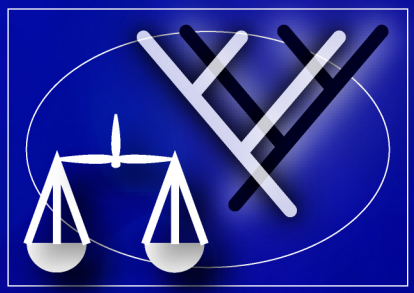

Merkmalsgewichtung in der Systematik

Zeichentheorie, Falsifikationismus und Phylogenetik

Dissertation

zur Erlangung des Doktorgrades

der Fakultät Biologie

der Universität Bielefeld

vorgelegt von

Lars Vogt

Bielefeld, Mai 2002

Merkmalsgewichtung in der Systematik

Zeichentheorie, Falsifikationismus und Phylogenetik

Dissertation
zur Erlangung des Doktorgrades
der Fakultät Biologie
der
Universität Bielelfeld

vorgelegt von
Lars Vogt

Bielefeld, Mai 2002

Besonderen Dank schulde ich dem Betreuer dieser Arbeit, Herrn Prof. Dr. Thomas Bartolomaeus, und der gesamten Arbeitsgruppe, die sich immer wieder mit Geduld und Diskussionsbereitschaft mit meinen Thesen auseinander gesetzt haben. Weiterhin möchte ich Dr. Wilko Ahlrichs und Christoph Bleidorn danken, in denen ich kritische Gesprächspartner gefunden habe. Björn Quast, meinem "Zimmergenossen", danke ich darüber hinaus für seinen freundschaftlichen Beistand. Prof Dr. H. Cruse danke ich sehr für die freundliche Übernahme des Korreferats.

Ein weiterer Dank geht an meine Familie, auf deren Rückhalt ich mich in jeder Situation verlassen konnte.

Das wissenschaftliche Projekt, in dessen Rahmen diese Dissertation möglich war, wurde gefördert durch die *Deutsche Forschungsgemeinschaft (Ba 1520/4-1).*

Inhaltsverzeichnis

Zusammenfassende Abhandlung

Einleitung

Klassifikationen

Dem Menschen scheint eine Systematisierung seiner Sinneseindrücke ein besonderes Anliegen zu sein. Dies mag über die notwendigen Bedingungen für ein reflektiertes Handeln begründet sein und somit auch über die Fähigkeit des Urteilens. Eine besondere Rolle in diesem Zusammenhang spielt aber sicherlich auch die besondere Fähigkeit des Menschen zur sprachlichen Kommunikation. Sprachliche Kommunikation bedarf einer begrifflichen Grundlegung, welche ebenso eine Systematisierung der Erscheinungen voraussetzt.

Sinneseindrücke können nur innerhalb eines in uns bereits bestehenden Rahmens, einer Art Schubladensystems, entstehen und unterliegen somit bereits einem ersten, uns nicht weiter bewussten, ordnenden Prinzip (z.B. durch verschiedene Sinneszelltypen und ihre nervöse Verschaltung, sowie ihre Repräsentation im Gehirn und vieles andere mehr – allesamt gehören sie zu den Bedingungen für die Möglichkeit von Erfahrung). Die hier angesprochenen Schubladen ließen sich als Erfahrungs-Kategorien ansprechen, durch die Beobachtungen erst möglich gemacht werden. Über diese Kategorien werden im weiteren dann Begriffs-Klassen gebildet und die Beobachtungen diesen zugeordnet. D.h. durch die Klassifizierung unserer Beobachtungen werden anhand von Begriffs-Klassen und den ihnen zugrundeliegenden Kategorien unsere Beobachtungen *von uns* geordnet und durch diesen Schritt auch erst für uns miteinander vergleichbar gemacht. So konstruieren wir uns eine Abbildung der materiellen Welt aufgrund der durch unseren spezifischen Sinnesapparat vermittelten Erfahrungen und *Eindrücke*. Eine solche Klassifizierung dient im weiteren als Grundlage für die Errichtung eines Ordnungs-Systems.

Welche Begriffsklassen nun zum Aufbau einer Systematik der belebten Natur herangezogen werden, liegt im Ermessen der Menschen. Dass dies immer wieder geschehen ist und auch zu den unterschiedlichsten Systemen geführt hat, zeigt die Geschichte. So wurde die belebte Natur sicherlich schon sehr früh in die Klasse der

essbaren und die der ungenießbaren Organismen systematisiert. Ein weiteres System wäre dasjenige zu Klassen von Heilpflanzen, von Giften und von solchen ohne erkennbare Wirkung. So ließe sich für jedes Anliegen ein eigenes System erstellen, wobei das jeweilige Anliegen auch als das ordnende Kriterium fungiert.

Es wird deutlich, dass es bei der Menge an Möglichkeiten einer Systematisierung einer allgemeinen Systematik bedarf, die diese verschiedenen Systeme beinhalten sollte (oder zumindest leicht zugänglich machen sollte) und die Funktion einer effektiven Wissensverwaltung übernehmen sollte. Ein solches System muss dem Anspruch genügen, alles Lebendige in sich erfassen zu können, und es sollte einen enzyklopädischen Charakter haben. Das System von Linné (Linné, 1758) ist ein solches System. Als ordnendes Prinzip wird ein Vergleich der Klassen bemüht in Bezug auf die Identitäten und Unterschiede der Erscheinungen ihrer Mitglieder. Dabei sollen die in der Hierarchie oben angeordneten Klassen (Elementarklassen: die Spezies) eine weitest gehende Ähnlichkeit ihrer jeweiligen Mitglieder aufweisen, und mit der Hierarchie absteigend sollen die größeren Klassen immer mehr Unterschiede zulassen. Darin begründet sich auch die Wahl des Binomens, bestehend aus dem Artnamen und dem Gattungsnamen, für die Bezeichnung einer Art – der Artname soll auf die Identität der Mitglieder dieser Klasse verweisen, wohingegen der Gattungsname den Unterschied der Mitglieder dieser Klasse zu denen der nächst-ähnlichen Klasse (andere Arten) aufzeigen soll. Die größeren Klassen stellen somit Vereinigungen der Elementarklassen, d.h. der Arten, dar.

Dieses System ist vor dem Hintergrund der Annahme einer natürlichen Hierarchie entworfen worden. Diese Hierarchie hat die menschliche Art *Homo sapiens* als Krönung der Schöpfung an die Spitze des Systems gestellt und alles Lebendige in Beziehung zu den Eigenschaften dieser Art bewertet und eingestuft (siehe auch Foucault, 1997). Eine solche Hierarchisierung diente u.a. als Legitimation des Systems als ein nicht vom Menschen willkürlich konstruiertes System, sondern eines, welches in der Natur eine Entsprechung finden sollte. Überbleibsel dieser Vorstellung finden sich noch heute im Vokabular vieler Biologen wieder und äußern sich in der Vorstellung einer evolutionären Leiter, die eine Zielgerichtetheit der Geschichte impliziert, sowie Ausdrücke wie „hochentwickelte“ und „primitive“ Lebensformen. Der Schöpfer wurde bemüht, da man sich keiner anderen Ursache für die beobachtbare Ordnung in der

belebten Natur gewiss werden konnte. Die Ordnung selber wurde zu diesem Zeitpunkt bereits, und wahrscheinlich schon weitaus früher, als empirisches Faktum angesehen.

Erklärung, Klassifikation und die Phylogenetische Systematik

Am Anfang einer jeden wissenschaftlichen Untersuchung steht ein *Problem*. Dieses *Problem* ist in der Regel unserer uns offenkundig gewordenen Unkenntnis geschuldet und erweckt in uns einen Erklärungsbedarf. Auf eine solche Unkenntnis werden wir aufmerksam gemacht, indem wir überraschende Beobachtungen/Erfahrungen machen, die im Widerspruch zu unserer bisherigen Sicht der Dinge stehen. Tritt ein solcher Fall auf, so fehlt es uns offensichtlich an geeigneten Erklärungsmodellen, um diese Befunde in ein in sich kohärentes Erfahrungsgebäude zu integrieren.
Das System von Linné hat einen solchen Erklärungsansatz für das beobachtbare Muster von Organismen, ihrer räumlichen Verteilung und ihren untereinander vergleichbaren und einander ähnlichen Eigenschaften geliefert. Nur musste für dessen Begründung eine empirisch nicht systematisch zu untersuchende Entität angenommen werden, die metaphysische Konzeption eines Schöpfers, welcher für die überraschende Ordnung verantwortlich gemacht worden ist. Mit dem Aufkommen der Evolutionstheorie von C. Darwin und der Synthetischen Theorie (auch *Modern Synthesis*) und ihrer steigenden Akzeptanz innerhalb der biologischen Wissenschaften, trat ein zum Schöpfer alternatives, wissenschaftlich zufriedenstellenderes Erklärungsmodell für die Ursache des Vorhandenseins einer Ordnung in der belebten Natur auf den Plan.

Fragt man sich nun jedoch, warum das tatsächliche Muster identischer und unterschiedlicher organismischer Eigenschaften genauso beschaffen ist, wie es für uns beobachtbar ist und nicht anders, so liefert die Phylogenetik die passenden Erklärungen. Sie stellt Hypothesen über singuläre geschichtliche Ereignisse auf, indem sie einen Vergleich der beobachtbaren Eigenschaften der betreffenden Organismen und ihrer Verteilung über die gesamten Organismen hinweg durchführt. Dieser Vergleich liefert die empirische Basis, welche nun unter Berücksichtigung der Folgerungen aus der Evolutionstheorie, generellen evolutionären Mechanismen, wie sie u.a. in der Populationsgenetik und der molekularen Biologie beschrieben werden, und in Hinblick auf die Aufstellung und Begründung erklärender phylogenetischer Hypothesen, interpretiert wird.

Auf diese Weise lassen sich jedoch nicht alle organismischen Eigenschaften erklären. Lediglich solche Eigenschaften, die von Generation zu Generation vererbt werden, sind potentiell offen für phylogenetische Erklärungen. Dazu gehören eine Vielzahl anatomischer Merkmale, DNA Sequenzen, Verhaltensmerkmale und viele andere mehr.

Beschäftigt man sich nun mit der Rekonstruktion der verwandtschaftlichen Beziehungen einer konkreten Gruppe von Organismen, so treten alsbald praktische Schwierigkeiten bei der Interpretation/Auswertung der Merkmale auf. Zum Beispiel bei der Verwendung von 18S rDNA Sequenzen fällt auf, dass, obwohl es sich bei der Nukleotidsequenz um eine geerbte Eigenschaft handelt, nicht alle Bestandteile des vorzufindenden Verteilungsmusters von identischen und unterschiedlichen Nukleotiden mit einer phylogenetischen Hypothese zu erklären sind. Es sind auch Übereinstimmung zwischen Vertretern verschiedener Arten zu finden, welche nicht auf Homologie (d.h. einen gemeinsamen und über Vererbung übertragenen Ursprung des Merkmals), sondern auf Homoplasie (d.h. einen nicht gemeinsamen Ursprung des Merkmals, verursacht durch Konvergenz oder Rückmutation), zurückzuführen sein müssen. Daraus ergibt sich ein methodisches Problem, da entschieden werden muss, welche Bestandteile des vorgefundenen Musters durch Homologie und welche durch Homoplasie erklärt werden sollten. Anders ausgedrückt: Es gibt eine Vielzahl von möglichen erklärenden Hypothesen zu ein und demselben Phänomen und alle stehen nicht im Widerspruch zur allgemeinen Evolutionstheorie oder dem zu beobachtenden Verteilungsmuster. Es stellt sich also die Frage, welche Identitäten gehen auf einen gemeinsamen geschichtlichen Ursprung zurück und welche haben eine andere Ursache. Dies lässt sich nur durch den Vergleich der Hypothesen untereinander und der Berücksichtigung ihrer jeweiligen Erklärungskraft klären, und damit handelt es sich um eine Frage, die nicht mehr allein empirisch zu lösen. Sie stellt eine methodologische Frage dar - die Frage nach der *Gewichtung von Merkmalen*.

Morphologie versus molekulare Daten

Bis zum Aufkommen der modernen molekularbiologischen Methoden, wie z.B. dem PCR-Verfahren, galten anatomische organismische Eigenschaften als die empirische Basis phylogenetischer Untersuchungen schlechthin, und die ersten streng phylogenetischen Untersuchungen wurden anhand morphologischer Befunde nach der

Methode von W. Hennig (1966) durchgeführt. Nach dieser Methode wird zu einer bestehenden Gruppe/Art die zugehörige Schwestergruppe (Adelphotaxon) gesucht, die sich dadurch auszuzeichnen hat, dass sie Merkmale aufweist, die nur ihre Vertreter und die Vertreter der betreffenden Gruppe/Art besitzen. Diese Merkmale werden Synapomorphien genannt. Synapomorphien sind es also, die innerhalb dieser Methode dazu verwendet werden, Schwestergruppenverhältnisse aufzudecken. Aus Schwesterngruppenverhältnissen wiederum lassen sich die Monophylien der neu entstandenen Gruppen ableiten. Nach dieser Methode kommen phylogenetische Untersuchungen der Suche nach Synapomorphien gleich. Verwandtschaftsbeziehungen werden dabei offensichtlich Schritt für Schritt rekonstruiert (sog. ‚*clustering*' Verfahren) – nicht alle möglichen Verwandtschaftsbeziehungen werden während der Bewertung der insgesamt vorhandenen empirischen Befunde berücksichtigt, da mit der ersten gefundenen Schwestergruppenbeziehung eine Vielzahl von Möglichkeiten für den weiteren Verlauf der Untersuchung ausgeschlossen werden. Das hat jedoch zur Folge, dass, wie auch andere ‚*clustering*'-Verfahren, die klassische hennigsche Methode den Nachteil besitzt, dass sie nicht zwingend die optimale Hypothese für die vorhandenen Daten liefert, sondern während der Analyse in lokalen Optima gefangen werden kann (zu ‚*clustering*'-Verfahren siehe z.B. Page und Holmes, 1998). Dennoch galt die klassische hennigsche Methode eine lange Zeit als der unter PhylogenetikerInnen bevorzugte methodische Ansatz, und eine Vielzahl von phylogenetischen Hypothesen und damit verbundenen Klassifikationen sind auf diesem Weg entstanden (siehe u.a. Brusca und Brusca, 1990).

Durch die beständige Publikation neuer Ergebnisse morphologischer Untersuchungen und insbesondere durch das Aufkommen molekularer Sequenzdaten war die hennigsche Vorgehensweise aufgrund der Menge der zu berücksichtigenden Daten nicht länger praktikabel. Durch die zu diesem Zeitpunkt inzwischen verfügbaren Rechnerleistungen war die Möglichkeit von computerunterstützten Analysen gegeben, und eine Vielzahl von Algorithmen und statistischen Verfahren mit der dazugehörigen Software sind seit dem entwickelt worden (einen guten und aktuellen Überblick liefert J. Felsensteins Homepage „Phylogeny Programs", http://evolution.genetics.washington.edu/phylip/software.html). Diese neuen Verfahren sind überwiegend von molekular arbeitenden ForscherInnen genutzt worden. Im Zuge dieser Entwicklung sind einige der morphologisch begründeten Verwandtschaftsbeziehungen aufgrund der molekularen Befunde revidiert und der überwiegende Teil zusätzlich bestätigt worden. In diesem

Zusammenhang sind auch ein Großteil der Ergebnisse der Analysen der 18S rDNA Sequenzen von Polychaeten zu sehen (siehe Artikel von C. Bleidorn, L. Vogt und T. Bartolomaeus: *A contribution to sedentary polychaete phylogeny using 18S rDNA sequence data* und *New insights into polychaete phylogeny (Annelida) inferred from 18S rDNA sequences*, und das Poster von C. Bleidorn, L. Vogt, N. Arnold und T. Bartolomaeus: *Zur Phylogenie der Polychaeta anhand von 18S rDNA Sequenzen*).

Revidierungen werden immer dann als verhältnismäßig unproblematisch angesehen, solange die morphologischen Befunde, die zur Begründung der bisherigen Verwandtschaftssysteme herangezogen worden sind, von den ForscherInnen als unbefriedigend oder zumindest nicht überzeugend gewertet worden sind (siehe Poster von C. Bleidorn, L. Vogt und T. Bartolomaeus: *Travisia ist ein Scalibregmatidae, kein Opheliidae (Annelida, Polychaeta)*).

Stehen jedoch die morphologischen Befunde im Widerspruch zu den molekularen, und beide Arten von Daten werden von den WissenschaftlerInnen als phylogenetisch aussagekräftig gewertet, wird die zur Zeit ungelöste methodologische Problematik offenkundig. Als Paradigma eines solchen Falls kann die Frage nach der Monophylie der Articulata verstanden werden. Die Articulata werden als monophyletische Gruppe in vielen Lehrbüchern angeführt und stellen ein Taxon dar, das auf eine, innerhalb einer von der Morphologie geprägten Systematik, lange Tradition zurückblicken kann (siehe z.B. Brusca und Brusca, 1990; Nielsen, 1995; Ax, 1999). Molekulare Untersuchungen von Vertretern der Articulata widersprechen jedoch einer solchen Auffassung und unterstützen vielmehr eine Gruppe, die sich morphologisch durch die Eigenschaft der Häutungsfähigkeit beschreiben ließe (Aguinaldo et al., 1997). Der aktuelle Verlauf der Diskussion um diese Verwandtschaftsbeziehungen macht deutlich, dass sich diese Frage nicht einfach durch das Hinzufügen neuer empirischer Befunde lösen lässt (z.B. Schmidt-Rhaesa et al., 1998; Giribet und Wheeler, 1999; Wägele et al., 1999; Manuel et al., 2000; Wägele und Misof, 2001; Zrzavý, 2001), sondern vielmehr ein methodologisches Problem darstellt.

Darüber hinaus hat das Auftreten der molekularen Daten eine Diskussion entfacht, wie man mit der Kombination morphologischer und molekularer Daten bei der phylogenetischen Rekonstruktion umgehen solle. Sollte man die Daten partitionieren und getrennt untersuchen oder eine kombinierte Analyse durchführen – und wenn kombiniert, sollten dann morphologische und molekulare Merkmale unterschiedlich gewertet werden (z.B. Eernisse und Kluge, 1993; Chippindale und Wiens, 1994; Hedges

und Maxson, 1996; Page, 1996 ; Lee, 1997; Littlewood et al., 1997; Wiens, 1998; Ballard et al., 1998; Kluge, 1998; Wiley et al., 1998; Flook et al., 1999)?
Es bedarf folglich einer Theorie und Methode der Bewertung des phylogenetischen Informationsgehalts von Merkmalen, um im Falle einer in sich widersprüchlichen Interpretation der vorhandenen empirischen Befunde eine Entscheidung fällen zu können. Der phylogenetische Informationsgehalt eines Merkmals wird dabei durch einen ihm zugewiesenen Wert ausgedrückt, den man das Gewicht des Merkmals nennt.

