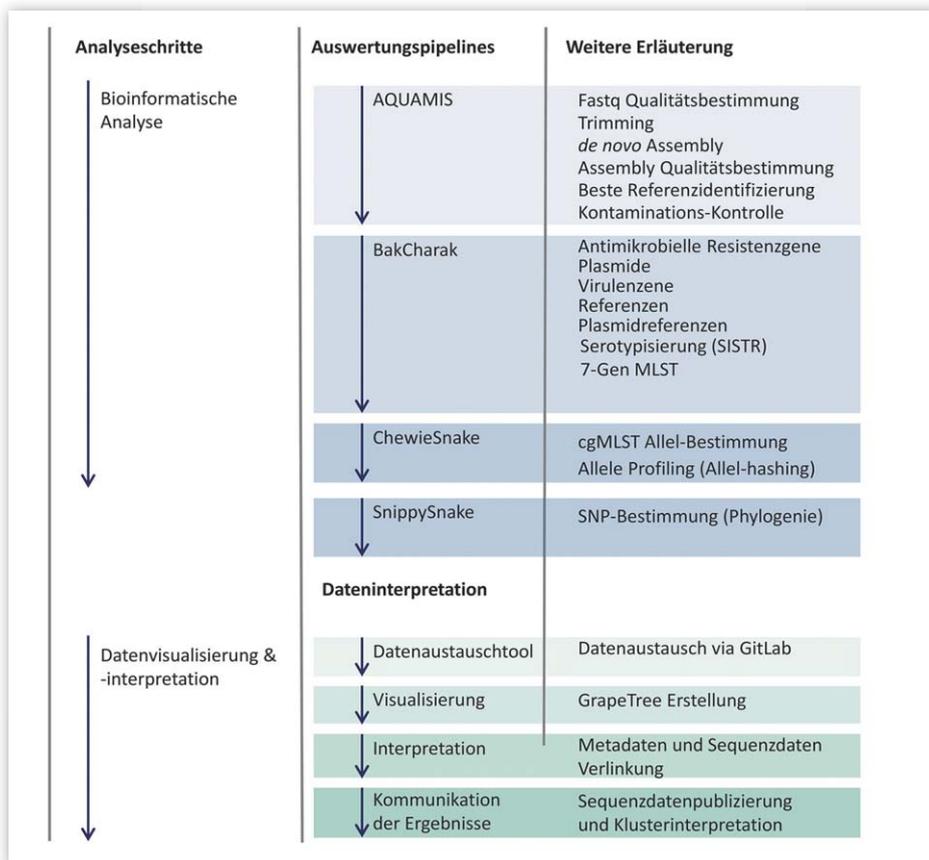


Abb. 2: Entwickelte und veröffentlichte bioinformatische Auswertungspipelines für die Qualitätsbestimmung, Typisierung und Charakterisierung von bakteriellen Genomsequenzen. Die Softwareprogramme sind frei zugänglich³. Die erzeugten Daten können einfach in Form von Reports elektronisch ausgetauscht und in Verbindung mit weiteren Metadaten interpretiert werden.

die automatische Analyse von Bakteriengenomen im Hochdurchsatz. Diese Pipelines sind frei verfügbar und können von interessierten Laboratorien genutzt werden (Abb. 2). Sie berechnen Qualitätskontrollparameter der generierten Sequenzen und charakterisieren genomische Bereiche, wie Resistenzgene und Virulenzgene. Zudem ermöglichen sie die Identifizierung von mobilen genetischen Elementen sowie Sero- und Multi-locus-Sequenztypen⁴. Für die Clustererkennung von eng verwandten Isolaten, z. B. bei der Aufklärung von Krankheitsausbrüchen, können hochauflösende cgMLST (engl.: core genome multi locus sequence typing) und SNP-Analysen (engl.: single nucleotide polymorphism) durchgeführt werden. Unterstützung bietet das Studienzentrum auch für die Etablierung der Technologie in den Untersuchungslaboratorien der Bundesländer durch die Ausrichtung von Workshops, individuellen Trainingskursen und Beratung.

Das Studienzentrum beteiligt sich an nationalen und internationalen Forschungsprojekten sowie an der Standardisierung der Methoden [5,

6]. Im nationalen GenoSalmSurv-Projekt (eng.: genome based surveillance of salmonellae) wird derzeit mit Partnereinrichtungen (Robert Koch-Institut und Bayerisches Landesamt für Gesundheit und Lebensmittelsicherheit) ein Modell für eine integrierte genombasierte Überwachung von Salmonellen auf Basis des Einsatzes von Genomsequenzierungsverfahren entwickelt [7]. Das Modell verfolgt einen dezentralen Ansatz, in dem jedes teilnehmende Labor Genomsequenzen der eigenen Isolate erzeugt und die Analyse zur Clustererkennung anschließend durch eine gemeinsam genutzte Software erfolgt.

