



Chimären

Was sie sind, wie sie entstehen und welche Probleme sie machen

Department of Biodiversity

Faculty of Biology, University of Duisburg Essen, Universitätsstr.5,45141 Essen, Germany

@biodiversityUDE

Wenn Sie an den Begriff Chimäre denken, assoziieren sie vermutlich direkt ein Mischwesen, vielleicht eine Sphinx, halb Mensch halb Löwe oder ein anderes Fabelwesen. Diese Assoziation ist verständlich, da der Term Chimäre tatsächlich seinen Ursprung in der (griechischen) Mythologie besitzt. Die Entstehung der Chimären steht hierbei nicht im Vordergrund. Wenn nun eine Chimäre, sagen wir halb Schnabeltier, halb Chamäleon inmitten einer Gruppe von Schnabeltieren, sowie Chamäleons auf einer Wiese stehen würde, wozu würden Sie die Chimäre zählen? Zu den Schnabeltieren, den Chamäleons, oder sie als eigene Lebensform beschreiben? Vermutlich ist ihnen hier ganz klar, dass es sich um eine eigene Lebensform handeln muss, die aus beiden Organismen hervorgegangen ist. Wir hätten also drei verschiedene Arten auf dieser Wiese. Warum ist dieser Gedankengang nun von Belang? Organismen werden schon lange nicht mehr alleinig anhand ihres Erscheinungsbildes identifiziert, sondern auch mithilfe von molekularen Markern und Referenzdatenbanken erkannt. Was nun sehr kryptisch klingt, ist im Prinzip relativ simpel. Kennt oder erkennt ein Wissenschaftler Lebewesen nicht, sei es, weil sie nicht mit optischen Verfahren zuzuordnen sind, oder weil das Wissen zur Zuordnung der Lebewesen fehlt, kann er die Erbinformation (DNS) der Organismen zur Identifikation nutzen. Mithilfe eines sogenannten Lyseverfahrens kann die DNS aus den Zellen freigesetzt werden.