Das Gewichten von Merkmalen in der Systematik

Die Frage nach der Gewichtung morphologischer gegenüber molekularer Merkmale stellt ein spezielles Problem einer weitaus allgemeineren Fragestellung dar: Wie kann der phylogenetische Informationsgehalt von Merkmalen gemessen werden – oder anders ausgedrückt: Wie kann man die empirischen Befunde im Hinblick auf ihren Beitrag zur Rechtfertigung einer Apomorphie-Hypothese quantifizieren? Es sind die verschiedensten Vorschläge und Meinungen zum Themengebiet der Gewichtung von Merkmalen geäußert (z.B. Neff, 1986; Wheeler, 1986; Bryant, 1989; Chippindale und Wiens, 1994; Allard und Carpenter, 1996; Milinkovitch et al, 1996; Haszprunar, 1998; Källersjö et al., 1999; Wenzel und Siddall, 1999; O'Keefe und Wagner, 2001; Wiens, 2001) und auch einige konkrete Gewichtungssysteme vorgestellt worden:
Die bisher gängigen Gewichtungsverfahren lassen sich grob in *a priori* und *a posteriori* Verfahren gliedern (nach Kitching et al., 1998). Den *a posteriori* Verfahren liegen Überlegungen zur Optimierung von Merkmalsverteilungen anhand von aus der kladistischen Analyse (Verwandtschaftsanalyse) berechneten Parametern zugrunde. Zu ihnen gehören unter anderem das *successive approximations weighting* (Farris, 1969; Carpenter, 1988), das *reverse successive weighting* (Trueman, 1998) und das *implied weighting* (Goloboff, 1993, 1995). Den *a priori* Verfahren liegen meist Annahmen über unterschiedliche Evolutionsprozesse zugrunde (Rieppel, 1999). Zu ihnen gehören beispielsweise die Gleichgewichtung aller Merkmale (Kluge, 1997a), das unterschiedliche Gewichten von verschiedenen Kodonpositionen bei Sequenzdaten (z.B. Björklund, 1999), so wie die unterschiedliche Gewichtung von Transversionen gegenüber Transitionen (z.B. Broughton et al., 2000). Es wird auch versucht, morphologische Merkmale auf der Grundlage von Prozesswahrscheinlichkeiten zu bewerten (Lewis, 2001). Weitere Ansätze versuchen den Schwierigkeiten, die bei der

Erstellung der positionalen Korrespondenz (positionale Homologiehypothese, bzw. primäre Homologie nach de Pinna, 1991) auftreten können und sich bei Sequenzdaten dadurch ausdrücken, dass bestimmte Stellen im Alignment sich nicht eindeutig alignieren lassen, durch eine differentielle Gewichtung Rechnung zu tragen (z.B. Lutzoni et al., 2000).

Maximum Likelihood oder Maximum Parsimonie

Die Wahl eines bestimmten Gewichtungsansatzes hat weitgehende Konsequenzen für die Vorgehensweise innerhalb der kladistischen Analyse. Es existieren verschiedene Methoden und Verfahren zur kladistischen Analyse, wie z.B. Maximum Parsimonie, Maximum Likelihood und Distanz Verfahren (einen Überblick zu den verschiedenen Verfahren liefern Hillis et al., 1996; Wägele, 2000), die jeweils unterschiedliche Gewichtungssysteme und –ansätze umsetzen. Insbesondere zwischen Verfechtern der Maximum Parsimonie und der Maximum Likelihood Methode ist eine Diskussion um die Verlässlichkeit und Begründung der Wahl der jeweiligen Methode entbrannt. Dabei beziehen sich die Protagonisten der Parsimonie Methode auf erkenntnistheoretische und wissenschaftsphilosophische Argumente und Modelle, um ihre Position zu begründen und zu rechtfertigen, wobei sie sich insbesondere auf den popperschen Falsifikationismus berufen (z.B. Kluge, 1997, 1997a, 1998; Siddall and Kluge, 1997; Farris et al., 2001). Während sich die Verfechter des Maximum Likelihood Ansatzes auf empirische Studien und computergestützte Simulationen beziehen (Felsenstein, 1978; Huelsenbeck und Hillis, 1993; Huelsenbeck, 1995; Yang, 1996; Sullivan und Swofford, 2001; Swofford et al., 2001). Diese zwei konträren Positionen lassen sich auch als Struktur oder Muster basierte Ansätze gegenüber Prozess basierten Ansätzen charakterisieren.

Zu den Befürwortern der Maximum Parsimonie Methode gehören auch die sogenannten *Pattern Cladisten*, die, unter Bezugnahme auf Poppers Falsifikationismus, u.a. drei Thesen vertreten (z.B. Kluge, 1997, 1997a):

1) der Kongruenz-Test ist der entscheidende Schritt der phylogenetischen Analyse, da nur hier die Möglichkeit einer Falsifikation von phylogenetischen Hypothesen besteht

2) das angenommene Hintergrundwissen soll minimal gehalten werden
3) Annahmen über evolutionäre Prozeßwahrscheinlichkeiten sollen nicht in die Analyse einfließen, da sie zusätzliches Hintergrundwissen darstellen

Wobei sie zu dem Schluss kommen, dass nur die Maximum Parsimonie Methode konsistent zu diesen Forderungen ist und dass die empirischen Befunde nicht gewichtet, bzw. nur gleichgewichtet, in die Analyse eingehen dürfen.
Dieser Position habe ich mich insbesondere in den Artikeln *Testing and Weighting Characters*, *Weighting Indels as Phylogenetic Markers of 18S rDNA Sequences in Diptera and Strepsiptera* und *Process Probabilities and the Weighting of Characters in Systematics – Following the falsificationist program of phylogenetic research*, sowie dem Poster *Zur Logik des Gewichtens phylogenetischer Merkmale* gewidmet. Dies soll im Folgenden erläutert werden.

Die Rolle des Homologie- und Apomorphie-Konzepts in der Rekonstruktion von Phylogenien

Dem Homologie-Konzept kommt in der phylogenetischen Forschung eine zentrale Rolle zu. Umso verwunderlicher ist die Tatsache, dass es nicht *das* Homologie-Konzept gibt, sondern eine Vielzahl verschiedener Konzepte unter ein und demselben Etikett koexistieren.

Der Begriff *Homologie* ist ursprünglich von R. Owen (1843) in einer prä-darwinschen und prä-mendelischen Ära geprägt worden (Butler, 2000). Owen verstand unter einer Homologie *„the same organ in different animals under every variety of form and function“*. Nach dieser Definition wären sämtliche intraspezifischen und intraorganismischen ‘Homologien’ ausgeschlossen, sogenannte *general homologies* (siehe Schmitt, 1989) oder *iterative homologies* (*sensu* Wagner, 1989), wie z.B. die serielle Homologie, die ontogenetische Homologie, die sexuelle Homologie und die polymorphe Homologie. Es wären nur sogenannte supraspezifische Homologien (auch spezielle Homologien; siehe z.B. Schmitt, 1989) Homologien im Sinne von Owen. Bronn (1858) hat zum Zweck der Unterscheidung dieser beiden Gruppen ersteren den Begriff *Homonomie* zugewiesen.

Aber auch für die supraspezifischen Homologien existieren die verschiedensten Konzeptionen (einen Überblick über die aktuelle Diskussion liefern die Bände von Hall, 1994; Bock und Cardew, 1999).

So lässt sich Owens Konzept auch im Rahmen von vergleichenden Studien ohne die Annahme der Evolutionstheorie sinnvoll anwenden. Dies wird als *biological homology* bezeichnet (*sensu* Roth, 1984; Woese, 1987; Aboitiz, 1988; taxic homology *sensu* Patterson, 1982, und Carine und Scotland, 1999). Dabei wird *"the same organ in different animals"* als strukturelle oder funktionelle Gleichheit verstanden. Es bezieht sich auf gemeinsame entwicklungsbiologische Zwänge unter den Arten und versucht konservative Muster der Evolution morphologischer Merkmale kausal zu erklären (Wagner, 1989; Roth, 1991; Sluys, 1996).

Mit dem Aufkommen der Evolutionstheorie war jedoch auch die Möglichkeit für eine andere Interpretation des owenschen Ausdrucks *„the same organ in different animals"* gegeben. Remane (1952, 1954, 1961) gehört möglicherweise zu den ersten Autoren, die eine Homologie-Konzeption vorgestellt haben, welche heutzutage als *historical homology* (*sensu* Wagner, 1989) oder *phylogenetic homology* (*sensu* Roth, 1984) bezeichnet wird. *„The same organ in different animals"* wird hier als historisch-ontogenetische Gleichheit aufgefasst. Aber auch hier gibt es wieder verschiedene Ausprägungen dieser Homologie-Konzeption (siehe u.a. de Pinna, 1991; Brower und Schawaroch, 1996; Sluys, 1996).

Es existieren also eine Vielzahl von zum Teil sich fundamental voneinander unterscheidenden Konzepten, die alle mit einem einzigen Terminus in Verbindung gebracht werden – *Homologie*. Ob es nun sinnvoll ist, einen einzigen Terminus für all diese verschiedenen Konzepte zu verwenden, oder ob jedem einzelnen Konzept ein eigener Name zugewiesen werden sollte, um Missverständnissen vorzubeugen und die Transparenz zu erhöhen (Schmitt, 1989; Butler, 2000), soll nicht die Aufgabe der vorliegenden Arbeit sein. In *Testing and Weighting Characters* habe ich mich vielmehr damit beschäftigt, wie ein für die Phylogenetik geeignetes Homologie-Konzept beschaffen sein muss, um darüber hinaus auch noch den Anforderungen des popperschen Falsifikationismus an eine empirische Wissenschaft zu genügen.

Ausgangpunkt für *Testing and Weighting Characters* stellt also die Frage nach den notwendigen Bedingungen für die Möglichkeit dar, die Phylogenetik als eine empirische Wissenschaft *sensu* Popper (1983, 1994) zu etablieren. Dabei wird deutlich, dass eine konzeptuelle Unterscheidung von Homologie und Apomorphie sinnvoll ist (wenn ich in den Artikeln den englischen Ausdruck *synapomorphy* verwende, so liegt das daran, dass im englischsprachigen Raum *synapomorphy* weitestgehend synonym zum deutschen *Apomorphie* verwendet wird). Dabei sollte sich die Konzeption einer phylogenetischen Homologie nur auf die genealogischen Verwandschaftsbeziehungen, also auf die verwandtschaftlichen Beziehungen einzelner Individuen zueinander, beziehen. Als empirisches Test-Kriterium phylogenetischer Homologiehypothesen kann die Forderung nach *Identität* der homologen Eigenschaften herangezogen werden. Dies wäre also eine empirische, anhand von Beobachtungen und Experimenten überprüfbare Qualität.

Ausgehend von einer solchen Homologie-Konzeption kann eine eindeutige Differenzierung der Begriffe *Merkmal* und *Merkmalszustand* vorgenommen werden, wobei ein phylogenetisches Merkmal als eine evolutionsgeschichtliche Hypothese über ein singuläres Transformationsereignis verstanden wird, welches in seiner vollständigen Form immer aus zwei Komponenten bestehen muss, nämlich dem Zustand vor und dem Zustand nach der Transformation, welche durch zwei unterschiedliche Merkmalszustände repräsentiert werden. Somit dient eine Homologiehypothese als erkenntnistheoretisches Argument in der Rekonstruktion einzelner Transformationsereignisse.

Apomorphie unterscheidet sich von Homologie durch dessen Bezugnahme auf die verwandtschaftlichen Beziehungen von Spezies anstelle der von einzelnen Individuen. Aber da diese sich nur über die Elemente der Spezies, den einzelnen Individuen, erfassen lassen, besteht eine logische Verbindung zwischen den beiden Konzepten; Apomorphien sind zwingend auch immer Homologien – umgekehrt gilt dies jedoch nicht. Daraus folgt, dass auch für Apomorphiehypothesen der Identitäts-Test während der Merkmalsanalyse angewendet werden kann. Darüber hinaus ermöglicht die Konzeption von Apomorphie einen weiteren Test, den *Kongruenz*-Test, welcher während der kladistischen Analyse durchgeführt wird. Da dieser Test jedoch nicht gegen empirisch überprüfbare Eigenschaften vollzogen wird, sondern

Apomorphiehypothesen gegen Apomorphiehypothesen testet, stellen erfolgreich im Kongruenz-Test getestete Apomorphiehypothesen nur dann empirische Hypothesen dar, wenn sie zuvor gegen empirische Daten getestet worden sind. Sind sie das, so haben sie nach Popper durch diese zuvor erfolgreich bestandenen Tests bereits einen gewissen Bewährungsgrad erhalten, den sie als Gewicht in den Kongruenz-Test mit einbringen. Und nur auf diesem Weg käme Apomorphiehypothesen, welche den Kongruenz-Test erfolgreich bestanden haben, eine empirische Erklärungskraft zu. D.h., der Identitäts-Test liefert die Basis für die Gewichtung von Merkmalen für die kladistische Analyse, und die hier vorgestellte Homologie-Konzeption liefert das notwendige konzeptionelle Bindeglied zwischen den merkmalstragenden empirischen Entitäten, den einzelnen Individuen, und den Argumenten für die Stammesrekonstruktion, den Apomorphien.

Dabei nimmt die Merkmalsanalyse mit ihrem integrierten Identitäts-Test eine zentrale Rolle bei der Begründung eines *a priori* Gewichtungssystems innerhalb eines falsifikationistischen Ansatzes in der Phylogenetik ein. Darüber hinaus wird deutlich, dass die Gewichtung von Merkmalen vor der kladistischen Analyse zwingend notwendig ist, möchte man die Phylogenetik als eine empirische Wissenschaft *sensu* Popper verteidigen (insbesondere diese Thematik wird in *Process Probabilities and the Weighting of Characters in Systematics - Following the falsificationist program of phylogenetic research* wieder aufgegriffen und weiter erläutert).

Aus der Abhängigkeit der Bedeutung des Kongruenz-Tests der kladistischen Analyse von seiner strukturellen Beziehung zum Identitäts-Test der Merkmalsanalyse lässt sich eine Bewertung bezüglich der ersten These der *Patter Cladisten* formulieren. Die Behauptung, der Kongruenz-Test sei der entscheidende Test bei phylogenetischen Untersuchungen, da nur in diesem Test eine Falsifikationsmöglichkeit phylogenetischer Hypothesen gegeben sei, ist innerhalb einer falsifikationistischen Konzeption der Phylogenetik nicht zwingend. Vielmehr muss neben dem Kongruenz-Test eine weitere und konzeptionell näher an der empirischen Basis der Untersuchungen angesiedelte Falsifikationsmöglichkeit bestehen, wenn der Kongruenz-Test ein empirischer Test sein soll. Demnach muss der Kongruenz-Test konzeptionell auf dem Identitäts-Test aufbauen. Jedoch soll dies nicht Anlass zur Folgerung liefern, dass der Identitäts-Test eine dem Kongruenz-Test gegenüber übergeordnete Rolle spielte – beide

Falsifikationsmöglichkeiten stellen wichtige Bestandteile phylogenetischer Untersuchungen dar.

Erste Schlussfolgerungen aus diesem Ansatz werden in einem Gewichtungssystem für Sequenzdaten nicht proteinkodierender 18S rDNA exemplarisch umgesetzt und innerhalb einer empirischen Untersuchung getestet. Hierbei wird insbesondere einer differentiellen Gewichtung von Insertionen, Deletionen und Nukleotid-Transformationen die Aufmerksamkeit geschenkt. Die Ergebnisse dieser Untersuchung werden in *Weighting Indels as Phylogenetic Markers of 18S rDNA Sequences in Diptera and Strepsiptera* dargelegt.

Prozesswahrscheinlichkeiten und die Gewichtung von Merkmalen

Aufbauend auf die in *Testing and Weighting Characters* entwickelte Konzeption einer phylogenetischen Homologie und Apomorphie, sowie ihrer konzeptionellen Beziehung zueinander und den zwei möglichen Testkriterien Identität und Kongruenz, und aufbauend auf der Erkenntnis, dass die Gewichtung von Merkmalen von der Härte des von ihnen innerhalb der Merkmalsanalyse bestandenen Identitäts-Tests abhängt, wird in *Process Probabilities and the Weighting of Characters in Systematics - Following the falsificationist program of phylogenetic research* die Rolle von Prozesswahrscheinlichkeiten bei der Evaluation von Merkmalsgewichten untersucht. Darüber hinaus lassen sich aus den Ergebnissen dieser Analyse weitere kritische Schlüsse bezüglich der drei Thesen der *Pattern Cladisten* ziehen, und es lässt sich nachweisen, dass sie nicht zwingend aus der Anwendung des popperschen Falsifikationismus in der Phylogenetik folgen.

Um in die weitreichenden Auseinandersetzungen um die Umsetzung des popperschen Falsifikationismus in der Phylogenetik ein wenig mehr Licht zu bringen, unterscheide ich in dieser Arbeit das Konzept der Falsifikation und das der Bewährung. Die Umsetzung des Falsifikationskonzeptes innerhalb der phylogenetischen Untersuchungen ist bereits in *Testing and Weighting Characters* ausreichend dargestellt.

Der Bewährungsgrad stellt die Basis für die Wahl der zur Zeit am besten begründeten kladistischen Hypothese dar. Eine Quantifizierung der Bewährungsgrade ist in phylogenetischen Untersuchungen immer dann notwendig, wenn die Merkmalsverteilung verschiedener Merkmale ihrer hypothetisierten Apomorphie widersprechen. Dem relativen Grad der Bewährung einer Apomorphiehypothese wird dabei innerhalb der kladistischen Analyse durch ein spezifisches Gewicht des entsprechenden Merkmals bzw. der entsprechenden Transformation Rechnung getragen. Dabei hängt das jeweilige Gewicht von der Härte des Identitäts-Tests ab.
Betrachtet man nun Poppers Vorschlag für eine mathematische Formulierung des Bewährungsgrades, so wird deutlich, dass die einzige sinnvolle Interpretation dieser Formel durch die Anwendung von Prozesswahrscheinlichkeiten für die entsprechenden evolutionären Transformationsereignisse ermöglicht wird. Dabei tritt folgende Beziehung zu Tage: Je unwahrscheinlicher ein evolutionäres Transformationsereignis, welches eine neue vererbbare organismische Eigenschaft verursacht, umso Härter der Identitäts-Test und somit umso höher das Merkmalsgewicht, das vergeben werden muss, sollte der Test von der Apomorphiehypothese erfolgreich bestanden werden.

Damit wird deutlich, dass, entgegen der dritten These der *Pattern Cladisten*, Prozesswahrscheinlichkeiten eine zentrale Rolle innerhalb der Gewichtung von Merkmalen zukommt und somit integraler Bestandteil der kladistischen Analysen sein muss – insbesondere bei der Verfolgung eines falsifikatorischen Wissenschaftsansatzes. Und es lässt sich keine Begründung innerhalb des falsifikationistischen Ansatzes für die Anwendung der sogenannten *non-weighting* Maximum Parsimonie Methode finden (im Widerspruch zu z.B. Kluge, 1997a). Auch lässt sich auf dieser Basis die Zurückweisung der Maximum Likelihood Methode als eine gegenüber dem Falsifikationismus vermeidlich inkonsistente Methode nicht begründen (siehe hierzu auch die Diskussion von de Queiroz und Poe, 2001; Faith und Trueman, 2001; Farris *et al.*, 2001; Kluge, 2001).

In einer Simulationsstudie zu *Process Probabilities and the Weighting of Characters in Systematics - Following the falsificationist program of phylogenetic research* wird mit Hilfe von einer Computersoftware die Evolution von Sequenzen simuliert und die daraus erhaltenen Daten als empirische Grundlage herangezogen, um das Verhalten der Maximum Parsimonie Methode unter Verwendung verschiedener Merkmals- und

Transformationsgewichte zu untersuchen. Dabei wird deutlich, dass diejenigen Analysen, welche den Produkten von unwahrscheinlichen Transformationsprozessen ein verhältnismäßig hohes Gewicht zuweisen, die den Daten zugrundeliegende Ereignisabfolge signifikant besser rekonstruieren als Analysen, die eine andere Gewichtung anwenden. Dies ist als eine weitere Bestätigung der zentralen Rolle, die evolutionären Prozesswahrscheinlichkeiten bei der Stammesrekonstruktion zukommt, zu werten.