Mit NGS das Pathogenitätspotenzial von Erregern besser verstehen

In Zusammenarbeit mit dem NRL für Salmonellen konnte das mit Schafinfektionen assoziierte *Salmonella enterica* subsp. *diarizonae* Serovar 61:k:1,5,(7) erstmalig genetisch umfassend analysiert werden. Das Serovar zeigt eine hohe Prävalenz in Schafherden weltweit. Gewöhnlich ver-

laufen Infektionen bei Schafen asymptomatisch, haben aber das Potenzial, Diarrhoe, Aborte oder chronisch proliferative Rhinitis zu verursachen. Die vollständige Charakterisierung des Genoms mit der long-read-Sequenzierung im Studienzentrum zeigte einen hohen Anteil an Pseudogenen⁵ im Vergleich zu eher ubiquitär vorkommenden Serovaren wie *Salmonella* Typhimurium. Es konnte nachgewiesen werden, dass sich das Serovar allmählich an das Schaf adaptiert hat (Abb. 3). Zur Virulenz des Serovars trägt höchstwahrscheinlich ein in den Bakterien stabil vorkommendes Plasmid bei, auf dem sich Gene für die Ausbildung eines Typ-IV-Sekretionsapparats⁶ und weitere Virulenzgene befinden [8].

In Zusammenarbeit mit dem Forschungszentrum für Infektionsmedizin an der Stiftung Tierärztliche Hochschule Hannover werden mit der dritten Generation der NGS-Technologie neue Arten der Gattung *Arcanobacterium* sequenziert [9] und charakterisiert [10]. Die Gattung besteht aus zehn Arten von fakultativ anaeroben, asporogenen, gram-positiven Bakterien, die in einer Vielzahl von Wirten vorkommen, z. B. Robben, Katzen, Hunden, Pferden und Menschen. Das Bakterium ist bisher in Zusammenhang mit Krankheiten wie Pharyngitis, Bakteriämie und Sepsis beim Menschen (*A. haemolyticum*), Mastitis bei Rindern (*A. pluranimalium*) oder Dermatitis bei Nerzen (*A. phocae*) beschrieben worden. Mit der Sequenzierung von *A. phocae*-Isolaten konnten Fälle von Dermatitis bei neun Nerzen über einen Zeitraum von mehr als 2 Jahren auf einer Farm in Finnland verfolgt werden. Die Sequenzanalyse der Genome ergab eine enge genetische Verwandtschaft der untersuchten Isolate, was auf einen dauerhaften Infektionsdruck durch einen bestimmten Klon bei den Nerzen hindeutet [10].

Identifizierung und Charakterisierung neuer Resistenzgene mittels NGS

Dass sich antimikrobielle Resistenzen in der Bakterienflora von Tier und Mensch entwickeln, bleibt weiterhin eine bedeutende Herausforderung. Besonders wichtig für die Risikobewertung ist hierbei, wie hoch das Potenzial von antimikrobiellen Resistenzgenen ist, zwischen natürlich vorkommenden, kommensalen Bakterien und pathogenen Mikroorganismen übertragen werden zu können. Wichtig dabei ist v. a. herauszufinden, ob Resistenzgene auf mobilen DNA-Elementen (z. B. Plasmiden oder Transposons⁷) lokalisiert sind. Auch dabei können NGS-

³ **AQUAMIS** (Qualitätskontrolle): https://gitlab.com/bfr_bioinformatics/AQUAMIS; **ChewieSnake** (Isolatvergleich und Ausbruchsanalyse): https://gitlab.com/bfr_bioinformatics/chewieSnake; **Bakcharak** (Typisierung und Charakterisierung): https://gitlab.com/bfr_bioinformatics/bakcharak; **SnippySnake** (Isolatvergleich über SNPs) https://gitlab.com/bfr_bioinformatics/snippy-snake

⁴ **Sequenztyp**: Numerische Bezeichnung für ein bestimmtes allelisches DNA-Sequenzprofil.

⁵ **Pseudogene**: DNA-Abschnitte, die wie ein Gen aufgebaut sind, jedoch nicht mehr als Vorlage für ein funktionales Protein dienen.