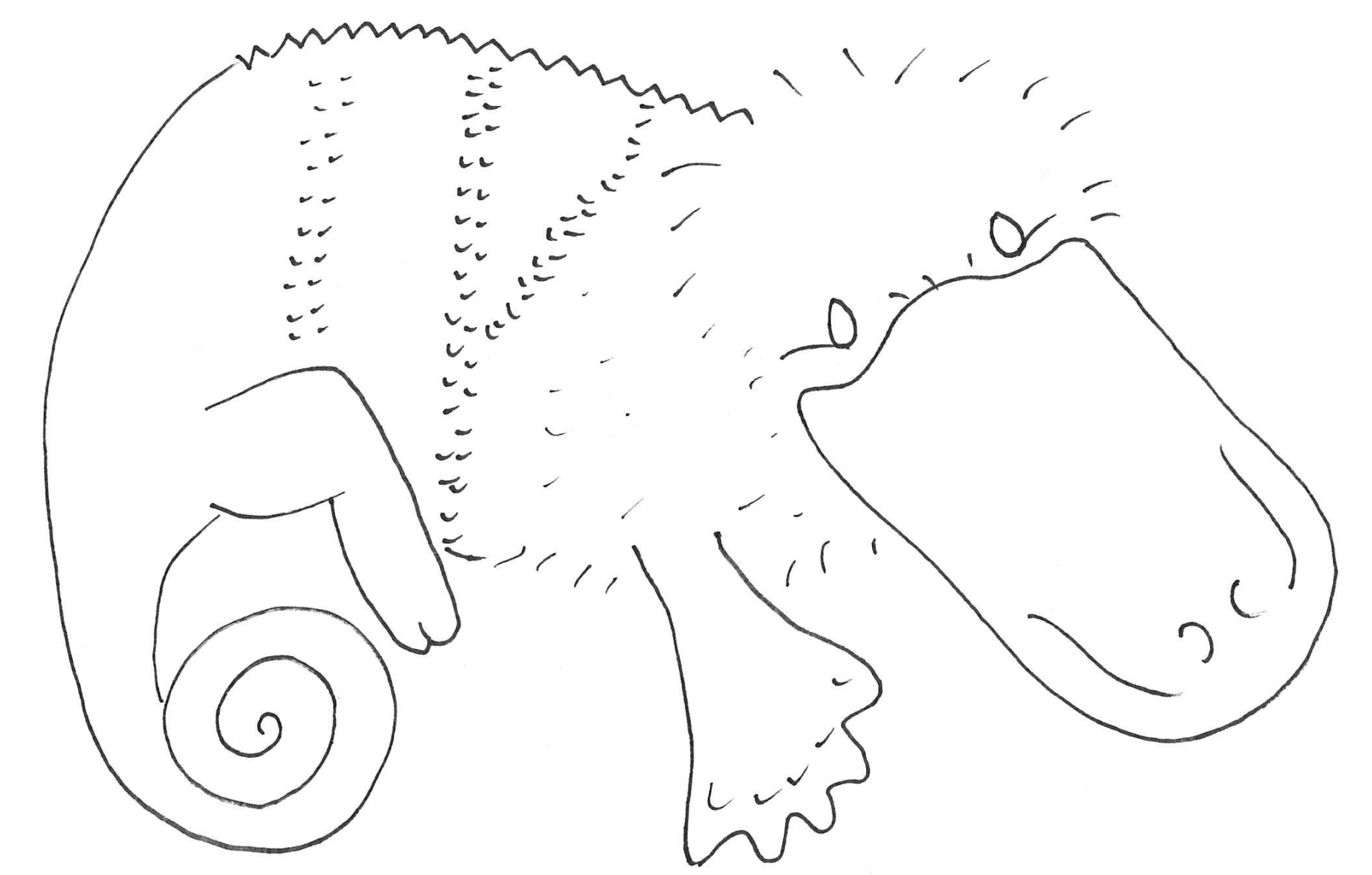


Fig.1 Beispiel einer Chimäre. Ein Bimer aus Schnabeltier und Chamäleon

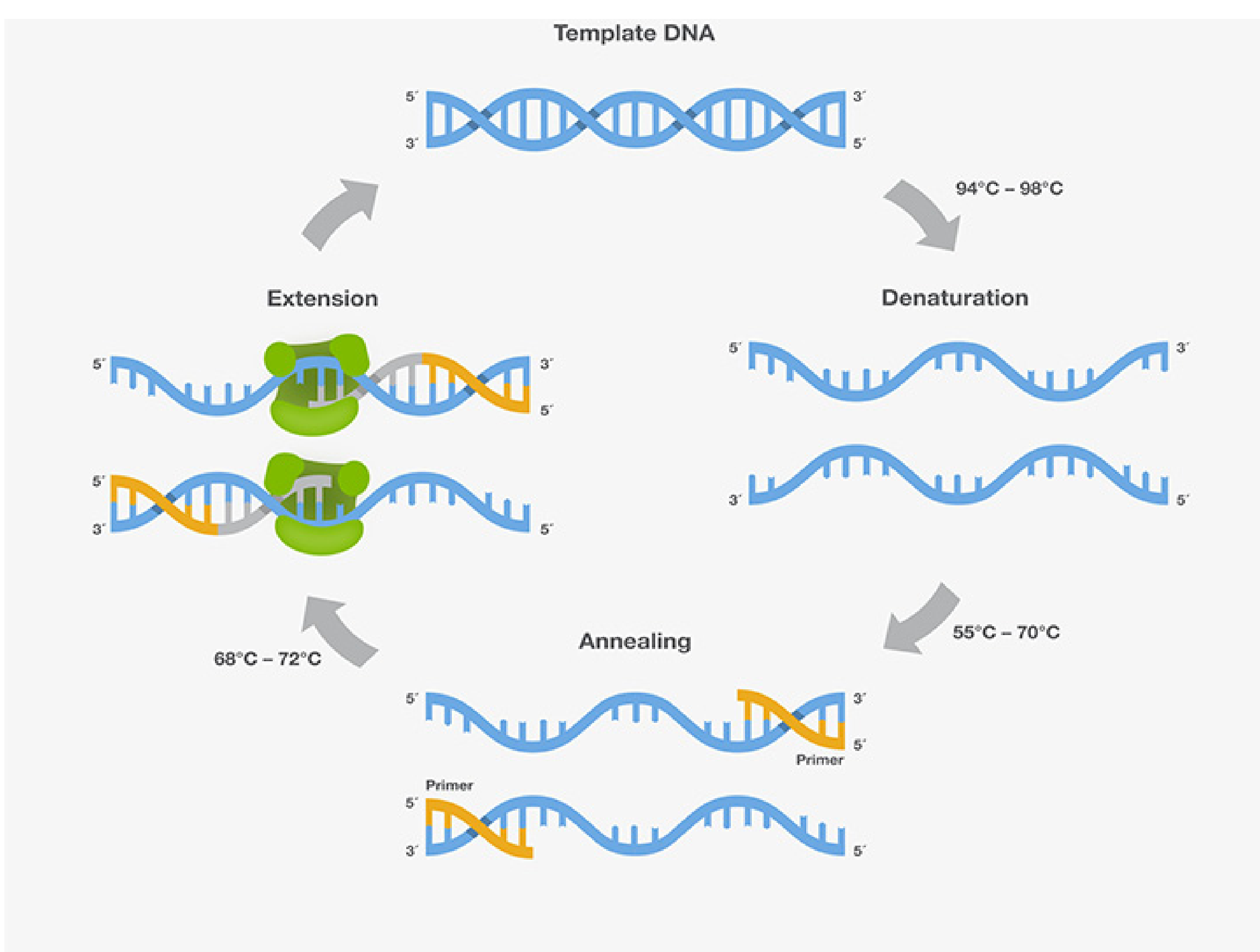


Fig. 2 Prozess einer PCR (Thermofischer)

Die freigesetzte DNS wird danach einem Prozess unterzogen, der sich Polymerase-Ketten-Reaktion nennt (Polymerase Chain Reaction (PCR)). Dabei begrenzen Primer, die genau vor und hinter der DNS Region von Interesse binden, die Amplifizierungsregion und die Polymerase fährt ab der Bindungsstelle der Primer die DNS ab und füllt sie komplementär auf. Dadurch entstehen viele sogenannte Amplicons. Mithilfe dieses Verfahrens ist es möglich eine Region in der DNS der Organismen zu vervielfältigen (amplifizieren). Diese Region sollte jeder Organismus der untersucht werden soll besitzen. Derzeit werden z.B. für mikrobielle Bestimmungen Abschnitte eines Gens genutzt, das für die kleine ribosomale Untereinheit kodiert, also einem Teil eines Ribosoms. Dies ist ein konservierter Bereich der kaum Mutationen unterliegt, der jedoch wiederum so unterschiedlich ist, dass Organismen unterschieden werden können. Gehen wir also davon aus, dass wir zwei Mikroorganismen in einer Wasserprobe haben, die wir bestimmen möchten. Sie werden molekular aufbereitet, wie zuvor beschrieben und nach der Übersetzung der