Tatort Evolution

Es bestehen gewisse Parallelen zwischen dem Vorgehen und der Arbeitsweise einer/eines KriminalkommissarIn oder DetektivIn und dem einer/eines PhylogenetikerIn. Wie Sherlock Holmes müssen auch PhylogenetikerInnen die am *Tatort* verbliebenen Spuren aufnehmen und möglichst unverändert dokumentieren, die relevanten aus der Menge der irrelevanten Spuren extrahieren, um diese dann im Hinblick auf ihre Aussagekraft zu bewerten und möglichst schlüssig zu interpretieren. Ihnen ist also die Suche nach empirischen Indizien gemeinsam, um eine möglichst gut begründete Rekonstruktion des *Tathergangs* zu ermöglichen und eine, so weit dies die Tatbestände zulassen, vorläufig gerechtfertigte Hypothese aufzustellen. Wurde die Tat nicht direkt beobachtet, so ist eine Überführung nicht möglich und wie in der Phylogenetik, so ist auch in der Kriminologie ein echter Beweis nicht existent und jede Hypothese prinzipiell fallibel. Es geht in beiden Arbeitsfeldern also um das Deuten und Interpretieren von Spuren, oder auch Zeichen, die ein bestimmter zu rekonstruierender Prozess hinterlassen hat. Dabei können diese Spuren als bestimmte Typen eines allgemeiner gefassten Zeichenkonzepts verstanden werden, wie es C. S. Peirce in seiner allgemeinen Zeichentheorie (Semiotik) vorgestellt hat.

In dem Artikel *Signs and Phylogeny - A Semiotic Approach to Systematics* stelle ich die Anwendung der Theorie der Semiotik auf die phylogenetische Methodologie und Arbeitsweise vor. Dabei wird auf die semiotische Erkenntnistheorie von C.S. Peirce als Grundlage zurückgegriffen. Der Vorteil des semiotischen Ansatzes von Peirce ist hierbei nicht zuletzt, dass er einem analytische Werkzeuge zur Hand gibt, die nicht nur eine kritische Analyse der vermuteten evolutionären Prozesse und ihrer konzeptionellen Beziehungen untereinander ermöglicht, sondern auch die verschiedenen Schritte der phylogenetischen Untersuchung selbst systematisch und kritisch analysiert und die

unterschiedlichen Schlussformen, die sie für ihre Analysen bemüht, aufdeckt. Hierbei ist insbesondere die von Peirce vorgenommene Unterscheidung von Abduktion, Induktion und Deduktion und ihre Rolle innerhalb wissenschaftlicher Untersuchungen von entscheidender Bedeutung. Damit wird eine andere Perspektive auf das phylogenetische Arbeiten eröffnet, die u.a. eine Bewertung der Diskussionen um die Konsequenzen der Anwendung des popperschen Falsifikationismus auf die Phylogenetik ermöglicht.

Der Ansatz sieht nicht nur in der Evolution der Organismen und ihren Merkmalen einen Zeichenprozess (Semiosis) - eine Evolution von natürlichen Zeichen, sondern auch in der von PhylogenetikerInnen betriebenen Stammesrekonstruktion einen solchen. Dabei geht es darum, einige der Zeichen *aus* der Phylogenie als Zeichen *zur* phylogenetischen Rekonstruktion richtig zu *interpretieren*. Hierfür ist es zwingend erforderlich, eine allgemeine Konzeption des phylogenetischen Zeichens, also des Zeichens zur phylogenetischen Rekonstruktion, zu entwickeln. Die natürlichen Zeichen müssen bestimmte Eigenschaften aufweisen, um als phylogenetische Zeichen verwendet werden zu können. So müssen sie z.B. aus einem Mechanismus der organismischen Vererbung hervorgehen, der eine gewisse Neigung zur Konservierung der Struktur aufweist. Haben sie diese Eigenschaften, so stellen sie potentielle Indizes zur phylogenetischen Vergangenheit der entsprechenden Spezies dar.

Aus der Untersuchung zu den notwendigen Eigenschaften phylogenetischer Zeichen, um sie als empirische Argumente innerhalb der Stammesrekonstruktion verwenden zu können, erfolgt eine Bestätigung der in *Testing and Weighting Characters* aufgestellten These der Konzeption des kladistischen Merkmals: ein vollständiges Argument für die Rekonstruktion eines evolutionären Transformationsereignisses besteht immer aus zwei und nicht mehr Bestandteilen – dem Zustand vor der Transformation und demjenigen nach der Transformation. Bei einem kladistischen Merkmal, das in der Stammesrekonstruktion verwendet werden soll, handelt es sich notwendigerweise um ein *empirisches Argument* und somit um eine Hypothese, da es zwischen dem prinzipiell nicht direkt beobachtbaren Prozess der Phylogenie und den beobachtbaren organismischen Eigenschaften *vermitteln* soll. Es besteht also immer aus zwei beobachtbaren organismischen Eigenschaften, dem *plesiomorphen* und dem *apomorphen* Merkmalszustand.

Auf der Basis der semiotischen Theorie und der bisherigen Kenntnisse über generelle Prozesse der Evolution, lässt sich das Konzept eines idealen phylogenetischen Zeichens formulieren. Auf der Grundlage der spezifischen Eigenschaften dieses idealen Zeichens lassen sich Testkriterien ableiten, anhand derer man die „Ähnlichkeit“ der materiellen Zeichen zu diesem Ideal überprüfen kann. Dabei finden die aus den bisherigen Studien gewonnen Kriterien der *Identität* und der *Kongruenz* auch im semiotischen Ansatz eine Bestätigung.

Unter Bezugnahme auf die wechselseitige Erhellung, die das Wissen um generelle evolutionäre Mechanismen und dasjenige um spezifische singuläre phylogenetische Ereignisfolgen aufeinander ausüben, wird deutlich, dass es bei phylogenetischen Untersuchungen ratsam ist, alles genügend gut bewährte relevante Hintergrundwissen innerhalb eines jeden Analyseschrittes zu berücksichtigen. Da der semiotische Ansatz in seinen Grundzügen im Einklang zum popperschen Falsifikationismus steht, wird auch die zweite These der *Pattern Cladisten* widerlegt und die Forderung nach der Minimierung des Umfangs des angenommenen Hintergrundwissens als unhaltbar offengelegt.

Schlussfolgerung

Die vorliegende Arbeit verfolgt nicht den Anspruch, ein konkretes Verfahren zur Gewichtung von kladistischen Merkmalen vorstellen zu können. Vielmehr soll die konzeptionelle Struktur eines allgemeinen Gewichtungsschemas offengelegt werden. Sowohl die Arbeiten zur Anwendung des Falsifikationismus wie auch diejenigen zur Anwendung der Semiotik in der Phylogenetik verdeutlichen, dass Annahmen zu den entsprechenden evolutionären Prozesswahrscheinlichkeiten bei der Erstellung und Begründung von Merkmalsgewichten unumgänglich sind. Gänzlich ungeklärt bleibt, wie man auf diese Wahrscheinlichkeiten schließen kann. Doch gibt es eine Vielzahl vielversprechender statistischer Ansätze, die allerdings insbesondere für die Bewertung von molekularen Daten entwickelt worden sind.

Des weiteren soll die Arbeit einen Gegenpol zu der bisher ein wenig einseitigen Interpretation des popperschen Falsifikationismus in der Phylogenetik liefern, die der Komplexität des Themas, sowohl was die phylogenetische als auch die

erkenntnistheoretische Seite betrifft, gerecht zu werden versucht. Dabei wird deutlich, dass die Gründe, die von einigen Autoren gegen die Verwendung von Likelihood Methoden angeführt werden, sich aus einem falsifikatorischen Ansatz allein nicht rechtfertigen lassen und demnach, in der bisher vorgebrachten Form, nicht länger haltbar sind. Und was als vermeidliche Schwäche den Likelihood Methoden diagnostiziert worden ist, nämlich die explizite Verwendung von Transformationswahrscheinlichkeiten, mag sich noch als eine ihrer wesentlichen Stärken erweisen.

Mit dem semiotischen Ansatz ist es gelungen, eine neue Perspektive auf bestehende und bekannte Probleme der phylogenetischen Theorie und Methodologie zu eröffnen. Dadurch, dass sie sowohl die Konzeptualisierung des Prozesses der Phylogenie als auch die des Prozesses der phylogenetischen Rekonstruktion innerhalb eines gemeinsamen epistemologischen System ermöglicht, eröffnet sie die Basis für eine kohärente Konzeptualisierung aller Ebenen der Phylogenetik, einhergehend mit ihrer kritischen Analyse. Dies betrifft Konzepte wie das der Homologie, der Apomorphie und des kladistischen Merkmals, wie auch die Systematisierung phylogenetischer Untersuchungsschritte innerhalb der Merkmalsanalyse und der kladistischen Analyse, sowie die von ihnen bemühten Folgerungs- und Schlussformen und deren Verlässlichkeit.

Weitere Anwendungen dieses Ansatzes liegen auf der Hand und würden z.B. ein in sich kohärentes allgemeines *Spezies* Konzept mit den ihm eigenen spezifischen Eigenschaften oder z.B. ein allgemeines Konzept für *biologisches Merkmal* mit all seinen in den verschiedenen biologischen Disziplinen unterschiedlichen Ausprägungen ermöglichen. Die Anwendungen der semiotischen Theorie innerhalb der Phylogenie und den Biowissenschaften insgesamt sind bei weitem noch nicht ausgeschöpft.

Literatur

Aboitiz, F. (1988): Homology: A Comparative or a Historical Concept? *Acta Biotheoretica* **37**, 27-29.

Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A., und Lake, J. A. (1997): Evidence for a clade of nematodes, arthropods and other moulting animals. *Nature* **387**, 489-493.

Allard, M. W. und Carpenter, J. M. (1996): On weighting and congruence. *Cladistics* **12**, 183-198.

Ax, P. (1999): *Das System der Metazoa II - Ein Lehrbuch der phylogenetischen Systematik*. Akademie der Wissenschaften und der Literatur, G. Fischer, Stuttgart.

Ballard, J. W. O., Thayer, M. K., Newton, A. F. J. und Grismer, E. R. (1998): Data sets, partitions, and charakters: Philosophies and procedures for analyzing multiple data sets. *Systematic Biology* **47**, 367-396.

Björklund, M. (1999): Are third positions really that bad? A test using vertebrate cytochrome b. *Cladistics* **15**, 191-197.

Bock, G. R. und Cardew, G. (1999): *Homology*. Novartis Foundation Symposium 222, John Wiley & Sons, Chichester.

Bronn, H. G. (1858): *Morphologische Studien über die Gestaltungsgesetze*. Leipzig und Heidelberg.

Broughton, R. E., Stanley, S. E. und Durrett, R. (2000): Quantification of homoplasy for nucleotide transitions and transversions and a reexamination of assumptions in weighted phylogenetic analysis. *Systematic Biology* **49**, 617-627.

Brower, A. V. Z. und Schawaroch, V. (1996): Three Steps of Homology Assessment. *Cladistics* **12**, 265-272.

Brusca, R. C. und Brusca, G. J. (1990): *Invertebrates*. Sinauer Ass. Inc., Sunderland, Massachusetts.

Bryant, H. N. (1989) : An evaluation of cladistic and character analyses as hypothetico-deductive procedures, and the consequences for character weighting. *Systematic Zoology* **38**, 214-227.

Butler, A. B. (2000): Book reviews: The complexity of sameness. *Trends in Ecology and Evolution* **15**, 39-40.

Carine, M. A. und Scotland, R. W. (1999): Taxic and Transformational Homology: Different Ways of Seeing. *Cladistics* **15**, 121-129.

Carpenter, J. M. (1988): Choosing among multiple equally parsimonious cladograms. *Cladistics* **4**, 291-296.

Chippindale, P. T. und Wiens, J. J. (1994): Weighting, partitioning, and combining characters in phylogenetic analysis. *Systematic Biology* **43**, 278-287.

de Pinna, M. C. C. (1991) : Concepts and tests of homology in the cladistic paradigm. *Cladistics* **7**, 367-394.

de Queiroz, K. und Poe, S. (2001) : Philosophy and phylogenetic inference : A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. *Systematic Biology* **50**, 305-321.

Eernisse, D. J. und Kluge, A. G. (1993): Taxonomic congruence versus total evidence, and amniote phylogeny inferred from fossils, molecules, and morphology. *Molecular Biology and Evolution* **10**, 1170-1195.

Faith, D. P. und Trueman, J. W. H. (2001): Towards an inclusive philosophy for phylogenetic inference. *Systematic Biology* **50**, 331-350.

Farris, J. S. (1969): A successive approximations approach to character weighting. *Systematic Zoology* **18**, 374-385.

Farris, J. S., Kluge, A.G. und Carpenter, J. M. (2001): Popper and likelihood versus "Popper*". *Systematic Biology* **50**, 438-444.

Felsenstein, J. (1978): Cases in which parsimony or compatibility methods will be positively misleading. *Systematic Zoology* **27**, 401-410.

Flook, P. K., Klee, S. und Rowell, C. H. F. (1999): Combined molecular phylogenetic analysis of the Orthoptera (Arthropoda, Insecta) and implications for their higher systematics. *Systematic Biology* **48**, 233-253.

Foucault, M. (1997): *Die Ordnung der Dinge: eine Archäologie der Humanwissenschaften*. 14. Auflage, aus dem Franz. übersetzt von Ulrich Köppen, Suhrkamp-Taschenbuch Wissenschaft 96, Suhrkamp, Frankfurt a.M..

Giribet, G. und Wheeler, W. C. (1999): The position of arthropods in the animal kingdom: ecdysozoa, islands, Trees, and the "parsimony ratchet". *Molecular Phylogenetics and Evolution* **13**, 619-623.

Goloboff, P. A. (1993): Estimating character weights during tree search. *Cladistics* **9**, 83-91.

Goloboff, P. A. (1995): Parsimony and weighting: a reply to Turner and Zandee. *Cladistics* **11**, 91-104.

Hall, B. K. (1994): *Homology: The hierarchical basis of comparative biology*. Academic Press, San Diego, CA.

Haszprunar, G. (1998): Parsimony analysis as a specific kind of homology estimation and the implications for character weighting. *Molecular Phylogenetics and Evolution* **9**, 333-339.

Hedges, S. B. and Maxson, L. R. (1996): Re: Molecules and morphology in amniote phylogeny. *Molecular Phylogenetics and Evolution* **6**, 312-319.

Hennig, W. (1966): *Phylogenetic Systematics*. Übers. D. D. Davis und R. Zangerl, University of Illinois Press, Urbana.

Hillis, D.M., Moritz, C. und Mable, B. (1996): *Molecular Systematics*. 2nd Edition, Sinauer Ass. Inc., Sunderland, Massachusetts.

Huelsenbeck, J. P. (1995): Performance of phylogenetic methods in simulation. *Systematic Biology* **44**, 17-48.

Huelsenbeck, J. P. und Hillis, D. M. (1993): Success of phylogenetic methods in the four taxon case. *Systematic Biology* **42**, 247-264.

Källersjö, M., Albert, V. A. und Farris, J. S. (1999): Homoplasy *increases* phylogenetic structure. *Cladistics* **15**, 91-93.

Kitching, I. J., Forey, P. L., Humphries, C. J. und Williams, D. M. (1998): *Cladistics. The Theory And Practice Of Parsimony Analysis*. 2nd Edition, Oxford University Press, Oxford.

Kluge, A. G. (1997): Testability and the Refutation and Corroboration of Cladistic Hypotheses. *Cladistics* **13**, 81-96.

Kluge, A. G. (1997a): Sophisticated falsification and research cycles: Consequences for differential character weighting in phylogenetic systematics. *Zoologica Scripta* **26**, 349-360.

Kluge, A. G. (1998): Total evidence or taxonomic congruence: Cladistics or consensus classification. *Cladistics* **14**, 151-158.

Kluge, A. G. (2001): Philosophical conjectures and their refutation. *Systematic Biology* **50**, 322-330.

Lee, M. S. Y. (1997): Re: Molecules, and phylogeny: A response to Hedges and Maxson. *Molecular Phylogenetics and Evolution* **7**, 394-395.

Lewis, P. O. (2001): A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology* **50**, 913-925.

Linné, C. von (1758): *Systemae naturae*. Leyden.

Littlewood, D. T. J., Smith, A. B., Clough, K. A. und Emson, R. H. (1997): The interrelationships of the echinoderm classes: morphological and molecular evidence. *Biological Journal of the Linnean Society* **61**, 409-438.

Lutzoni, F., Wagner, P., Reeb, V. und Zoller, S. (2000): Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. *Systematic Biology* **49**, 628-651.

Manuel, M., Kruse, M., Müller, W. E. G. und Le Parco, Y. (2000): The comparison of β-Thymosin homologues among Metazoa supports an Arthropod-Nematode clade. *Journal of Molecular Evolution* **51**, 378-381.

Milinkovitch, M. C., LeDuc, R. G., Adachi, J., Farnir, F., Georges, M. und Hasegawa, M. (1996): Effects of character weighting and species sampling on phylogeny reconstruction: A case study based on DNA sequence data in Cetaceans. *Genetics* **144**, 1817-1833.

Neff, N. A. (1986): A rational basis for a priori character weighting. *Systematic Zoology* **35**, 110-123.

Nielsen, C. (1995): *Animal evolution*. Oxford University Press, Oxford.

O'Keefe, F. R. und Wagner, P. J. (2001): Inferring and testing hypotheses of cladistic character dependence by using character compatibility. *Systematic Biology* **50**, 657-675.

Owen, R. (1843): *Lectures on comparative Anatomy and Physiology of the Invertebrate Animals*. Longman, Brown, Greens and Longmans, London.

Page, R. D. M. (1996) : Consensus, confidence, and "total evidence". *Cladistics* **12**, 83-92.

Page, R. D. M. und Holmes, E. C. (1998): *Molecular Evolution. A phylogenetic Approach*. Blackwell Science LTD, Oxford.

Patterson, C. (1982): Morphological characters and homology. *In* "Problems of phylogenetic reconstruction", K. A. Joysey und A. E. Friday (eds.), pp. 21-74. Academic Press, London.

Popper, K. R. (1983): *Realism and the aim of science; from the postscript to the logic of scientific discovery*. Routledge, London and New York.

Popper, K. R. (1994): *Logik der Forschung*. 10., verbesserte und vermehrte Auflage von "Die Einheit der Gesellschaftswissenschaften; Band 4, J. C. B. Mohr (Paul Siebeck), Tübingen.

Remane, A. (1952): *Die Grundlagen des Natürlichen Systems, der Vergleichenden Anatomie und der Phylogenetik*. Leipzig 1952.

Remane, A. (1954): *Morphologie als Homologienforschung*. Verhandlungen der Deutschen Zoologischen Gesellschaft in Tübingen (1954).

Remane, A. (1961): Gedanken zum Problem: Homologie und Analogie, Praeadaptation und Parallelität. *Zoologischer Anzeiger* **166**, 447-465.

Rieppel, O. (1999): *Einführung in die computergestützte Kladistik.* Verlag Dr. Friedrich Pfeil, München.

Roth, V. L. (1984): On Homology. *Biological Journal of the Linnean Society* **22**, 13-29.

Roth, V. L. (1991): Homology and hierarchies: problems solved and unresolved. *Journal of Evolutionary Biology* **4**, 167-194.

Schmidt-Rhaesa, A., Bartolomaeus, T., Lemburg, C., Ehlers, U. und Garey, J. R. (1998): The position of the arthropoda in the phylogenetic system. *Journal of Morphology* **238**, 263-285.

Schmitt, M. (1989): Das Homologie-Konzept in Morphologie und Phylogenetik. *Zoologische Beiträge der Naturforschenden Gesellschaft* **32**, 505-512.

Siddall, M. E. und Kluge, A. G. (1997): Probabilism and Phylogenetic Inference. *Cladistics* **13**, 313-336.