⁶ **Typ-IV-Sekretionsapparat**: In Bakterien hinterlegtes Ausscheidungssystem für Proteine oder DNA. Einige Typ-IV-Apparate können nach Kontakt mit Zielzellen eine Injektion der ausgeschiedenen Proteine in die Zielzelle bewirken.

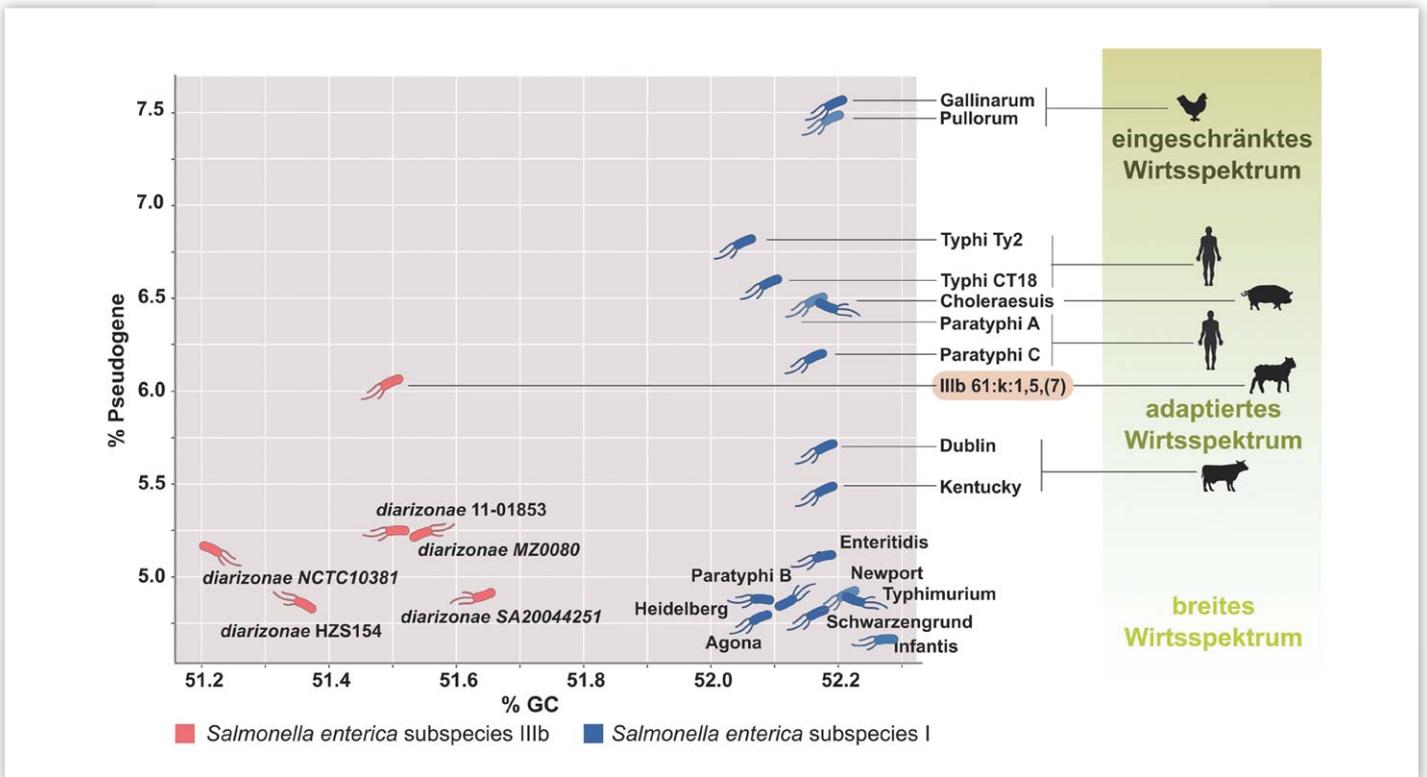


Abb. 3: Anteil von Pseudogenen in Genomen von *Salmonella* spp.-Serovaren. Es wird zwischen Serovaren, die ein breites, adaptiertes und eingeschränktes Wirtsspektrum besitzen, unterschieden. Das an das Schaf angepasste Serovar 61:k:1,5,(7) gehört entsprechend dem Anteil von Pseudogenen im Genom zu der Gruppe der adaptierten Serovaren.