Sequenzen in einen Buchstabencode (Sequenzieren) und einer Datenbankabfrage finden wir eine Sequenz die Organismus 1 zugeordnet wurde und eine die Organismus 2 zugeordnet wurde und ebenfalls eine die mit einer schlechten Übereinstimmung einem vollkommen anderem Organismus zugeordnet wurde.

Das Ergebnis ist hier das gleiche, wie im Beispiel mit dem Schnabeltier – Chamäleonhybriden. Wir haben jeweils 3 verschiedene Organismen. Die Frage ist nun sind diese Organismen echt? Im Bereich der SSU sind Chimärensequenzen eine Seltenheit. Sie können sich vorstellen, wenn dies häufiger geschieht, dass es zu einer Erhöhung der Diversitätsabschätzungen. Es liegt also eine Diversitätsüberbewertung vor und wohlmöglich besteht die Annahme, neue Arten seien gefunden worden. Dies kann vor allem beim Monitoring mittels molekularer Verfahren zum Problem werden.

Doch wie können solche Chimären nun überhaupt entstehen? Wenn die PCR der einzige Schritt in der molekularen Aufbereitung von Proben ist, in denen Chimären entstehen können, sind es vor allem die homologen (übereinstimmende) Regionen, die für deren Entstehung verantwortlich sind. Daher ist eine Quelle der Chimärenentstehung die unvollständige Primer Extension (Verlängerung).

Eine weitere bekannte Quelle ist das sogenannte Template Switching.