Sluys, R. (1996): The notion of homology in current comparative biology. *Zeitschrift für zoologische Systematik und Evolutionsforschung* **34**, 145-152.

Sullivan, J. und Swofford, D. L. (2001): Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated? *Systematic Biology* **50**, 723-729.

Swofford, D. L., Waddel, P. J., Huelsenbeck, J. P., Foster, P. G., Lewis, P. O. und Rogers, J. S. (2001): Bias in phylogenetic estimation and its relevance to the choice between parsimony and likelihood methods. *Systematic Biology* **50**, 525-539.

Trueman, J. W. H. (1998): Reverse successive weighting. *Systematic Biology* **47**, 733-737.

Wägele, J. W., Erikson, T., Lockhart, P. und Misof, B. (1999): The Ecdysozoa: Artifact or monophylum? *Journal of Zoological Systematics and Evolutionary Research* **37**, 211-223.

Wägele, J. W. und Misof, B. (2001): On quality of evidence in phylogeny reconstruction: a reply to Zrzavý's defence of the 'Ecdysozoa' hypothesis. *Journal of Zoological Systematics and Evolutionary Research* **39**, 165-176.

Wägele, J. W. (2000): *Grundlagen der Phylogenetischen Syste*matik. Verlag Dr. Friedrich Pfeil, München.

Wagner, G. P. (1989): The Biological Homology Concept. *Annual Reviews of Ecology and Systematics* **20**, 51-69.

Wenzel, J. W. und Siddall, M. E. (1999): Noise. *Cladistics* **15**, 51-64.

Wheeler, Q. D. (1986): Character weighting and cladistic analysis. *Systematic Zoology* **35**, 102-109.

Wiens, J. J. (1998): Combining data sets with different phylogenetic histories. *Systematic Biology* **47**, 568-581.

Wiens, J. J. (2001): Character analysis in morphological phylogenetics: Problems and solutions. *Systematic Biology* **50**, 689-699.

Wiley, E. O., Johnson, G. D. und Dimmick, W. W. (1998): The phylogenetic relationships of lampridiform fishes (Teleostei: Acanthomorpha), based on a total-evidence analysis of morphological and molecular data. *Molecular Phylogenetics and Evolution* **10**, 417-425.

Woese, C. R. (1987): Bacterial evolution. *Microbiological Reviews* **51**, 221-271.

Yang, Z. (1996): Phylogenetic Analysis Using Parsimony and Likelihood Methods. *Journal of Molecular Evolution* **42**, 294-307.

Zrzavý, J. (2001): Ecdysozoa versus Articulata: Clades, artifacts, prejudices. *Journal of Zoological Systematics and Evolutionary Research* **39**, 159-163.

Poster

Titel:

Zur Phylogenie der Polychaeta anhand von 18S rDNA Sequenzen

Vorgestellt: Jahresversammlung der Deutschen Zoologischen Gesellschaft (DZG) in Osnabrück, Juni 04.-08. 2001.

Abstract publiziert in: (2001) Zoology 104 (Suppl. IV), 69.

Autor: Bleidorn, C., Vogt, L., Bartolomaeus, T
Abt. Zoomorphologie und Systematik, 33501 Bielefeld, Germany;
email: lars.vogt@biologie.uni-bielefeld.de

Abstract:

The phylogeny of the Annelida is still a matter of debate (Rouse & Fauchald 1997 Zool Scripta 26:139, McHugh 2000 Can J Zool 78:1873), which certainly arises from the organisational heterogeneity within this group that reflects specific environmental adaptations of certain annelid taxa as well as the enormous age of this group. More morphological and molecular data than presently known are needed for an attempt to solve the phylogenetic relationships of Annelida. While morphological data have been collected for more than a century, molecular data is rare. When analysed, it is in conflict with the morphological data. In order to increase the data sets we sequenced 18SrDNA data from several sessile and hemisessile polychaete species, aligned them together with the known sequences and several representatives of the Bilateria. The data set was analysed by neighbour joining, maximum parsiomony and maximum likelihood methods. Using the Clitellata as outgroup a smaller data set was analysed in the same manner. Special attention was focussed on results that were in accordance with the morphological data and supported by bootstrap values higher than 95. All analyses clearly show a low resolution between the different higher polychaete taxa, but also indicate several well substantiated entities within the polychaeta. One of them is a cladeconsisting of representatives of the Orbiniinae, Protoaricinae and Questidae. This result is in accordance with phylogenetic analyses based on morphological characters (Rouse & Fauchald 1997: Zool Scr 26:139). These characters, however, were interpreted as homoplasies, so that our molecular data provide some further support for a common ancestry of the three taxa in question. The study is part of a larger attempt to gain more data on “sedentary” polychaete molecular markers.

Supported by the Deutsche Forschungsgemeinschaft (Ba 1520/4-1)

Titel:

Travisia ist ein Scalibregmatidae, kein Opheliidae (Annelida, Polychaeta)

Vorgestellt: Jahresversammlung der Deutschen Zoologischen Gesellschaft (DZG) in Halle, Mai 20.-24. 2002.

Abstract publiziert in: (2002) Zoology 105 (Suppl. V), 60.

Autor: Bleidorn, C., Vogt, L., Arnhold, N., Bartolomaeus, T
Abt. Zoomorphologie und Systematik, 33501 Bielefeld, Germany;
email: lars.vogt@biologie.uni-bielefeld.de

Abstract:

Opheliids are worldwide distributed polychaets. Up to now, more than 150 species have been described. Morphologically, three distinct groups can be recognized: A taxon Opheliinae comprising those genera with distinct body regions (Euzonus, Lobochesis and Ophelia), a taxon Ophelininae including all genera with an anal tube consisting of several reduced segments (Ammotrypanella, Antiobactrum, Armandia, Ophelina, Polyophthalmus and Tachytrypane) and the taxon Travisia, which closely resembles scalibregmatid species (Bellan G et al. 1990 C. R. Acad. Sci. Paris 310: 175, Hartmann-Schröder G 1996 Buchtitel.). However, there is no convincing apomorphy for the Opheliidae (Fauchald K & Rouse GW 1997 Zool Scripta 26: 71) and Rouse & Pleijel (2001 Polychaetes, Oxford University Press) annotated that they might be paraphyletic with regard to Scalibregmatidae. Our cladistic analyses of the 18S rDNA gene across those polychaetes traditionally assigned as “Sedentaria” also included representatives of all three opheliid groups and the Scalibregmatidae. The dataset was analysed with Maximum Parsimony (MP) and Maximum Likelihood (ML); node support has been validated with bootstrap methods and decay index for MP. The results of this analysis give strong support (100% bootstrap-value) for a clade consisting of scalibregmatids and Travisa. This result supports Blake’s assumption of a close relationship between both taxa (Blake 2000 In: Taxonomic atlas of the benthic fauna of the Santa Maria Basin and the Western Santa Barabara Channel Vol. 7 (4), Santa Barabara, California). A clade of Opheliinae + Ophelininae is also supported (99% bootstrap-value). After removing Travisia from the Opheliidae it has to be tested whether lateral grooves and a strong ventral groove might support the monophyly of the remaining Opheliidae. A rugose epidermis represents a possible autapomorphy for a taxon consisting of the Scalibregmatidae and Travisia.

Supported by the Deutsche Forschungsgemeinschaft (Ba 1520/4-1)

A contribution to sedentary polychaete phylogeny using 18S rDNA sequence data

C. Bleidorn, L. Vogt and T. Bartolomaeus*

Zoomorphologie & Systematik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany. E-mail:christoph.bleidorn@biologie.uni-bielefeld.de

** Systematik und Evolution der Tiere, Institut für Biologie, FU Berlin, Königin-Luise-Straße 1-3, 14-195 Berlin, Germany.*

eingereicht in:

Journal of Zoological Systematics and Evolutionary Research

Abstract

The phylogenetic position of the Annelida as well as their ingroup relationships are a matter of ongoing debate. A molecular phylogenetic study of sedentary polychaete relationships was conducted based on 70 sequences of 18S rDNA, including unpublished sequences of 18 polychaete species. The data set was analyzed with maximum parsimony and maximum likelihood methods. Clade robustness was estimated by parsimony-bootstrapping and -jackknifing, decay index, and clade support, as well as posteriori probabilities, which were calculated by a Bayesian inference. Irrespective of the applied method, some traditional sedentary polychaete taxa, as the Cirratulidae, Opheliidae, Orbiniidae, Siboglinidae, and Spionidae, were recovered by our phylogenetic reconstruction. A close relationship between Orbiniidae and *Questa* received a particularly strong support. As an unexpected result, Echiura appears to be a polychaete ingroup taxon which is closely related to *Dasybranchus* (Capitellidae). In correspondence to previous molecular analyses no support was found for the monophyly of Annelida nor for that of Polychaeta. However, our inference suggests that an increase in taxon sampling may yield additional resolution in the reconstruction of polychaete ingroup phylogeny.

Key words: Annelida – Polychaeta – phylogeny – 18S rDNA – Echiura - *Questa* - Bayesian inference

Introduction

The phylogenetic position of the Annelida is still a matter of debate (Schmidt-Rhaesa et al. 1998; Westheide et al. 1999; Wägele & Miesof 2001). The quest for the sister group of the Annelida is intimately connected to the question for the interrelationships within the Annelida, which are largely unresolved (Rouse & Fauchald 1997, 1998; Westheide 1997; Bartolomaeus 1998), as well as to the question which taxa are to be included to the Annelida. According to traditional classifications the Annelida consist of Polychaeta and Clitellata. While the latter is clearly supported as a monophylum in both, molecular and morphological analyses, the status of the Polychaeta is still controversely discussed (Rouse & Fauchald 1998, Westheide et al. 1999). Cladistic analyses of the classical morphological data recognize the Polychaeta as a monophylum (see Rouse & Pleijel 2001), while scenario based analyses regard them as paraphyletic (Westheide 1997). We are convinced that the attempt to bring a solution to this debate has to start with phylogenetic analyses on lower taxonomic levels. This is because the Annelida are highly diverse and represent an evolutionary old taxon that presumably radiated at or before the Precambrian-Cambrian border (Butterfield 1990; Conway-Morris & Peel 1995).

Diversity in annelids is expressed by tremendous morphological differences that become obvious when comparing different representatives of the polychaetes. In the past, therefore, taxonomists created up to 24 “orders” within this taxon (Fauchald 1977). As an “order” is a relatively high taxonomic rank, their large number likewise expressed the tremendous morphological differences between different polychaete taxa. A few of them could be substantiated as monophyletic groups by morphological data, but most of them, however, have never attained such support. Thus, more recent literature divides the polychaetes into more than 80 taxa, giving them the classical rang of families (Rouse & Fauchald 1997, Rouse 1999, Rouse & Pleijel 2001). However, there is still some doubt, whether all of them are monophyletic (Fauchald & Rouse 1997).

Recent detailed analyses of the ultrastructure and of the formation of certain organs could substantiate the hypothesis of monophyly for certain annelid taxa and provide evidence for a closer relationship among some of them (Meyer & Bartolomaeus 1996;

Purschke & Tzetlin 1996; Purschke 1997; Bartolomaeus 1998; Hausen & Bartolomaeus 1998; Hausam & Bartolomaeus 2001; Hausen 2001). These analyses also provided first evidence that the Pogonophora and Vestimentifera are sister taxa to certain polychaetes, supporting the hypothesis that both are derived polychaete taxa (Kojima et al. 1993, Bartolomaeus 1995, Rouse & Fauchald 1995, 1997, Black et al 1997, McHugh 1997). Echiura seem to be another candidate for a possible inclusion into the Annelida (McHugh 1997, Ax 2000, Hessling & Westheide 2002).

Recent molecular frameworks (Koijima 1998, Brown et al. 1999, McHugh 2000) could not provide any evidence for a probable monophyly of the Polychaeta or even of the Annelida. The results of theses surveys were conflicting with regard to the position of the traditional polychaete families, irrespective of the molecules these studies were based on. Merely the Pogonophora and Vestimentifera turned out to be monophyletic (Halanych et al. 2001), supported by statistically significant bootstrap values. No such support was provided in most of the analyses for the remaining polychaete entities.

If one compares these analyses, it becomes obvious that despite of their comparatively low diversity a relative large amount of molecular data exists for the Pogonophora (Halanych et al. 2001). A similar observation can be made for the Clitellata (Martin 2001, Siddall et al. 2001). Compared to the diversity in polychaetes, sequences of only a few polychaete taxa are available and in most cases only a single representative stands for a larger taxon (Kojima 1998, McHugh 2000). It therefore does not surprise that even those taxa that can be supported as monophyla on the basis of morphological data do not form a single clade in molecular analyses. We are strongly convinced that the low number of polychaete species analyzed causes the low resolution and want to show that an increased taxon sampling increases the resolution, so that molecular data can be used to substantiate the monophyly of certain polychaeta taxa. In this study we want to investigate the usefulness of the 18S rDNA for the inference of polychaete phylogeny and, therefore, we analyzed the 18S rDNA sequences of additional hemisessile and sessile taxa. Simulation studies have shown that an increase of the taxon sampling can improve the resolution of the phylogentic signal in a data set (Graybeal 1998). Special emphasis was laid in this study on possible monophyletic taxa within the polychaeta. In accordance to our assumption on the influence of taxon sampling on the resolution, we do also expect to gain evidence for possible sister group relationships.

Material and methods

Taxon sampling

Eighteen species of polychaetous annelids and one sipunculid species were collected from various sites (see table 1) and the complete 18S rDNA-sequence of each species was analyzed. For alignment and phylogenetic analyses a total of 70 metazoan 18S rDNA sequences, including nearly all available polychaete 18S rDNA sequences, were chosen from GenBank (see table 2). The sequence of *Capitella capitata* (U67323) was excluded from the analysis, as a reanalysis of this sequence (unpublished) suggests that U67323 is erroneous and probably from a misidentified specimen.

Table 1: Accessions of taxa sequenced for the present study.

Species	18S rDNA (bp)	Collection site	GenBank Accesssion-numbers
Polychaeta			
Aonides oxycephala	1735	Concarneau, France	AF448149
Apistobranchus typicus	1814	Qeqertarsuaq, Greenland	AF448150
Caulleriella parva	1807	Concarneau, France	AF448151
Clymenura clypeata	1805	Concarneau, France	AF448152
Dasybranchus caducus	1819	Concarneau, France	AF448153
Dodecaceria atra	1804	Concarneau, France	AF448154
Eteone longa	1814	Sylt, Germany	AF448155
Ophelia neglecta	1804	Concarneau, France	AF448156
Ophelia rathkei	1815	Sylt, Germany	AF448157
Ophelina acuminata	1681	Helgoland, Germany	will be handed in
Orbinia bioreti	1828	Concarneau, France	AF448158
Orbinia latreilii	1847	Concarneau, France	AF448159
Owenia fusiformis	1809	Concarneau, France	AF448160
Polyophthalmus pictus	1811	Banyuls-sur-mer, France	AF448161
Proscoloplos cygnochaetus	1965	Roscoff, France	AF448162
Scalibregma inflatum	1833	Helgoland, Germany	AF448163
Scolelepis squamata	1848	Sylt, Germany	AF448164
Telepsavus spec.	1814	Concarneau, France	AF448165
Sipunculida			
Sipunculus nudus	1817	Arcachon, France	AF448166

Table 2: List of sequences retrieved from GenBank

Species	18S rDNA (bp)	Source	GenBank Accesssion-numbers
Brachiopoda			
Lingula anatina	1813	GenBank	X81631
Phoronida			
Phoronis australis	1767	GenBank	AF119079

Bryozoa			
Plumatella repens	1813	GenBank	U12649
Kinorhyncha			
Pycnophyes kielensis	1806	GenBank	U67997
Priapulida			
Priapulus caudatus	1750	GenBank	AF025927
Nematomorpha			
Gordius aquaticus	1799	GenBank	X80233
Arthropoda			
Lepisma saccharina	1828	GenBank	X89484
Limulus polyphemus	1787	GenBank	U91490
Nebalia spec.	1805	GenBank	L81945
Mollusca			
Aplysia spec.	1826	GenBank	X94268
Lepidochitona corrugata	1821	GenBank	X91975
Ostrea edulis	1821	GenBank	L49052
Kamptozoa			
Barentsia hildegardae	1759	GenBank	AJ001734
Nemertini			
Prostoma eilhardi	1834	GenBank	U29494
Sipunculida			
Aspidosiphon misakiensis	1766	GenBank	AF119090
Phascolosoma granulatum	1841	GenBank	X79874
Themiste alutacea	1753	GenBank	AF119075
Echiurida			
Ochetostoma erythrogrammon	1814	GenBank	X79875
Urechis caupo	1772	GenBank	AF119076
Annelida			
Clitellata			
Enchytraeus sp.	1831	GenBank	Z83750
Glossiphonia spec.	1890	GenBank	Z83751
Hirudo medicinalis	1891	GenBank	Z83752
Lumbricus terrestris	1813	GenBank	AJ272183
Polychaeta			
Aphrodita aculeata	1810	GenBank	Z83749
Chaetopterus variopedatus	1692	GenBank	U67324
Dinophilus gyrociliatus	1784	GenBank	AF119074
Dodecaceria concharum	1701	GenBank	U50967
Glycera americana	1814	GenBank	U19519
Harmothoe impar	1736	GenBank	U50968
Lanice conchilega	1816	GenBank	X79873
Magelona mirabilis	1728	GenBank	U50969
Nephtys hombergii	1764	GenBank	U50970
Nereis virens	1814	GenBank	Z83754
Paralvinella palmiformis	1752	GenBank	AF168747
Polydora ciliata	1684	GenBank	U50971
Proceraea cornuta	1839	GenBank	AF212179
Protula sp.	1749	GenBank	U67142

Pygospio elegans	1758	GenBanK	U67143
Questa paucibranchiata	1788	GenBank	AF209464
Sabella pavonina	1726	GenBank	U67144
Scoloplos armiger	1769	GenBank	U50972
Siboglinidae			
Escarpia spicata	1764	GenBank	AF168741
Galathealinum brachiosum	1820	GenBank	AF168738
Lamellibrachia barhami	1759	GenBank	AF168742
Oasisia alvinae	1764	GenBank	AF168743
Polybrachia sp.	1820	GenBank	AF168739
Riftia pachyptila	1765	GenBank	AF168745
Ridgeia piscesae	1828	GenBank	X79877
Siboglinum fiordicum	1844	GenBank	X79876
Spirobrachia sp.	1754	GenBank	AF168740
Tevnia jerichonana	1763	GenBank	AF168746

DNA extraction

Collected specimens were identified and then preserved in 100% ethanol for later extraction. Genomic DNA was extracted from specimens using Qiagen DNeasy™ Tissue Kit.

PCR amplification, purification and sequencing

PCR amplification of the 18S rDNA gene was performed in three overlapping fragments of ~900bp each or in a whole with modificated primer pairs from Giribet et al. (1996) by using standard cycle sequencing protocols. Amplification reaction mixtures for 18S rDNA contained 25 ìl Qiagen Taq PCR Master Mix, 2 ìl Template-DNA, 4 ìl of each primer and 15 ìl H_2O. Amplifications were carried out using an Eppendorf Mastercycler gradient. The following PCR temperature file was used: 95 C for 3 min; 35 cycles with 94 C for 35 seconds, 45-50 C for 45 seconds to 1 min, and 72 C for 1 min; final extension at 72 C for 10 min. After detection by gel electrophoresis the products were purified with the Qiaquick PCR Purification Kit (Qiagen). Sequencing of all amplified fragments in both directions was carried out by the *IIT* Biotech/Bioservice of the University of Bielefeld. Overlapping fragments of the 18S rDNA were combined by using BioEdit (Hall 1999). Disagreement among these fragments was corrected by reference to the original chromatograms. All sequences were submitted to Genbank (for accession numbers see table 1).