Verfahren helfen, wie z. B. kürzlich für mobile Colistinresistenz-vermittelnde Gene gezeigt wurde. Colistin-resistente Enterobacteriaceen in der Nutztierhaltung waren in den vergangenen Jahren durch die horizontale Weitergabe der Colistinresistenz über das sogenannte *mcr-1*-Gen ein viel diskutiertes Thema [11]. Die Untersuchung von Colistin-resistenten *Salmonella*- und *E. coli*-Isolaten aus der Stammsammlung der NRL des BfR mittels Genomsequenzierung führte zur Beschreibung einer bisher unbekanntenen mobilen Colistin-Resistenzgen-Variante (*mcr-5*) [12]. Das Gen wurde sowohl in kommensalen *E. coli* als in *Salmonellen* aus der Geflügel- und Schweineproduktion nachgewiesen und trat erstmals 2011 in Deutschland auf [13, 14]. Mittlerweile ist das Gen weltweit in den unterschiedlichsten Proben und Mikroorganismen identifiziert worden [15]. Mithilfe der Genomsequenzierung konnten inzwischen auch mehrere verschiedene Übertragungsmechanismen von *mcr-5* aufgezeigt werden. Dieses Beispiel verdeutlicht die Vorteile von NGS-Verfahren bei der Untersuchung der genetischen Vielfalt von Resistenzgenen und die Komplexität ihrer Übertragungswege mittels mobiler genetischer Elemente.

Herausforderungen für den Einsatz des NGS in der Routine

Der Einsatz des NGS bei der Erforschung von Mikroorganismen bringt viele Vorteile, aber auch die Anwendungsmöglichkeiten in diag-

nostischen Routinelaboratorien sind zu berücksichtigen: NGS-Verfahren können z. B. in Monitoring- und Überwachungsprogrammen und insbesondere bei der Aufklärung und Prävention von Krankheitsausbrüchen eingesetzt werden [16]. Wenn NGS-Methoden für die Routinediagnostik angewendet werden, müssen im Sinne der Nachhaltigkeit verschiedene Anforderungen erfüllt werden. Der Umgang mit großen Datenmengen, die Bereitstellung von Rechenkapazität für die Analyse sowie Konzepte für den laborübergreifenden Austausch von Sequenzdaten müssen geschaffen werden. Die Standardisierung von NGS-Verfahren ist notwendig, um vergleichbare Daten zu generieren. Hier kommen zentrale oder dezentrale Sequenzierplattformen und Auswertungsmöglichkeiten als Lösungsansatz infrage. Das Studienzentrum skizziert und kommuniziert den Landeslaboratorien der Bundesländer dafür Ansätze, die in Workshops weiterführend diskutiert werden. Wir gehen davon aus, dass in den nächsten Jahren harmonisierte Prozesse ausgearbeitet werden, die eine effiziente Nutzung der NGS-Technologien in diagnostischen Laboratorien sektorenübergreifend ermöglichen.

Nicht zuletzt ist der Austausch von Sequenzdaten international für die Aufklärung von länderübergreifenden bzw. globalen Krankheitsausbrüchen entscheidend. Bestehende Konzepte hierfür wurden in der Vergangenheit von der Initiative „Global Microbial Identifier“ (www.globalmicrobialidentifier.org), der Europäischen Le-

bensmittelbehörde (EFSA) und dem Zentrum für Prävention und Kontrolle von Krankheiten (ECDC) erarbeitet und sehen den Aufbau einer internationalen Datenbank für den Austausch von Sequenzdaten vor [17, 18]. Diese Konzepte würden es den Landeslaboratorien und Bundesbehörden u. a. ermöglichen, Krankheitsausbrüche auch über Ländergrenzen hinaus effektiv zu erkennen, verdächtigen Infektionsquellen nachzugehen und geeignete Maßnahmen für den Gesundheits- und Verbraucherschutz zu ergreifen.

Anmerkungen

Wir danken PD Dr. Amir Abdulmajwood und seinen Mitarbeitenden (Institut für Lebensmittelqualität und Lebensmittelsicherheit der Stiftung Tierärztlichen Hochschule Hannover) für die Zusammenarbeit bezüglich der Entschlüsselung von *Arcanobacterium*-Genomen und den Partnern des GenoSalmSurv-Projekts, das durch das Bundesministerium für Gesundheit gefördert wird (Entscheidung ZMVI1-2518FSB709 vom 26.11.2018, aufgrund eines Beschlusses des Deutschen Bundestags durch die Bundesregierung).

Korrespondierender Autor

PD Dr. Burkhard Malorny



Abteilung Biologische Sicherheit, Bundesinstitut für Risikobewertung, Max-Dohrn-Str. 8-10, 10589 Berlin, burkhard.malorny@bfr.bund.de

⁷ **Transposon:** DNA-Abschnitt, der seine Position im DNA-Genom verändern kann (Transposition).