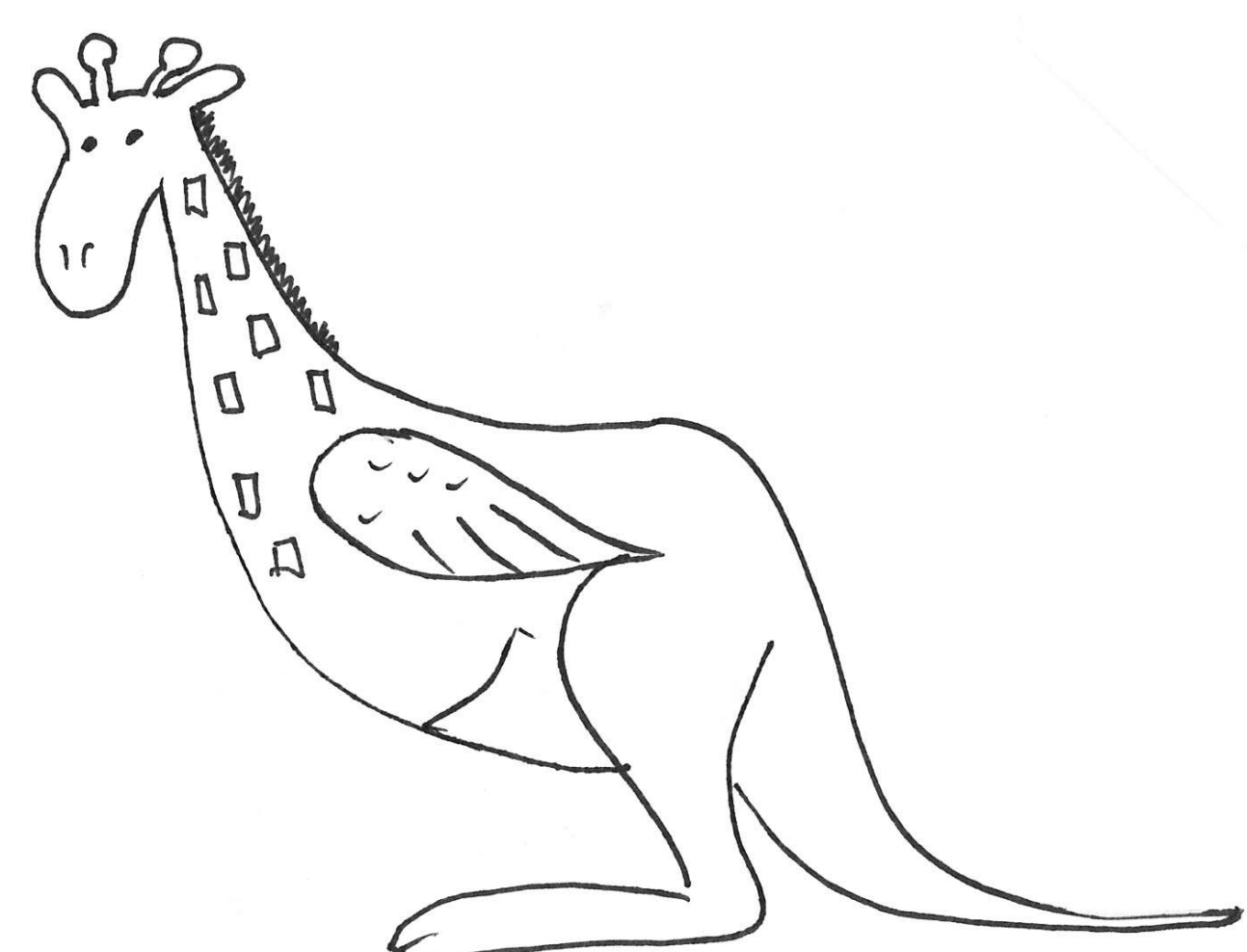


Fig. 4 Beispiel einer Multimere

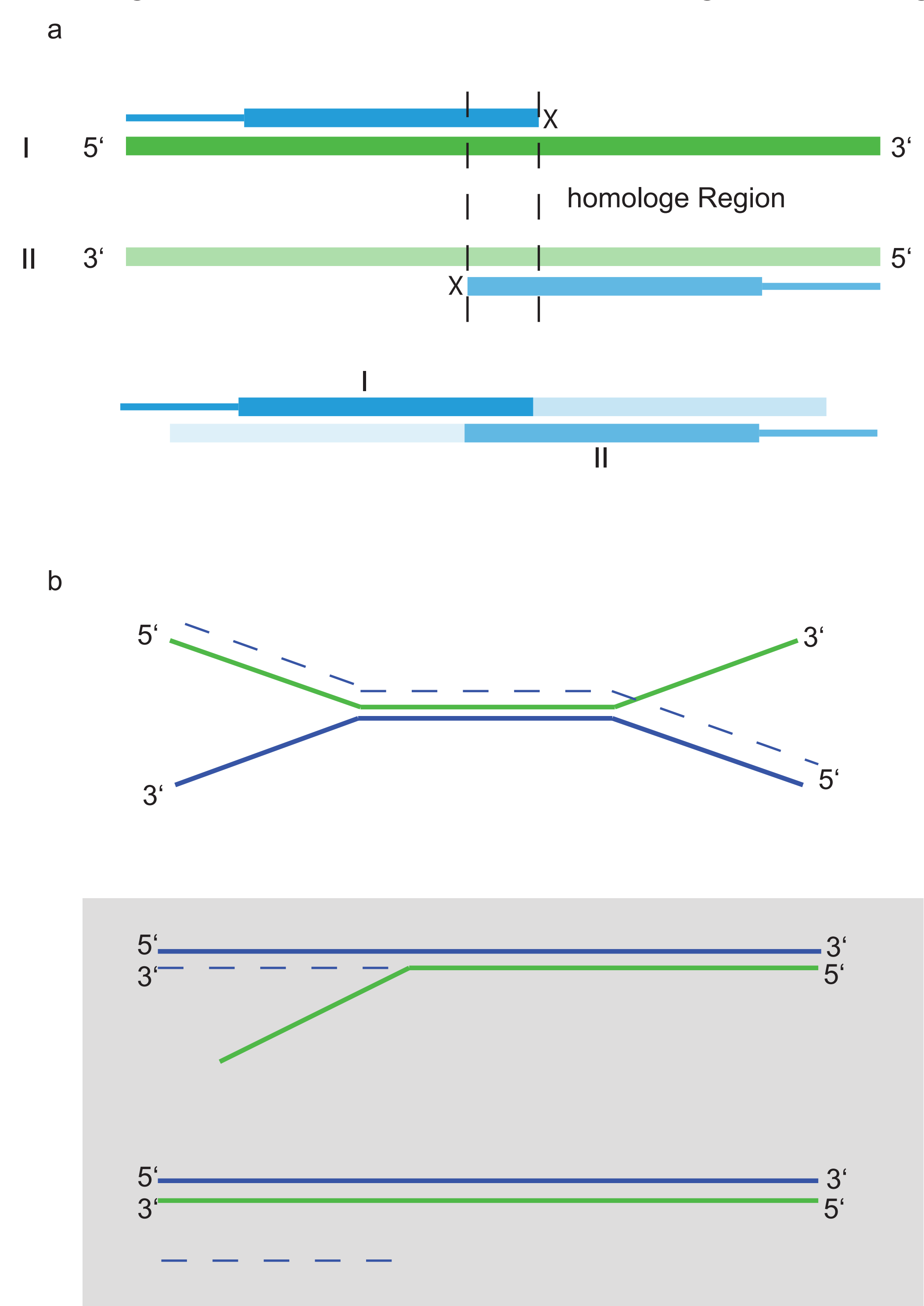


Fig.3 Chimerenentstehung durch a) Unvollständige Primerextension und b) Template Switching



Umgang mit Chimären in Amplicondatensätzen aus Umweltproben

Department of Biodiversity

Faculty of Biology, University of Duisburg Essen, Universitätsstr.5,45141 Essen, Germany