Sequence Alignment

All sequences were aligned by using CLUSTAL W (Thompson et al. 1994) and subsequently manually edited by eye using BioEdit (Hall 1999). Gap positions and regions that could not be aligned unambiguously were excluded from the analysis.

Data Analysis

All phylogenetic analyses were conducted with PAUP* version 4.0b8 (Swofford 2001). A chi-square test of homogeneity of base frequencies across taxa was performed. The program TreeView (Page 1996) was used for tree visualization. All trees were rooted *a posteriori* the analysis using the sequence of *Gordius aquaticus* (Nematomorpha).

Maximum parsimony and clade support

All maximum parsimony searches were run with 1000 random addition replicates, heuristic search option with tree-bisection-reconnection (TBR) branch swapping, holding one tree per step, and keeping all most-parsimonious trees. Two separate analyses, using two different weighting schemes were conducted. In the first analysis, all transformations were weighted equally, in the second analysis, transversions were weighted three times as much as transitions.

Bootstrap as well as Jackknife analyses were performed (Felsenstein 1985; Farris 1997) and the bootstrap and jackknife values were determined from 1000 replicates subject to full heuristic searches with simple taxon addition to provide measures of relative clade support. Additionally, decay analyses (Bremer 1994) were performed for selected clades. The decay indices were estimated with converse constraint heuristic searches based on 100 random sequence addition replicates (Baum et al., 1994). We evaluated whether the most parsimonious (MP) trees that include the selected clades are significantly better supported than trees that lack them (Whitlock & Baum, 1999; Lee, 2000). This was achieved by comparing the pool of MP trees from a converse constraint search with the unconstrained MP trees using a Wilcoxon signed-rank test (Templeton, 1983) as implemented in PAUP*. For each clade, the P value (clade significance) reported is the highest obtained across the pair-wise comparisons. A clade is considered significantly supported if $P < 0.1$* (Lee, 2000).

Maximum Likelihood

For estimating the appropriate model of sequence evolution, different models were tested using the program modeltest version 3.06 (Posada & Crandall 1998, 2001). Both test criteria (hLRT and AIC) indicate that the Tamura Nei substitution model (Tamura & Nei, 1993) with equal base frequencies, invariant sites and gamma distribution (TrNef+I+Ã) represents the optimal model in respect to the data.

A maximum likelihood analysis was performed under the likelihood settings suggested by the result of the modeltest using the heuristic search option with TBR branch swapping and simple sequence addition.

Bayesian inference

For Bayesian analysis of the data set we used MrBayes 2.01 (Huelsenbeck & Ronquist 2001). The ML parameters in MrBayes were set as follows: "lset nst=6" (GTR model), "rates=invgamma" and "basefreq=estimate". Each Markov chain, three heated and one cold, was started from a random tree and all four chains ran simultaneously for $2.5 * 10^6$ generations, with trees being sampled every 100 generations for a total of 25,000 trees. After the likelihood of the trees of each chain converged, we discarded the first 5,000 trees as burnin. The majority-rule consensus tree containing the posterior probabilities of the phylogeny was determined from 20,000 trees.

Results

Sequence data

The alignment of the 70 18S rDNA sequences resulted in 2,207 positions. After the exclusion of ambiguous sites, the remaining 1,519 positions were taken on into our data matrix. Overall, the data matrix consists of 865 variable positions (57%), of which 588 positions are parsimony informative (39%). Since the chi-square test of homogeneity of base frequencies across taxa resulted in no significant *P*-values (chi-square=89.1117, df=207, *P*=1.0), assuming that compositional bias has no effect on the recovery of phylogenetic signal seems justifiable.

Maximum parsimony and clade support

The equal weighted parsimony analysis resulted in 33 MP trees (length=4,590; consistency index [CI]=0.3316; consistency index excluding uninformative characters [CI']=0.2798; retention index [RI]=0.4170). The strict consensus tree from this analysis together with the bootstrap and jackknife frequencies is illustrated in Fig. 1. Decay indices and clade significance values for selected groups are shown in Fig. 2.

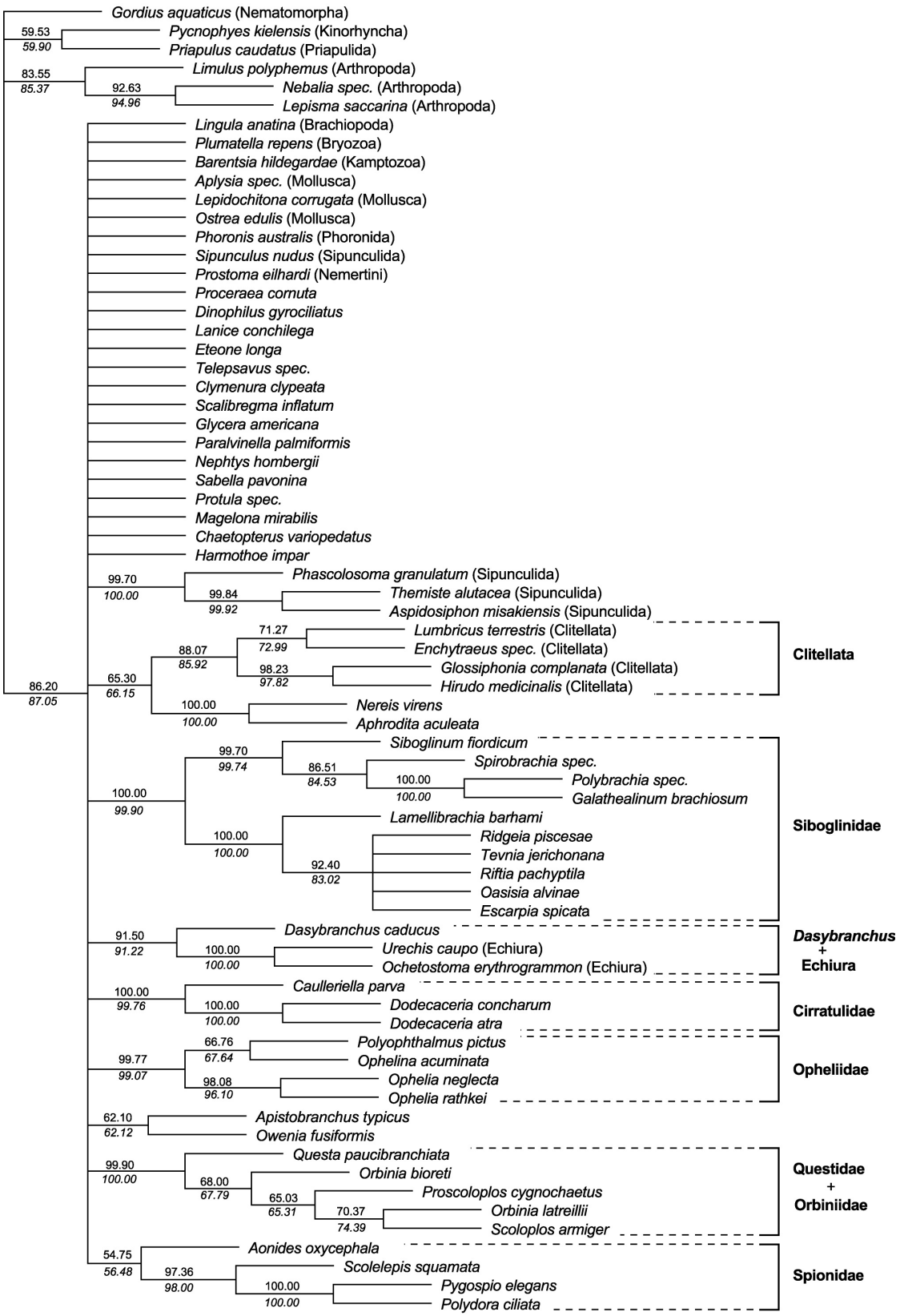

Figure 1: Strict consensus of 33 most-parsimonious trees of the equally weighted parsimony analysis (length= 4,590 steps; CI=0.33; RI=0.42). Bootstrap- and jackknife-frequencies are given above and below the branches.

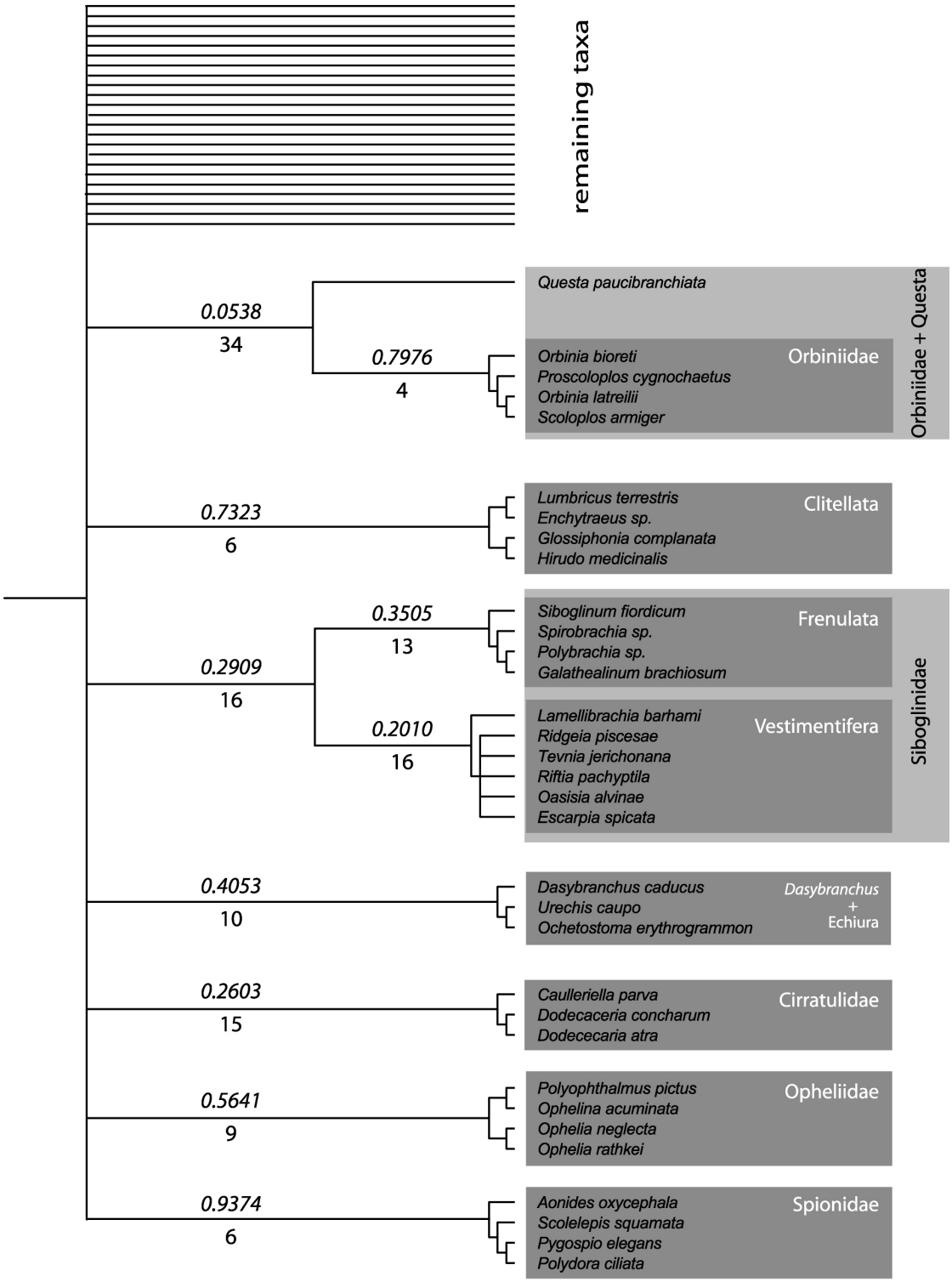

Figure 2: Phylogenetic relationship of selected groups (based on the strict consensus of the equally weighted parsimony analysis), with clade significance values and decay indices above and below the branches.

Some of the traditional polychaete-"families" represent well supported clades: **Opheliidae** (Bootstrap [BT]=99.8%; Jackknife [JK]=99.1%; decay index [DI]=9; clade

significance [*P*]=0.5641), **Cirratulidae** (BT=100.0%; JK=99.8%; DI=15; *P*=0.2603), **Spionidae** (BT=54.7%; JK=56.5%; DI=6; *P*=0.9374), **Orbiniidae** (BT=68.0%; JK=67.8%; DI=4; *P*=0.7976). Whereas *Orbinia* seems to be paraphyletic, since *Orbinia latreillii* appears to be closer related to other orbiniids than to *Orbinia bioreti* (*Orbinia latreillii* + *Scoloplos armiger*: BT=70.4, JK=74.4; *Orbinia latreillii* + *Scoloplos armiger* + *Proscoloplos cygnochaetus*: BT= 65.0, JK=65.3). The strongest support among all clades receives a relationship between the **Orbiniidae** and ***Questa*** (BT=99.9%; JK=100.0%; DI=34; *P*=0.0538*). A clade consisting of the **Echiura** and ***Dasybranchus caducus*** (Capitellidae) is also supported (BT=91.5%; JK=91.2%; DI=10; *P*=0.4053). Furthermore, as the results of the analysis of Halanych *et al.* (2001) already have shown, the **Siboglinidae** (BT=100.0%; JK=99.9%; DI=16; *P*=0.2909) are well supported. They consist of the two well supported sistergroups **Vestimentifera** (BT=100.0%; JK=100.0%; DI=16; *P*=0.2010), and the **Frenulata** (BT=99.7%; JK=99.7%; DI=13; *P*=0.3505). In concordance with the traditional view of annelid systematics, the **Clitellata** (BT=88.1%; JK=85.9%; DI=6; *P*=0.7323) are also well supported.

The unequally weighted parsimony analysis yielded in four MP trees. All trees differ in detail from the results of the equally weighted parsimony analysis. However, most of the well supported clades described above are recovered in both analyses. Only the **Spionidae** and **Orbiniidae** receive no support above the 50%-level in BT and JK from the unequally weighted analysis. Nevertheless, further on, a monophylum consisting of the **Orbiniidae** and ***Questa*** receives high support.

Comparing the two results of the bootstrap and the jackknife analyses, all yield in a high resolution for the relationships within those well supported groups, thereby exhibiting almost no conflicting evidence. In contrast to these findings are all other relationships generally weakly supported and many even do not reach the 50% support level.

Maximum Likelihood

The Tamura Nei substitution model (Tamura and Nei 1993) with equal base frequencies, invariant sites and gamma distribution (TrNef+I+Ã) represents the best fitting model for an explanation of the data of all the models that were considered in the modeltest. The most likely tree has a log-likelihood value of -22747.98730 and is illustrated in Fig. 3. The groups mentioned above are also supported by the likelihood

analysis. Their ingroup topologies are congruent to those found in the consensus tree of the equally weighted parsimony analysis.

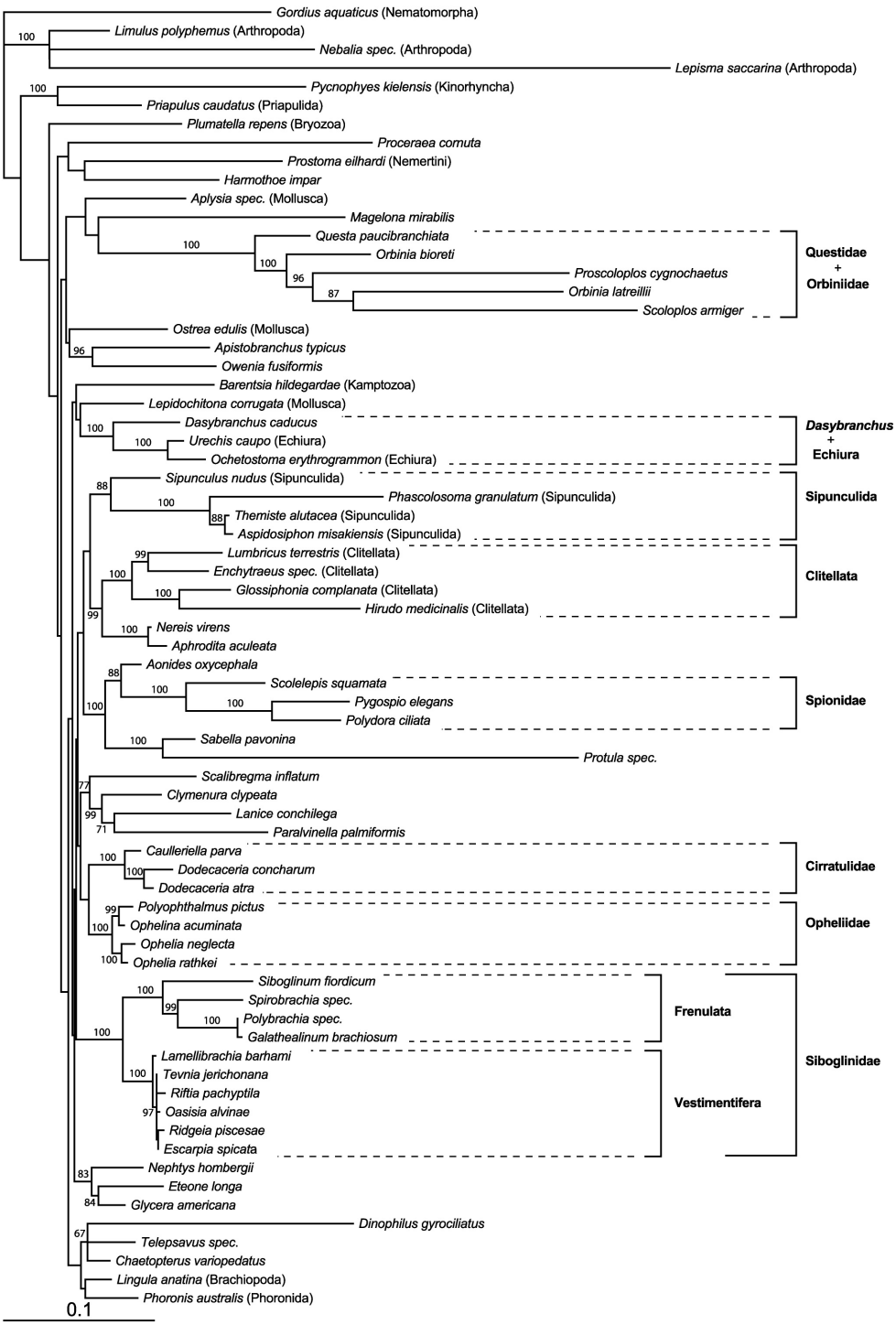

Figure 3: Maximum likelihood tree based on the TrNef+É+Ã model of sequence evolution (-log*L*=22747.9873). The posteriori probabilities of the Bayesian analysis are given above the branches.

Bayesian inference

The results of the Bayesian inference are presented as posteriori probabilities on the branches of the most likely tree (Fig. 3). Most obvious is, that all clades with a high bootstrap support are also strongly supported in the Bayesian analysis. In concordance with Whittingham et al. (2002), we found that some clades with a low bootstrap support are nevertheless well supported in the Bayesian analysis. For example, the clade formed by the Spionidae, which is supported by a low bootstrap value of 54,7%, is supported with a Bayesian probability of 88%. Another example are the Orbiniidae, who receive a support of 68% in the bootstrap analysis and a Bayesian probability of 100%. There is no generally accepted procedure of interpreting these findings, since the nature of posteriori probabilities is not well enough understood yet (Leaché & Reeder 2002). However, it seems to exists a general tendency of Bayesian analyses to yield higher support values than bootstrap analyses do (see also Buckley et al. 2002), rendering this method a less conservative test.