@biodiversityUDE

Arbeiten mit molekularen Datensätzen basiert viel auf Vertrauen. Vor allem in das Vertrauen in die Richtigkeit der Sequenzen, dass sie real sind und für Organismen stehen, die sich tatsächlich in der Probe, die man untersuchen möchte, tatsächlich befinden. Bioinformatische Pipelines mit eingebauten Filterschritten stärken das Vertrauen, da sie Reads mit schlechter Qualität herausfiltern, PCR Fehler, wie z.B. Chimeären detektieren sollen. Doch tun sie das wirklich? Je nach Fragestellung kann dies unterschiedlich wichtig sein, wenn jedoch die Biodiversität eines Systems im Vordergrund steht, ist sie vermutlich die Wichtigste. Anhand einer Untersuchung wichtiger Süßwasserbakterien aus der Gruppe der Betaproteobakterien, *Polynucleobacter*, fanden wir heraus, dass ein Anteil an OTUs nicht zu *Polynucleobacter* gehören kann, da die Sequenzen nicht mit den Referenzsequenzen clusterten aus den phylogenetischen Analysen.



Fig.1 In House Pipeline

Um alle Sequenzen auf Chimären zu untersuchen, die trotz Nutzung von UCHIME2 nicht herausgefiltert wurden nutzten wir folgende Methode. Wir nahmen die Sequenzen aus der Pipeline und schnitten sie jeweils 60% am 3' und 5' Ende, sodass wir pro OTU zwei Sequenzen erneut gegen die SILVA Datenbank blasten konnten.