Discussion

Like already shown for other taxa (Blaxter et al. 1998), extended taxon sampling helps to increase the resolution of comparative 18S rDNA sequence analysis. Provided that enough sequences are available, this molecule seems to be suitable for the inference of some aspects of polychaete phylogeny. Insofar, our initial assumption has been confirmed. However, our analysis also reflects the problems that arise when attempting to infer evolutionary events that took place during or before the Cambrium, which especially appears to apply to the 18S rDNA sequence data (Abouheif et al. 1998). Our data show that the phylogenetic information content of the 18S rDNA is not sufficient for giving significant support for neither the hypothesis of the monophyly of the Annelida nor for that of the Polychaeta. This low resolution may be due to a rapid radiation of the Annelida that has intensely been discussed elsewhere (Brown et al. 1999, Rota et al. 2001) and is generally credited to an erosion of information during time. Provided that such an explosive radiation occurred in annelids, it also influenced morphological characters leading to the known problems in tree reconstruction (Rouse & Fauchald 1997). If these difficulties actually hint at a rapid radiation, at least some of the groups we analyzed must have radiated more recently. In addition to this, mutational events that occurred in certain stem lineages turn out to be rather conservative, as these

taxa are highly supported as monophyletic in our analyses. Increasing the taxon sampling of 18S rDNA sequences of different polychaete groups thus indicates a possible solution of yet unsolved questions in annelid phylogeny by reconstructing polychaete ingroup relationships.

Our analysis confirms that the pogonophoran and vestimetiferan species cluster in a single clade, representing the taxon Siboglinidae (see Mc Hugh 2000, Halanych et al. 2001). According to morphological analyses Siboglinidae represent a subordinate polychaete taxon (Bartolomaeus 1998, Rouse & Fauchald 1997). Reduction of the gut lumen during development and persistence of its cells to house endosymbiontic bacteria, as well as an extremely elongated first segment are strong arguments derived from morphological analyses which support the monophyly of this Siboglinidae. The Clitellata also form a monophyletic group when 18S rDNA and other molecular data sets are analyzed (Rota et al. 2001, but see Martin 2001). A large number of morphological characters, like hermaphroditism, restriction of gonads to the anterior segments, sperm ultrastructure, modified and direct development, re-location of the brain from the prostomium into a more posterior position (Ferraguti 1984, Purschke et al. 1993, Rouse & Fauchald 1997) support the monophyly of this taxon. A specific glandular region posterior to the gonads, the clitellum, which produces a cocoon that encloses the eggs, also is apomorphic for clitellates.

Formation of clitellar material by a special glandular region and the restriction of gonads to a few segments is also characteristic for the Questidae and lead to the hypothesis of a questid-clitellate relationship (Giere & Riser1981). Subsequent studies, however, argued against such a position of the aberrant taxon *Questa* (Jamieson & Webb 1984, Rouse & Fauchald 1997, Giere & Erseus 1998). So far, their phylogenetic relationships remained uncertain. Our analysis, which is the first to entail several orbiniid sequences together with a sequence of *Questa*, provides strong evidence for the position of aberrant *Questa* as being closely related to the Orbiniidae (supporting Rota et al. 2001).

The support of the monophyly of Orbiniidae is highly dependend on the choice of method. Although Bayesian probalities yield high support (100%), this taxon is not supported by bootstrap analysis of the unequally weighted parsimony analysis (<50%). Elevated parapodia in the posterior body region are regarded as a morphological

autapomorphy for this taxon (Fauchald & Rouse 1997). However, this character is not present in all orbiniid taxa (Rouse & Pleijel 2001) and so the knowledge of ingroup relationships is essentiall for a correct phylogenetic interpretation of this character.

Beside this, all available 18S rDNA sequences of the cirratulids cluster together in a highly supported clade. The same is true for the opheliids and spionids. These results support the monophyly hypothesis gained from morphological data, and thus support traditional taxa within the polychaetes.

A final and very interesting result concerns the position of the Echiura. Most textbooks regard them as taxon outside the Annelida. Based on alpha 1 elongation factor sequences, McHugh (1997), however, provided some evidence that the Echiura belong to the Annelida. The only morphological feature that might support inclusion of the Echiura into the Annelida are the chaetae (see Ax 2000). However, our analysis now provides evidence for a common ancestry of the Echiura and Capitellidae and this result is supported with a high bootstrap-value (91.5%) and a Bayesian probality of 100%. We have to await for additional morphological and molecular data to severely test this hypothesis.

In this study we could confirm our initial assumption that an increased taxon sampling increases the resolution of the 18S rDNA data sequences in annelids. As this group is an evolutionary old group, we are sure that 18SrDNA sequences from additional species as well as further molecules will help to resolve polychaete phylogeny.

Zusammenfassung

Beitrag zur Phylogenie sedentärer Polychaeten unter Verwendung von 18S rDNA Sequenzdaten

Die Stellung der Anneliden im phylogenetischen System und die Verwandtschaftsbeziehungen ihrer Innengruppentaxa sind Gegenstand aktueller Diskussionen. Eine molekulare Analyse der Phylogenie sedentärer Polychaeten wurde unter Verwendung von 70 Sequenzen der 18S rDNA durchgeführt, unter denen sich bisher unpublizierte Sequenzen von insgesamt 18 Polychaeten-Arten befinden. Der Datensatz wurde mittels Maximum Parsimonie- und Maximum Likelihood-Methoden analysiert. Die Stabilität der einzelnen Knotenpunkte wurde unter Verwendung von Parsimonie-*Bootstrapping* und *Jackknifing*, *Decay Index* und *Clade Robustness* getestet; a posteriori-Wahrscheinlichkeiten wurden mit Hilfe einer Bayesianischen Analyse ermittelt. Unabhängig von der benutzten Methode, fanden traditionelle Taxa sedentärer Polychaeten, wie die Cirratulidae, Opheliidae, Orbiniidae, Siboglinidae und Spionidae bei der phylogenetischen Rekonstruktion eine Unterstützung. Eine besonders starke Unterstützung findet sich für eine nahe Verwandtschaft der Orbiniidae mit *Questa*. Unerwarteter Weise wird eine Stellung der Echiuren als Innegruppentaxon der Polychaeten mit einer nahen Verwandtschaft zu *Dasybranchius* (Capitellidae) unterstützt. Für die Monophylie der Anneliden und der der Polychaten hingegen lassen sich keine Hinweise aus dem Datensatz entnehmen; dieses Ergebnis steht im Einklang mit den bisher zu dieser Fragestellung veröffentlichten molekularen Analysen. Es zeigt sich jedoch, dass eine Erhöhung des *taxon samplings* zu einer besseren Auflösung bei der Rekonstruktion der Verwandtschaftsbeziehungen innerhalb der Polychaeten beitragen kann.

References

Abouheif, E.; Zardoya, R.; Meyer, A., 1998: Limitations of metazoan 18S rRNA sequence data: Implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. J. Mol. Evol. **47**, 394-405.

Ax, P., 2000: Multicellular animals. The phylogenetic system of the Metazoa. Volume II. Berlin: Springer Verlag. pp. 1-396.

Bartolomaeus, T., 1995: Structure and formation of the uncini in larval *Pectinaria koreni*, *Pectinaria auricoma* (Terebellida) and *Spiorbis spirobis* (Sabellida): Implications for annelid phylogeny and the position of the Pogonophora. Zoomorphology **115**, 161-177.

Bartolomaeus, T., 1998: Chaetogenesis in polychaetous Annelida – significance for annelid systematics and the position of the Pogonophora. Zoology **100**, 348-364.

Baum, D.A.; Sytsma, K.J.; Hoch, P.C., 1994: A phylogenetic analysis of *Epilobium* (Onagraceae) based on nuclear ribosomal DNA sequences. Syst. Bot. **19**, 363-388.

Black, M.B.; Halanych, K.M.; Maas, P.A.Y.; Hoeh, W.R.; Hashimoto, J.; Desbruyeres, D.; Lutz, R.A.; Vrijenhoek, R.C., 1997: Molecular systematics of vestimentiferan tubeworms from hydrothermal vents and cold-water seeps. Mar. Biol. **130**, 141-149.

Blaxter, M.L.; De Ley, P.; Garey, J.R.; Liu, L.X.; Scheldemann, P.; Vierstraete, A.; Vanfleteren, J.R.; Mackey, L.Y.; Dorris, M.; Frisse, L.M.; Vida, J.T.; Kelley Thomas, W., 1998: A molecular evolutionary framework for the phylum Nematoda. Nature **392**, 71-75.

Bremer, K., 1994: Branch support and tree stability. Cladistics **10**, 295-304.

Brown, S.; Rouse, G.W.; Hutchings, P.; Colgan, D., 1999: Assessing the usefulness of histone H3, U2 and snRNA and 28S rDNA in analyses of polychaete relationships. Aust. J. Zool. **47**, 499-516.

Buckley, T.R.; Arensburger, P.; Simon, C.; Chambers, G.K., 2002: Combined data, Bayesian inference, and the origin of the New Zealand Cicada genera. **51**, 4-18.

Butterfield, N.J., 1990: A reassesment of the enigmatic Burgess Shale fossil Wiwaxia corrugata (Matthew) and its relationship to the polychaet Canadia spinosa Walcott. Paleobiology **16**, 287-303.

Conway-Morris, S.; Peel, J.S., 1995: Articulated halkieriids from the lower Cambrian of north Greenland and their role in early protostome evolution. Phil. Tran. R. Soc. Lond. B **247**, 305-358.

Farris, J.S., 1997: The future of phylogeny reconstruction. Zoologica Scripta **26**, 303-311.

Fauchald, K., 1977: The polychaete worms. Definitions and keys to the orders, families and genera. Nat. Hist. Mus. LA County Sci. Ser. **28**, 1-188.

Fauchald, K.; Rouse, G.W., 1997: Polychaete systematics: past and present. Zool. Scr. **26**: 71-138.

Felsenstein, J., 1985: Confidence limits on phylogenys: an approach using the bootstrap. Evolution **39**, 783-791.

Ferraguti, M., 1984: The comparative ultrastructure of sperm flagella central sheath in Clitellata reveals a new autapomorphy of the group. Zool Scr. **13**, 201-207.

Giere, O; Riser, N.W., 1981: Questidae – Polychaetes with oligochaetoid morphology and development. Zool. Scr. **10**, 95-103.

Giere, O.; Erseus, C., 1998: A systematic account of the Questidae (Annelida, Polychaeta), with description of new taxa. Zool. Scr. **27**, 345-360.

Giribet,G.; Carranza,S.; Baguna,J.; Riutort,M.; Ribera,C., 1996: First molecular evidence for the existence of a Tardigrada + Arthropoda clade. Mol. Biol Evol. **13**, 76-84.

Graybeal, A.,1998: Is it better to add taxa or characters to a difficult phylogenetic problem? Syst. Biol. **47**, 9-17.

Halanych, K.M.; Feldman, R.A.;Vrijenhoek, R.C., 2001: Molecular evidence that *Sclerolinum brattstromi* is closely related to Vestimentiferans, not to frenulate Pogonophorans (Siboglinidae, Annelida). Biol. Bull. **201**, 65-75.

Hall, T.A., 1999: BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT.Nucl. Acids. Symp. Ser. **41**, 95-98.

Hausam, B.; Bartolomaeus, T., 2001: Ultrastructure and development of forked and capillary setae in *Orbinia bioreti* and *Orbinia latreillii* (Annelida: Orbiniidae). Inv. Biol. **120**, 13-28.

Hausen, H., 2001: Untersuchungen zur Phylogenie “spiomorpher” Polychaeten (Annelida). Berlin: Logos Verlag. pp. 1-145.

Hausen, H.; Bartolomaeus, T., 1998: Setal structure and chaetogenesis in Scolelepis squamata and Malacoceros fuliginosus (Spionida, Annelida). Acta Zool. **79**, 146-161.

Hessling, R.; Westheide, W., 2002: Are Echiura derived from a segmented ancestor? - Immunohistochemical analysis of the nervous system in developmental stages of *Bonellia viridis*. J. Morphol. (in press).

Huelsenbeck, J.P.; Ronquist, F.R., 2001: MrBayes: Bayesian inference of phylogeny. Biometrics **17**, 754-755.

Jamieson, B.G.M.; Webb, R.I., 1984: The morphology, spermatozoal ultrastructure and phylogentic affinities of new species of questid (Polychaeta, Annelida). In P.A. Hutchings (ed.), *Proceedings of the 1st International Polychaete Conference, Sidney, Australia* (pp. 21-34). Sydney, Australia: Linnean Society of New South Wales.

Kojima, S., 1998: Paraphyletic status of Polychaeta suggested by phylogenetic analysis based on the amino acid sequences of elongation factor 1-alpha. Mol. Phyl. Evol. **9**, 255-261.

Kojima, S.; Hashimoto, T.; Hasegawa, M.; Murata, S.; Ohta, S.; Seki, H.; Okada, N., 1993: Close phylogenetic relationship between Vestimentifera (tubeworms) and Annelida revealed by the amino acid sequence of elongation factor-1a. J. Mol. Evol. **37**, 66-70.

Leaché, A.D.; Reeder, T.W., 2002: Molecular systematics of the eastern Fence Lizard (*Sceloporus undulatus*): a comparison of Parsimony, Likelihood, and Bayesian approaches. Syst. Biol. 51, 44-68.

Lee, M.S.Y., 2000: Clade robustness and significance. Syst. Biol. **49**, 829-836.

Martin, P., 2001: On the origin of the Hirudinea and the demise of the Oligochaeta. Proc. R. Soc. Lond. B **268**, 1089-1098.

McHugh, D., 1997: Molecular evidence that echiurans and pogonophorans are derived annelids. PNAS **94**, 8006-8009.

McHugh, D., 2000: Molecular phylogeny of the annelids. Can. J. Zool. **78**, 1873-1884.

Meyer, K.; Bartolomaeus, T., 1996: Ultrastructure and formation of the hooked setae in *Owenia fusiformis* delle Chiaje, 1842: Implications for annelid phylogeny. Can. J. Zool. **74**, 2143-2153.

Page, R.D.M., 1996: TREEVIEW: An application to display phylogenetic trees on personal computers. Comp. Appl. Biosci. **12**, 357-358.

Posada, D.; Crandall, K.A., 1998: Modeltest: testing the model of DNA substitution. Bioinformatics **14**, 917-918.

Posada, D.; Crandall, K.A., 2001: Selecting the best-fit model of nucleotid substitution. Syst. Biol. **50**, 580-601.

Purschke, G., 1997: Ultrastructure of the nuchal organs in polychaetes (Annelida) – new results and review. Acta Zool. **78**, 123-143.

Purschke, G.; Tzetlin, A.B., 1996: Dorsal ciliary folds in the polychaet foregut: structure, prevalence and phylogenetic significance. Acta Zool. **77**, 33-49.

Purschke G.; Westheide, W.; Rhode, D.; Brinkhurst, R.O., 1993: Morphological reinvestigation and phylogenetic relationship of *Acanthobdella peledina* (Annelida, Clitellata). Zoomorphology **113**, 91-101.

Rota, E.; Martin, P.; Erseus, C., 2001: Soil-dwelling polychaetes: enigmatic as ever? Some hints on their phylogenetic relationships as suggested by a maximum parsimony analysis of 18S rRNA gene sequences. Contributions to Zoology **70**, 127-138.

Rouse, G.W., 1999: Trochophora concepts: ciliary bands and the evolution of larvae in spiralian Metazoa. Biol. J. Lin. Soc. **66**, 411-464.

Rouse, G.W.; Fauchald, K., 1995: The articulation of annelids. Zool. Scr. **24**, 269-301.

Rouse, G.W.; Fauchald, K., 1997: Cladistics and polychaetes. Zool. Scr. **26**, 139-204.

Rouse, G.W.; Fauchald, K., 1998: Recent views on the status, delineation and classification of the Annelida. Amer. Zool. **38**, 953-964.

Rouse, G.W.; Pleijel, F., 2001: Polychaetes. Oxford: University Press, pp. 1-354.

Schmidt-Rhaesa, A.; Bartolomaeus, T.; Lemburg, C.; Ehlers, U.; Garey, J.R., 1998: The position of the Arthropoda in the phylogenetic system. J. Morph. **238**, 263-285.

Siddall, M.E.; Apakupakul, K.; Burreson, E.M.; Coates, K.A.; Erséus, C.; Gelder, S.R.; Källersjö, M.; Trapido-Rosenthal, H., 2001: Validating Livanow: Molecular data agree that leeches, branchiobdellidans, and *Acanthobdella peledina* form a monophyletic group of oligochaetes. Mol. Phyl. Evol. **21**, 346-351.

Swofford, D.L., 2001: PAUP*. Phylogenetic analysis using parsimony, version 4.0b8. Sinauer, Sunderland, MA.

Tamura, K., Nei, M., 1993 Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol **10**, 512-526.

Templeton, A.R., 1983: Phylogenetic inference from restriction endonuclease cleavage site maps with particular reference to the humans and apes. Evolution **37**, 221-244.

Thompson, J.D.; Higgins, D.G; Gibson, T.J., 1994: CLUSTAL W: improving the sensivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. **22**, 4673-4680.

Wägele, J.W.; Miesof, B., 2001: On quality of evidence in phylogeny reconstruction: a reply to Zravý's defence of the "Ecdysozoa" hypothesis. J. Zool. Syst. Evol. Research **39**, 165-176.

Westheide, W., 1997: The direction of evolution within the Polychaeta. J. Nat. Hist. **31**, 1-15.

Westheide, W., McHugh, D., Purschke, G.; Rouse, G., 1999: Systematization of the Annelida: different approaches. Hydrobiologia **402**, 291-307.

Whitlock, B.A.; Baum, D.A., 1999: Phylogenetic relationships of *Theobroma* and *Herronia* (Sterculiaceae) based on sequences of the nuclear gene VICILIN. Syst. Bot. **24**, 128-138.

Whittingham, L.A.; Slikas, B.; Winkler, D.W.; Sheldon, F.H., 2002: Phylogeny of the tree swallow genus, Tachycineta (Aves: Hirudinidae), by Bayesian analysis of mitochondrial DNA sequences. Mol. Phyl. Evol. 22, 430-441.

New insights into polychaete phylogeny (Annelida) inferred from 18S rDNA sequences

C. Bleidorn, L. Vogt and T. Bartolomaeus*

Zoomorphologie & Systematik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany. E-mail:christoph.bleidorn@biologie.uni-bielefeld.de

** Systematik und Evolution der Tiere, Institut für Biologie, FU Berlin, Königin-Luise-Straße 1-3, 14-195 Berlin, Germany.*

eingereicht in:

Molecular Phylogenetics and Evolution

Abstract

The ingroup relationships of polychaete annelids are still controversely discussed. A molecular phylogenetic study of sedentary polychaete relationships was conducted based on 94 sequences of 18S rDNA, including unpublished sequences of 13 polychaete species. The data set was analyzed with maximum parsimony and maximum likelihood methods. Clade robustness was estimated by parsimony-bootstrapping and -jackknifing, as well as posteriori probabilities, which were calculated by a Bayesian inference. No evidence has been found for possible monopyhly of Scolecida. In all analyses a placement of the Echiura as derived polychaete ingroup with a close relationship to the capitellids is confirmed. The orbiniids appear paraphyletic with regard to Questa. *Travisia* is transferred from Opheliidae to Scalibregmatidae and shows a close relationship to Scalibregma. The remaining opheliids include a yet undescribed cetnodrilid genus from Elba, whereas the other invetsigated ctenodrilid *Ctenodrilus serratus* a close affinity to the cirratulid genus *Dodecaceria* shows. A common ancestry of arenicolids and *Branchiomaldane vincenti* is confirmed by all analyses conducted; a sistergroup relationship between these taxa and the Maldanidae is recovered by only some analyses. Maximum Likelihood and Bayesian inference give evidence for the hypothesis that Myzostomida are derived polychaetes.