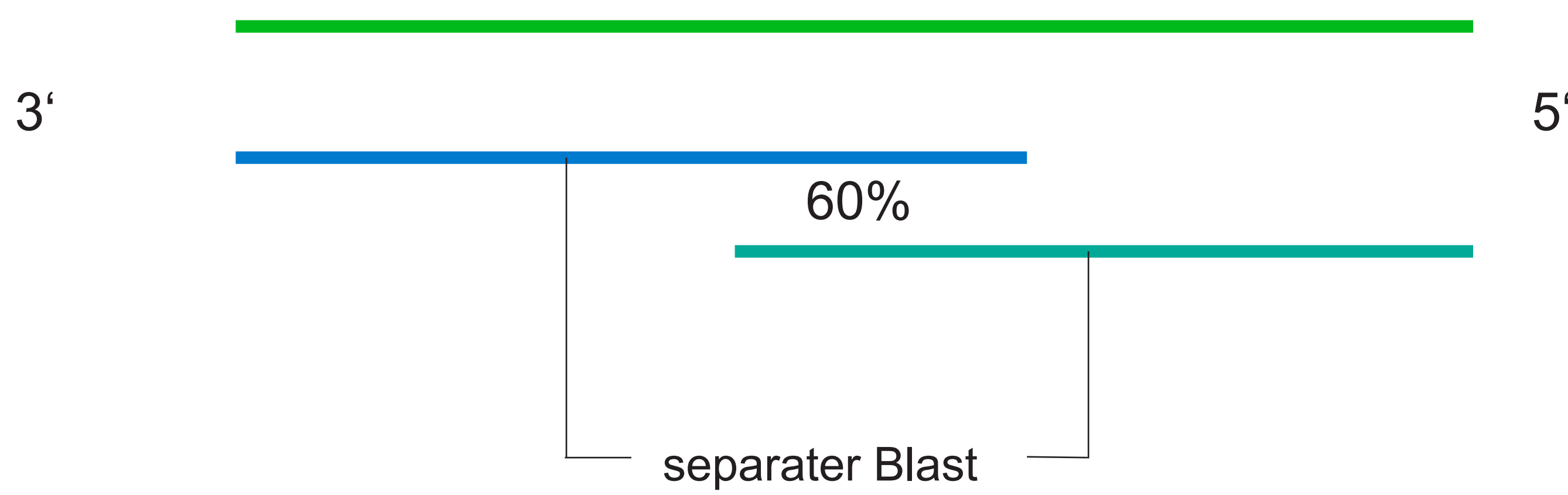


Fig.3 Strategie zur Untersuchung von Chimären

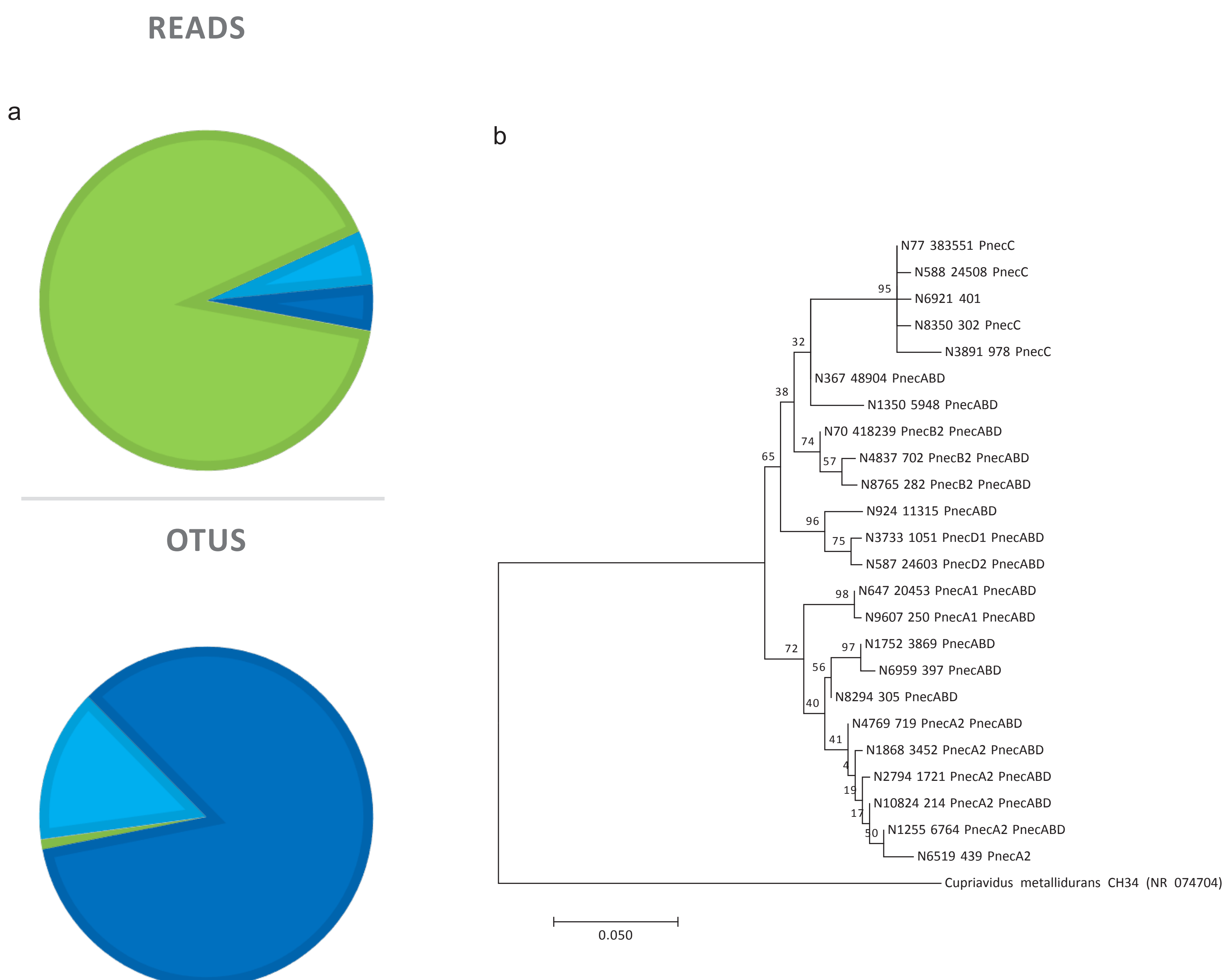


Fig.4 Read und OTU ANzahl nach Chimärencheck und OTU Endauswahl

Grüne Flächen zeigen den Anteil der nach dem Chimärencheck erhalten bleibt. Dunkelblaue Flächen zeigen den Anteil der nach dem ersten Filterschritt entfernt wurde und hellblau analog den Anteil der nach dem zweiten Filterschritt entfernt wurde