Key words: Annelida – Polychaeta – phylogeny – 18S rDNA – Echiura - Travisia - Branchiomaldane - Bayesian inference

Introduction

In the traditional classification polychaete annelids have been classified into over 80 families (Fauchald, 1977). The phylogenetic relationships of these polychaete taxa are matter of ongoing debates in recent papers on annelid morphology (Meyer and Bartolomaeus, 1996; Purschke and Tzetlin, 1996; Purschke, 1997; Rouse and Fauchald, 1997, 1998; Westheide, 1997; Bartolomaeus, 1998; Hausen and Bartolomaeus, 1998; Westheide *et al.*, 1999; Hausen, 2001). Rouse and Fauchald (1997) suggest that the Polychaeta comprise two major clades, the Scolecida and the Palpata. The Scolecida comprehend the traditionally as families classified polychaete taxa Arenicolidae, Capitellidae, Cossuridae, Maldanidae, Opheliidae, Orbiniidae, Paraonidae, Scalibregmatidae, and Questidae. This clade is weakly supported and the only known autapomorphies for such a taxon are the presence of two or more pygidial cirri and the possesion of similar rami by the parapodia. The highest ranked taxa within the Palpata are the Canalipalpata (comprising the remaining so called “Sedentaria”) and the Aciculata (formerly termed Errantia). This view was challenged by Bartolomaeus (1998) and Hausen (2001) who question the monophyly of Scolecida, Palpata and Canalipalpata. Instead they present evidence for the monophyly of a taxon comprising all polychaete taxa with hooked setae (sensu Bartolomaeus 1998) including the “scolecids” Arenicolidae, Maldanidae and Capitellidae, and the “canalipalpats” Oweniidae, Chaetopteridae, Sabellariidae, Terebellidae, Alvinellidae, Ampharetidae, Pectinariidae, Sabellidae, Serpulidae, and Siboglinidae, as well as the enigmatic Psammodrilidae.

Apart from the discussion about the phylogenetic relationships of polychaete taxa, it is also controversially discussid wether Echiura and Myzostomida represent derived polychaete taxa. Traditionally Echiura are regarded as closely related to the Annelida (Rouse & Fauchald 1995, 1997; Rouse 1999), but recent studies of the organisation of the nervous system (Hessling & Westheide 2002) and comparative analyses of molecular data (McHugh 1997, 1999; Brown et al. 1999) give evidence to include the Echiura into the Annelida, although the annelid sister taxon of the Echiura remains to be found. The Myzostomida are traditionally regarded as derived annelids (Nielsen 2000;

Rouse & Fauchald 1997), but analyses of molecular and morphological data show the opposite (see Zrzavy et al. 2001).

Within the Polychaeta Fauchald and Rouse (1997) and Rouse and Pleijel (2001) have shown that the monophyly and composition of some of the above mentioned polychaete families is obscure. The same holds true for the question of the monophyly and phylogenetic position of Ctenodrilidae, Opheliidae and Scalibregmatidae (Fauchald and Rouse, 1997; Blake, 2000a, 2000b; Rouse and Pleijel, 2001).

Whereas the above mentioned hypotheses on polychaete systematics are all achieved by analysing morphological data, in recent years many attempts were conducted to unravel polychaete relationships using molecular data (Kojima, 1998; Brown *et al*.; 1999; McHugh, 1997, 2000; Halanych *et al*., 2001; Rota *et al*., 2001; Struck *et al*., in press; Bleidorn *et al*. 2002). Whilst none of these analyses can be regarded as major breakthrough in polychaete systematics, some of these analyses have shown that traditional polychaete families are often well supported (Struck *et al*., 2002; Bleidorn *et al*., 2002). As molecular data turned out to be an important tool for solving problems on the relationships at lower taxonomic levels, we used 18S rDNA sequence data to study the phylogenetic relationships among polychaetes. Increased taxon sampling, particularly of "scolecid" taxa, gives the possibility to test the hypotheses on polychaete systematics developed through the cladistic analysis by Rouse and Fauchald (1997). Furthermore, the phylogenetic position of Ctenodrilidae, *Travisia* and *Branchiomaldane* is investigated, and evidence for the monophyly of Opheliidae, Scalibregmatidae and Cirratulidae is presented by analyzing our data set. Increasing the taxon sampling seems to be a promising method to find the sister taxon of the Echiura within the Annelida.

Material and Methods

Taxon sampling

Thirteen newly determined 18S rDNA sequences of several sedentary polychaete taxa were aligned together with 81 metazoan 18S rDNA sequences from GenBank, including nearly all avaible sequences of polychaete taxa from GenBank (see table 1). Collection sites for the newly sequenced annelids were as follows: Arcachon, France (*Arenicola*

marina, Ophelia bicornis), Buenos Aires, Argentina (*Protoariciella uncinata*, collected by Dr. Rodolfo Elias), Cape Town, South Africa (*Scoloplos (Leodamas) johnstonei*, collected by Bilke Hausam), Collioure, France (*Protoaricia oerstedii*), Concarneau, France (*Branchiomaldane vincenti, Lipobranchus jefreysii*), Helgoland, Germany (*Ctenodrilus serratus*), Kristinenberg, Sweden (*Travisia forbesii*), Roscoff, France (*Amphitritides gracilis*, *Notomastus latericeus*), Sylt, Germany (*Capitella capitata*, *Protodriloides symbioticus*). The following polychaete 18S rDNA sequences from GenBank have been excluded from the analysis because there exists evidence for their inaccuracy: *Pectinaria regalis* (AY040698), *Marphysia sanguinea* (AY040695), *Armandia maculata* (AY040681), *Nereis virens* (Z83754), *Aphrodite aculeata* (Z83749), *Capitella capitata* (U67323) and *Magelona mirabilis* (U50969).

Table 1: List of taxa used in the analysis with 18S rDNA sequence accession numbers.

Higher Taxon	Species	GenBank Accession No.
Annelida		
Aelosomatidae	*Aelosoma hemprichi*	AJ310500
	Aelosoma spec.	Z83748
Alvinellidae	*Paralvinella palmiformis*	AF168747
Ampharetidae	*Amphitritides gracilis*	AF508115
Amphinomida	*Eurythoe complanata*	AY040685
Aphroditoidea	*Harmothoe impar*	U50968
Apistobranchidae	*Apistobranchus typicus*	AF448150
Arenicolidae	*Arenicola marina (GenBank)*	AJ310502
	Arenicola marina (Arcachon, France)	AF508116
	Branchiomaldane vincentii	AF508117
Capitellidae	*Capitella capitata*	AF508118
	Dasybranchus caducus	AF448153
	Notomastus latericeus (GenBank)	AY040697
	Notomastus latericeus (Sylt, Germany)	AF508121
Chaetopteridae	*Chaetopterus variopedatus*	U67324
	Telepsavus spec.	AF448165
Clitellata	*Branchiobdella parasita*	AF310690
	Enchytraeus spec.	Z83750
	Glossiphonia complanata	Z83751
	Hirudo medicinalis	Z83752
	Lumbriculus variegatus	AF209457
	Lumbricus terrestris	AJ272183
	Tubificoides bermudae	AF209467
Cirratulidae	*Caulleriella parva*	AF448151
	Cirratulidae (GenBank)	AY040682
	Dodecaceria atra	AF448154
	Dodecaceria concharum	U50967
Ctenodrilidae	*Ctenodrilidae sp. Elba*	AJ310503
	Ctenodrilus serratus	AF508119
Dinophilidae	*Dinophilus gyrociliatus*	AF119074
Eunicidae	*Eunice pennata*	AY040684
Glyceridae	*Glycera americana*	U19519

Hrabeiella	*Hrabeiella periglandulata*	AJ310501
Maldanidae	*Clymenura clypeata*	AF448152
	Maldanidae (GenBank)	AY040694
Myzostomida	*Myzostoma cirriferum*	AF260585
	Myzostoma fissum	AF260584
	Myzostoma spec.	AF123305
Nephtyidae	*Nephtys hombergii*	U50970
Nereidae	*Nereis limbata*	U36270
Opheliidae	*Ophelia bicornis*	AF508122
	Ophelia neglecta	AF448156
	Ophelia rathkei	AF448157
	Ophelina acuminata	AY083310 + AY083311
	Polyophthalmus pictus	AF448161
	Travisia forbesii	AF508127
Orbiniidae	*Naineris laevigata*	AY040696
	Orbinia bioreti	AF448158
	Orbinia latreilii	AF448159
	Proscoloplos cygnochaetus	AF448162
	Protoaricia oerstedii	AF508123
	Protoariciella uncinata	AF508124
	Scoloplos armiger	U50972
	Scoloplos (Leodamas) johnstonei	AF508126
Oweniidae	*Owenia fusiformis*	AF448160
Paergodrilidae	*Paergodrilus heideri*	AJ310504
	Stygocapitella subterranea	AJ310505
Phyllodocidae	*Eteone longa*	AF448155
Protodrilida	*Protodrilus purpureus*	AJ310506
	Protodriloides symbisioticus	AF508125
Questa	*Questa paucibranchata*	AF209464
Sabellidae	*Sabella pavonina*	U67144
Scalibregmatidae	*Lipobranchus jeffreysii*	AF508120
	Scalibregma inflatum	AF448163
Serpulidae	*Protula spec.*	U67142
Siboglinidae	*Scerolinum brattstromi*	AF315061
	Siboglinum fiordicum	AF315060
	Riftia pachyptila	AF168745
Spionidae	*Aonides oxycephala*	AF448149
	Polydora ciliata	U50971
	Pygospio elegans	U67143
	Scolelepis squamata	AF448164
Syllidae	*Proceraea cornuta*	AF212179
Terebellidae	*Lanice conchilega*	X79873
	Loimia medusa	AY040694
Echiura	*Ochetostoma erythrogrammon*	X79875
	Urechis caupo	AF119076
Sipunculida	*Aspidosiphon misakiensis*	AF119090
	Phascolopsis gouldii	AF
	Phascolosoma granulatum	X79874
	Sipunculus nudus	AF448166
	Themiste alutacea	AF119075
Brachiopoda	*Lingula anatina*	X81631
Bryozoa	*Plumatella repens*	U12649
Phoronida	*Phoronis australis*	AF119079
Mollusca	*Lepidochitona corrugata*	X91975
	Ostrea edulis	L49052
Kamptozoa	*Barentsia hildegardae*	AJ001734

Arthropoda	*Lepisma saccarina*	X89484
	Limulus polyphemus	U91490
	Nebalia spec.	L81945
Kinorhyncha	*Pycnophyes kielensis*	U67997
Nematomorpha	*Gordius aquaticus*	X80233
Priapulida	*Priapulus caudatus*	AF025927

DNA extraction, PCR amplification, purification and sequencing

Collected specimen were identified and then preserved in 100% ethanol for later extraction. Genomic DNA was extracted from specimen using Qiagen DNeasy™ Tissue Kit according to the manufacture's instructions. PCR amplification of the 18S rDNA gene was performed in two overlapping fragments of ~900bp and ~1400bp each with modificated primer pairs from Giribet *et al*. (1996) by using standard cycle sequencing protocols. Amplification reaction mixtures for 18S rDNA contained 25 ìl Qiagen Taq PCR Master Mix, 2 ìl template-DNA, 4 ìl of each primer and 15 ìl H_2O.
Amplifications were carried out using an Eppendorf Mastercycler gradient. The following PCR temperature file was used: 95 C for 3 min; 35 cycles with 94 C for 35 seconds, 45-55 C for 45 seconds to 1 min, and 72 C for 1 min; final extension at 72 C for 10 min. After detection by gel electrophoresis the products were purified with the Qiaquick PCR Purification Kit (Qiagen). Sequencing of all amplified fragments in both directions was carried out by the *IIT* Biotech/Bioservice of the University of Bielefeld. Overlapping fragments of the 18S rDNA were combined by using BioEdit (Hall, 1999). Disagreement among these fragments was corrected by reference to the original chromatograms. All sequences were submitted to Genbank (for accession numbers see table 1).

Sequence Alignment

Sequences were aligned with CLUSTAL W (Thompson *et al*., 1994) and subsequently manually edited by eye using BioEdit (Hall, 1999). Gap positions and regions that could not be aligned unambiguously were excluded from the analysis.

Data Analysis

A chi-square test of homogeneity of base frequencies across taxa was performed. The program TreeView (Page, 1996) was used for tree visualization. All trees were rooted *a posteriori* the analysis using the sequence of *Gordius aquaticus* (Nematomorpha).

Maximum Likelihood

For estimating the appropriate model of sequence evolution, a hierarchical likelihood ratio test (hLRT) was carried out as implemented in the program modeltest version 3.06 (Posada and Crandall, 1998, 2001).The test criteria indicate that the Tamura Nei substitution model (Tamura and Nei, 1993) with equal base frequencies, invariant sites and gamma distribution (TrNef+I+Ã) represents the optimal model in respect to the data. A maximum likelihood analysis was performed with PAUP*, version 4,0b8 (Swofford, 2001) under the likelihood settings suggested by the result of the modeltest using the heuristic search option with TBR branch swapping and simple sequence addition.

Bayesian inference

A Bayesian analysis of the data set was conducted by using MrBayes 2.01 (Huelsenbeck and Ronquist, 2001). The ML parameters in MrBayes were set as follows: “lset nst=6” (GTR model), “rates=invgamma” and “basefreq=estimate”. Each Markov chain, three heated and one cold, was started from a random tree and all four chains ran simultaneously for $2.5 * 10^6$ generations, with trees being sampled every 100 generations for a total of 25,000 trees. After the likelihood of the trees of each chain converged, we discarded the first 5,000 trees as *burnin*. The majority-rule consensus tree containing the posterior probabilities of the phylogeny was determined from 20,000 trees.

Maximum parsimony

The data set was analyzed by parsimony, using PAUP*, version 4,0b8 (Swofford, 2001). Maximum parsimony searches were run with 1,000 random addition replicates, heuristic search option with tree-bisection-reconnection (TBR) branch swapping, holding one tree per step, and keeping all most-parsimonious trees. Bootstrap as well as Jackknife analyses were performed (Felsenstein, 1985; Farris, 1997) and the bootstrap and jackknife values were determined from 1,000 replicates subject to full heuristic searches with simple taxon addition to provide measures of relative clade support.

Results

Sequence data

After the exclusion of ambiguous sites, the alignment contains 1,574 positions, of which 709 are parsimony informative. The chi-square test of homogeneity of base frequencies across taxa resulted in no significant *P*-values (chi-square=144.546402, df=279, *P*=1.0). Assuming that compositional bias has no effect on the recovery of phylogenetic signal seems justifiable.

Maximum Likelihood and Bayesian inference

Of all the models that were considered in the modeltest, The Tamura Nei substitution model (Tamura and Nei, 1993) with equal base frequencies, invariant sites and gamma distribution (TrNef+I+Ã) represents the best fitting model for an explanation of the data. The most likely tree has a log-likelihood value of and is illustrated in Fig. 1. Fig.2 shows the results of the Bayesian inference which are presented as posteriori probabilities on the branches of the majority-rule consensus tree of the 20,000 inferred trees.

Whereas no evidence has been found for the monophyly of Annelida and Polychaeta, many of the traditional polychaete-"families" are recovered by ML and are well supported through Bayesian probabilities. So are Siboglinidae (Bayesian probability [BP]=100%), Chaetopteridae (BP=92%), Spionidae (BP=89%), Capitellidae (BP=100%), Parergodrilidae (BP=99), and Maldanidae (BP=100) also well supported. The inclusion of *Branchiomaldane vincenti* in the Arenicolidae is also well supported (BP=100). A monophyletic Terebellida including Terebellidae (represented by *Amphtitrides gracilis, Lanice conchilega* and *Loimia medusa*) and Alvinellidae (*Paralvinella palmiformis*) yields high support (BP=100), whereas the Terebellidae appear paraphyletic in regard to the Alvinellidae. *Travisia forbesii* joins the Scalibregmatidae *Lipobranchus jefreysii* and *Scalibregma inflatum* (BP=100), and a close relationship between *Scalibregma* and *Travisia* is also strongly supported (BP=100). The remaining Opheliids cluster together in a strong supported clade (BP=100). Monophyly of the Opheliniae as represented by *Polyophthalmus pictus* and

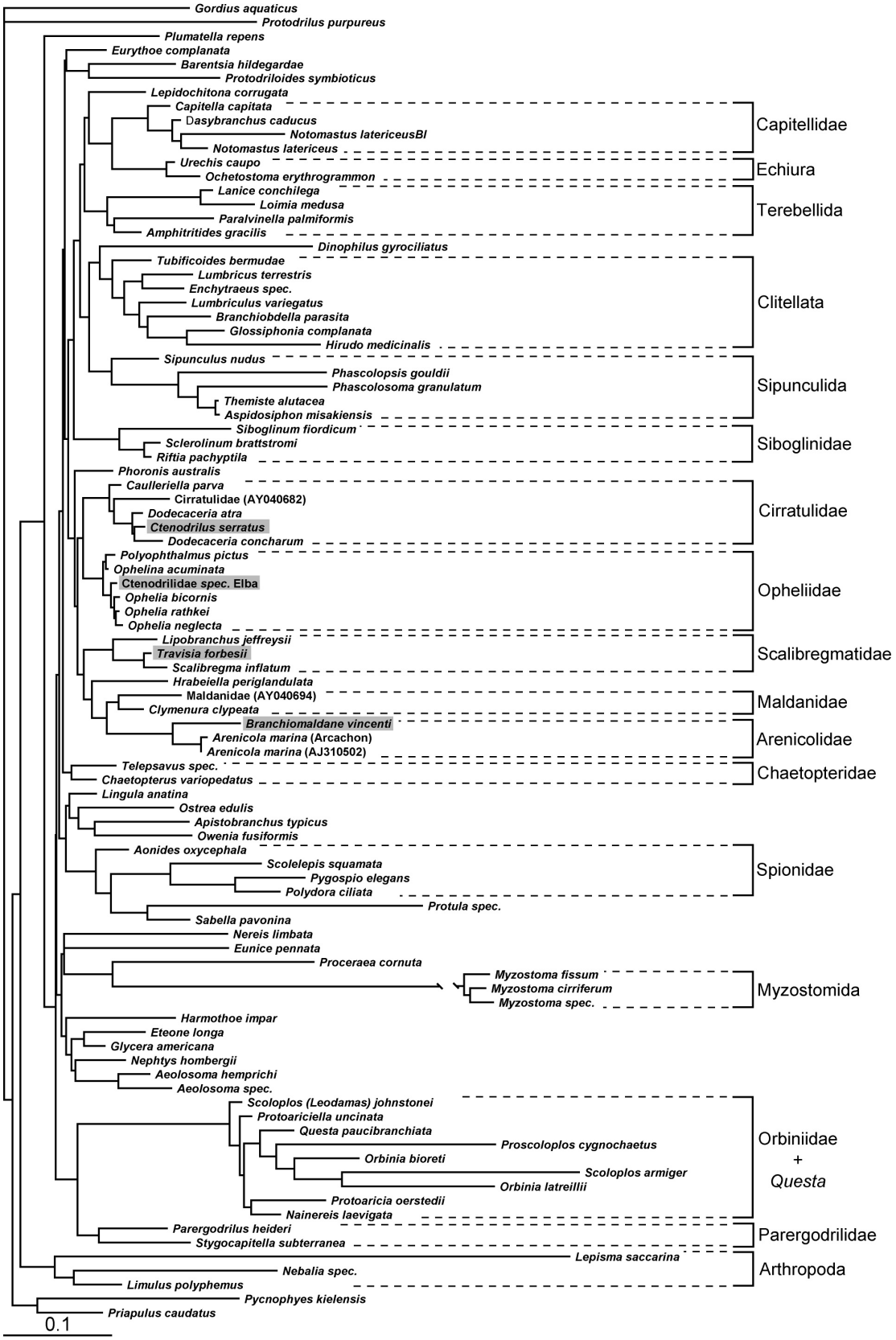

Figure1: Maximum likelihood tree based on the TrNef+I+Ã model of sequence evolution (-log*L*=244,060.72728). Taxa which are discussed in detail in the discussion are grey shaded.