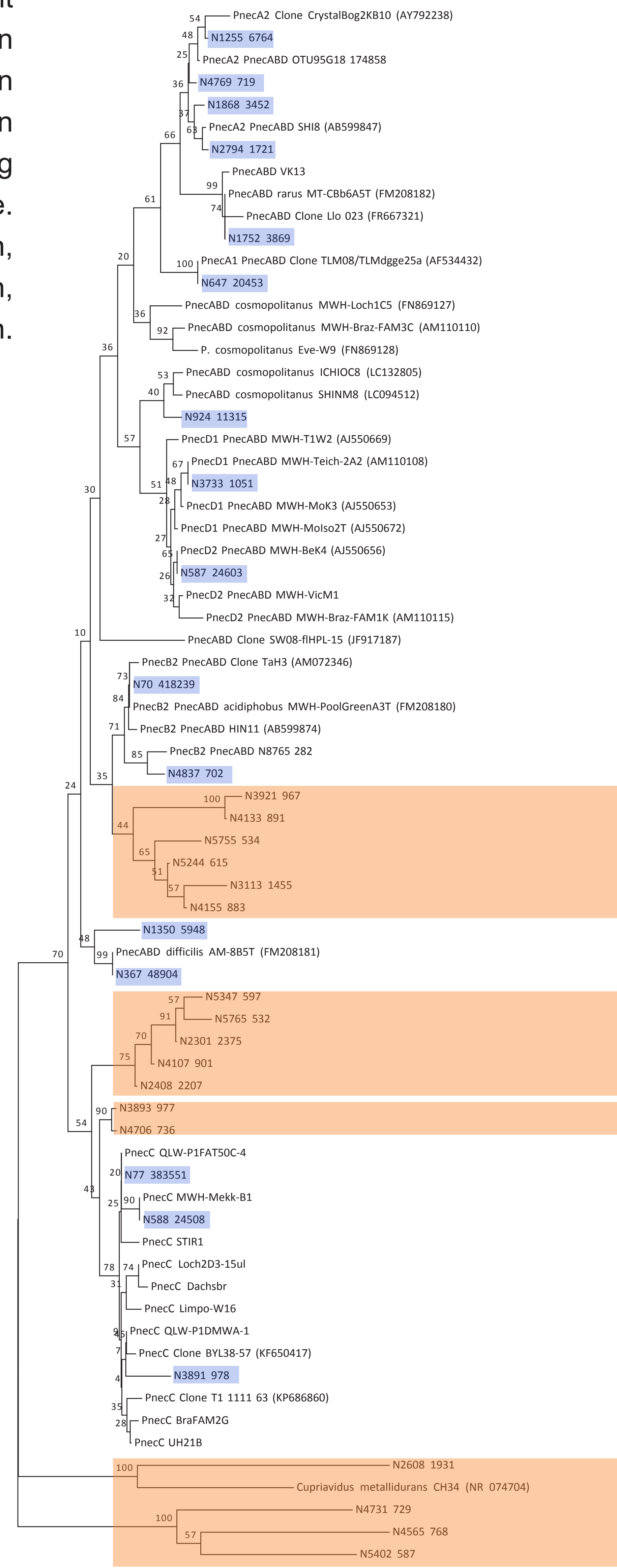


Fig.2 Clustering der OTUs mit Referenzsequenzen
OTUs mit über 500 Reads. Chimäre Sequenzen oder falsch zugeordnete Sequenzen aus der Datenbank sind orange hinterlegt. Nutzbare OTUS sind blau hinterlegt

NacheinemVergleichdergeschnittenenSequenzenstelltesichheraus, dass nur 4,4% der Reads betroffen waren. Allerdings machte das 85,4% der OTUs aus. Wir gingen in diesem Schritt von Bimeren aus. An dieser Stelle konnte gesichert werden, dass es sich um Betaproteobakterien handelte und definitiv um Burkholderiaceae. Anschließende phylogenetische Untersuchungen der geschnittenen Sequenzen zeigte jedoch zum Teil ungleiche Clusterungen mit den Referenzsequenzen. Eine Teilung der OTUs in insgesamt vier Untersequenzen ergab, dass sich ebenfalls Multimere unter den Chimären befanden. Diese Multimere bestanden jedoch aus Anteilen anderer Stämme dergleichen Art und sind daher nur wichtig zu detektieren, wenn die Artebene von Belang ist. Als Endresultat konnten wir für unsere Analysen 94,6 % der Reads nutzen, allerdings nur knapp 1% der OTUs waren für eine Untersuchung auf Artebene zuverlässig. Wir arbeitend derzeit an einer Pipeline, die noch zuverlässiger Chimären aus Amplicon-Datensätzen detektieren soll.