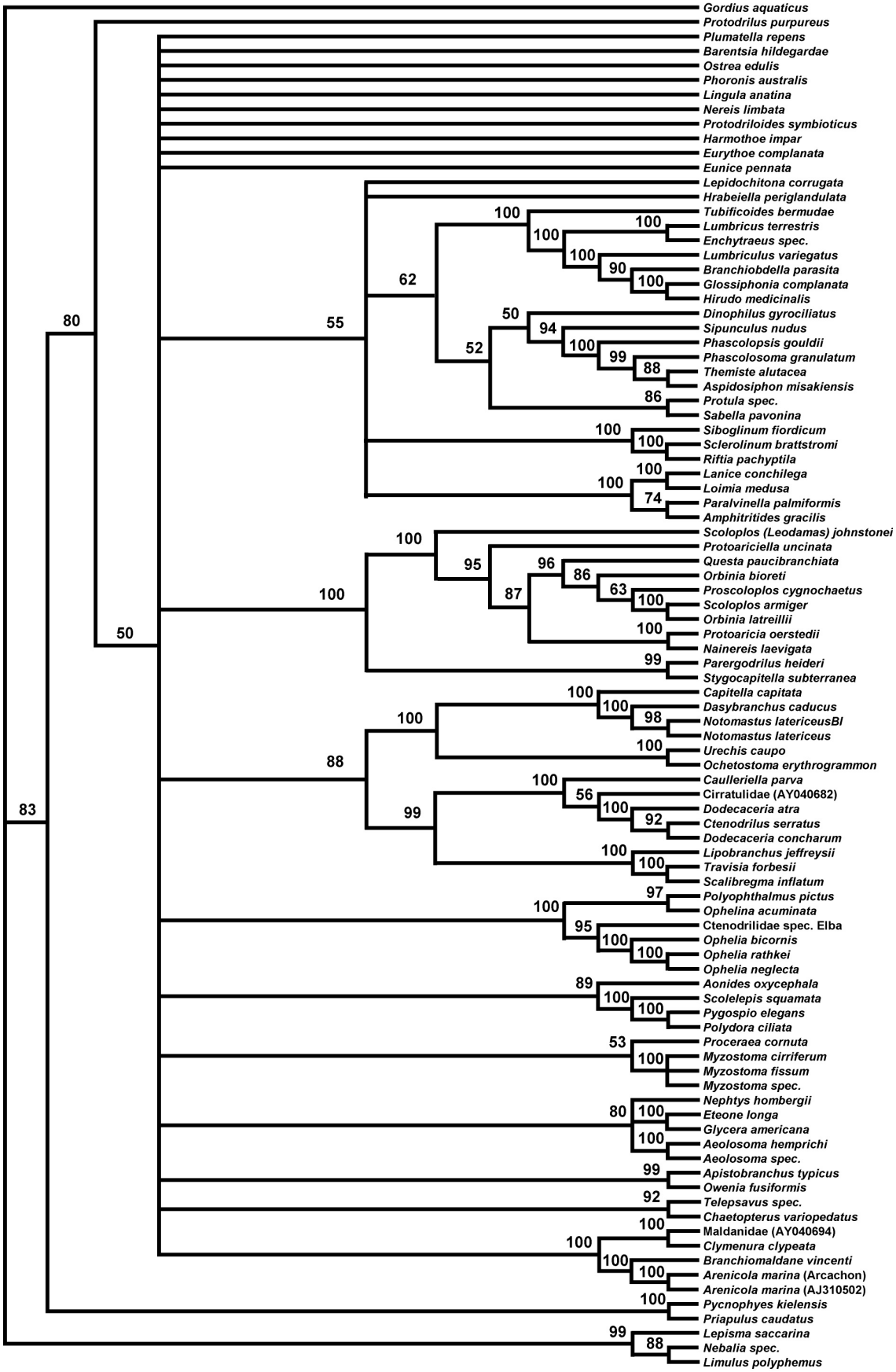

Figure 2: Majority-rule consensus tree of the Bayesian analysis. The posteriori probabilities are given above the branches.

Ophelina acuminata is also supported (BP=97). The Opheliidae are joined by the undescribed Ctenodrilid species from Elba which clusters togehter with the three considered *Ophelia* species (BP=95). In contrast to this, the other here regarded Ctenodrilid *Ctenodrilus serratus* joins the Cirratulidae (BP=100) and appears as ingroup of *Dodecaceria* (BP=92). A clade consisting of the orbiniids and *Questa* is well supported (BP=100), whereas the Orbiniidae appear paraphyletic in regard to *Questa* (BP=96). While the ingroup relationships of this taxon are only poorly resolved in reference to the branch-lenghts of the ML-analysis, they find good support in the Bayesian inference. The genera *Scoloplos* and *Orbinia* appear paraphyletic and a close relationship between *Naineris laevigata* and *Protoaricia oerstedii* yields high support (BP=100). A sistergroup relationship between the Orbiniidae + *Questa* cluster and the Parergodrilidae receives strong support (BP=100). Further relationships between polychaete families are recovered by ML as follows: A sistergroup relationship between Cirratulidae and Scalibregmatidae (BP=99); the echiurids cluster together with the Capitellidae (BP=100). A cluster consisting of *Apistobranchus typicus* (Apistobranchidae) and *Owenia fusiformis* (Oweniidae) receives high support (BP=99).
The monophyly of the Myzostomida is strongly supported and a close relationship to the syllid *Procerea cornuta* is recovered by ML as well as Bayesian inference (BP=53).
Monophyly of the Clitellata (BP=100) and Sipunculida is well supported (BP=94).
This analysis contradicts the hypothesis that Scolecida, Palpata and Canalipalpata is monophyletic.

Maximum parsimony

The equally weighted parsimony analysis results in 12 MP trees (length=6,598; consistency index [CI]=0.2766; consistency index excluding uninformative characters [CI']=0.2356; retention index [RI]=0.494). The bootstrap consensus tree from this analysis together with the bootstrap and jackknife frequencies is illustrated in Fig. 3.

Most of the strongly supported taxa of the Bayesian analysis are also supported through the MP Bootstrap- and Jackknifing-analyses, though often with a lower degree: Capitellidae (Bootstrap [BT]=98%; Jackknife [JK]=99%), Chaetopteridae (BT=74%; JK=72%), Maldanidae (BT=96%; JK=99%), Arenicolidae (including *Branchiomaldane vincenti*) (BT=100%; JK=100%), Parergodrilidae (BT=75%; JK=74%), Siboglinidae

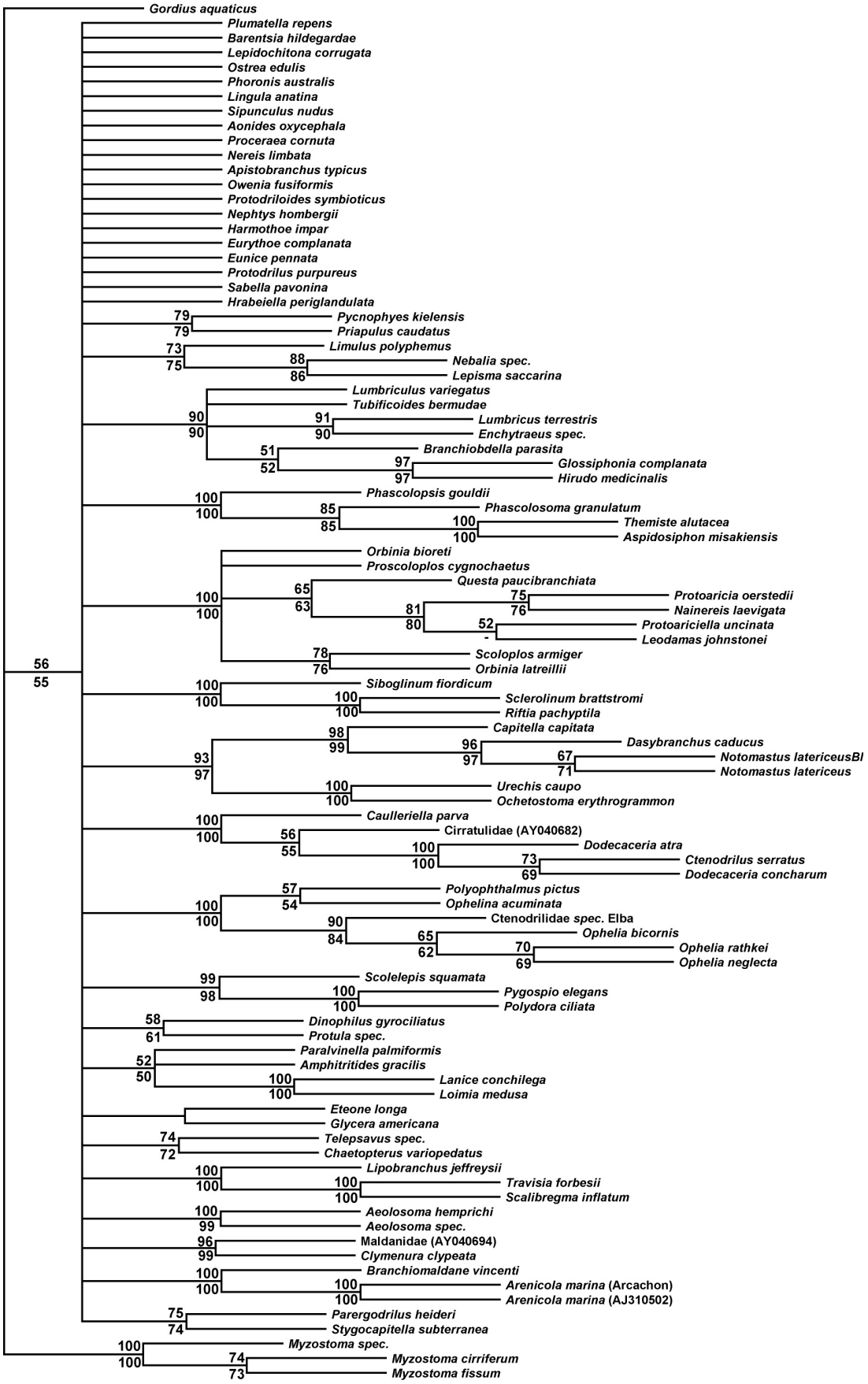

Figure 3: Bootstrap consensus tree of the Maximum Parsimony analysis. Bootstrap- and jackknife-frequencies are given above and below the branches.

(BT=100%; JK=100%). As well as in ML, *Ctenodrilus serratus* joins the cirratulids (BT=100%; JK=100%) and appears as ingroup of *Dodecaceria* (BT=100%; JK=100%). This analysis also recovers the inclusion of the undescribed Ctenodrilid from Elba in the Opheliids (BT=100%; JK=100%) and supports a cluster of this taxon together with *Ophelia* (BT=90%; JK=84%). *Travisia forbesii* as an ingroup of Scalibregmatidae yields high support (BT=100%; JK=100%) and clusters together with *Scalibregma inflatum* (BT=100%; JK=100%). The result of a capitellid - echiurid relationship is also confirmed in this analysis (BT=93%; JK=97%). A clade consisting of *Questa* and Orbiniidae yields strong support (BT=100%; JK=100%), whereas ingroup relationships of this taxon are only weakly supported. As in ML the Orbiniids appear paraphyletic in regard to *Questa*. In contrast to the Bayesian inference the Terebellida are only poorly supported (BT=52%; JK=50%) and no support is given for the Spionidae, Arenicolidae + Maldanidae, Orbiniidae + *Questa* + Parergodrilidae and Sipunculida. Monophyly of Myzostomida is well supported (BT=100%; JK=100%). In the parsimony analysis they group at a basal position within the tree. This basal position contradicts the results of the ML analysis, which can be explained by a long-branch effect.

Discussion

Bayesian probabilities vs. Bootstrap frequencies

A Bayesian analysis of phylogenies searches for the best set of trees that is consistent with a given model and the data set under investigation (Rannala and Yang, 1996; Mau *et al*., 1999). The consensus of this set of trees can be used to estimate probabilties for node support, which can be taken as the equivalent of bootstrap values (Hall, 2001). Whereas in the present analysis all nodes with a high bootstrap support (>80%) are also well supported by Bayesian probalities. This does not holds true vice versa: not all well supported nodes through Bayesian probabilities are suported by bootstrap. Bayesian probabilities tend to be higher than comparable bootstrap values for the same node. This has been noticed by some authors before (Bleidorn *et al*., 2002; Buckley *et al*., 2002; Whittingham *et al*., 2002), but there is presently no generally accepted procedure of interpreting these findings, since the nature of posteriori probabilities is not well enough understood yet (Leaché and Reeder, 2002). This flaw have to be kept in mind while

interpreting results which are “only” supported through high Bayesian probabilities, but which lack from sufficient bootstrap support.

Annelids

As many analyses of molecular data sets in the past (Kojima, 1998; Brown *et al*., 1999; McHugh, 2000; Rota *et al*., 2001; Struck *et al*., 2002; Bleidorn *et al*., 2002), this analysis gives no support for a monophyletic Polychaeta nor a monophyletic Annelida. The low resolution may be due to a rapid radiation of the Annelida that has intensely been discussed elsewhere (Brown *et al*., 1999; Rota *et al*., 2001) and is generally credited to an erosion of information during time (Abouheif *et al*., 1998).

The focus of this analysis was to study ingroup relationships of the Polychaeta in comparison to the hypotheses proposed by the cladistic analysis of Rouse and Fauchald (1997) on the basis of morphological data, which has been adopted in the comprehensive polychaete monographs of Beesley *et al*. (2000) and Rouse and Pleijel (2001).

Position of the Echiura

Most remarkable is the confirmation of our former result (Bleidorn *et al*.,2002) that Echiura are a polychaete ingroup taxon with a close affinity to the Capitellidae. This relationship is well supported through bootstrap values and Bayesian probabilities. The phylogenetic position of these unsegmented marine worms is controversially discussed. As Nielsen (2000) pointed out, the Echiura resemble annelids in anatomy and embryology, with the exception that they show no trace of segmentation. In the analyses of Rouse and Fauchald (1995, 1997) and Rouse (1999) they are treated as sister taxon of the Articulata (Annelida + Arthropoda). Molecular analyses instead place them as derived polychaetes (McHugh, 1997, 1999; Brown *et al*., 1999; Bleidorn *et al*., 2002). This view is congruent with the findings of Hessling and Westheide (2002) that Echiura show serially repeated units in their nervous system which correspond to typical metameric ganglia of the Annelida. The placement as derived polychaetes favor the hypotheses of a secondary lost of segmentation in Echiura. Up to now there are no morphological synapomorphies of a possible common ancestry of Capitellidae and Echiura found.

"Scolecida" and "Orbiniidae"

Contradicting Rouse and Fauchald (1997), no evidence has been found for the monophyly of "Scolecida". Our analyses include a broad sampling of scolecid taxa including Capitellidae (4 sequences), Arenicolidae (3), Maldanidae (2), Scalibregmatidae (3), Opheliidae (6), *Questa* (1), and Orbiniidae (8). None of our analyses yield support for a close relationship of these taxa. Instead, we find the Capitellidae clustering together with Echiura. The inferred phylogenetic position of the Opheliidae depends on the choice of method. Whereas ML recovers a cirratulid - opheliid relationship, Bayesian inference gives no hint for a possible sistergroup of Opheliids and, instead, supports a scalibregmatid - cirratulid clade. In accordance to previous analyses (Kojima, 1998; Brown *et al*., 1999; McHugh, 2000; Rota *et al*., 2001; Struck *et al*., 2002; Bleidorn *et al*., 2002), most hypotheses on relationships between polychaete families are only weakly supported. Orbinidae + *Questa* are closely related to the Parergodrilidae, a result supporting the analysis of Struck *et al*. (2002). Interestingly this relationship was also found in some of the cladistic analyses by Rouse and Fauchald (1997). Irrespective of the method used, our analysis suggests a probable paraphyly of the Orbiniidae in regard to *Questa*. The ingroup relationships of the Orbiniidae inferred by the 18S rDNA are in contrast to the traditional assumption (see Hartman, 1957) and a recent cladistic analysis of Blake (2000c). Confirming Blake (1996, 2000c), there seems to be no justification for distinguishing the two subfamilies Orbiniinae (here represented by *Scoloplos* (*Leodamas*) *johnstonei, Scoloplos armiger, Naineris laevigata, Orbinia bioreti,* and *Orbinia latreilii*) and Protoariciinae (*Proscoloplos cygnochaetus, Protoaricia oerstedii* and *Protoariciella uncinata*), that were morphologically seperated on the presence of one or two achaetigerous rings between prostomium and first chaetiger (Eisig, 1914).

According to Rouse and Fauchald (1997) the results of our ML-analysis confirm the maldanid - arenicolid relationship, a result which lacks support in the MP Bootstrap- and Jackknifing analyses. The placement of *Branchiomaldane*, an arenicolid taxon with maldanid-like anatomy (Rouse and Pleijel, 2001), within Arenicolidae is confirmed by all analyses with strong support and recjects a possible inclusion of this taxon in the Maldanidae.

Monophyly of Opheliidae and Scalibregmatidae and the position of Travisia

Opheliids are worldwide distributed with more than 150 described species. However, there is no convincing known apomorphy for the Opheliidae (Fauchald and Rouse, 1997) and Rouse and Pleijel (2001) annotated that they might be paraphyletic with regard to Scalibregmatidae. Three distinct morphological groups can be recognized within the Opheliidae: A taxon Opheliinae comprising those genera with distinct body regions (*Euzonus*, *Lobochesis* and *Ophelia*), a taxon Ophelininae including all genera with an anal tube consisting of several reduced segments (*Ammotrypanella*, *Antiobactrum*, *Armandia*, *Ophelina*, *Polyophthalmus,* and *Tachytrypane)* and the taxon *Travisia*, which closely resembles scalibregmatid species (Bellan *et al*., 1990; Hartmann-Schröder, 1996) and is closely related to Scalibregmatidae as mentioned in Blake (2000a). Dauvin and Bellan (1994) studied the systematics of Travisiinae and synonymisized *Dindymenides* and *Kesun* with *Travisia*. They also found that ventral and lateral grooves are generally absent or only poorly developed if present. In contrast, a well developed ventral groove can be found in all other opheliid taxa. All *Travisia* species (except the fusiform species *Travisia hobsonae* and *Travisia fusiformis* (Santos, 1977)) are maggot-shaped and resemble scalibregmatids like *Polyphysia*, while all other Opheliids are fusiform or cylindrical shaped (Bellan *et al*., 1990). Storch (1988) pointed out that *Travisia* possesses a stratified epidermis, which is unusual for invertebrates.

Scalibregmatids are worldwide distributed polychaets with 55 nominal species. According to Fauchald and Rouse (1997) there is no autapomorphy for this taxon known and typical scalibregmatid characters as the rugose epidermis and segmental annuli can be found in the Opheliidae, too. In their polychaete “meta-tree” Rouse and Pleijel (2001) show Scalibregmatids as a sistergroup to a taxon consisting of arenicolids, capitellids, maldanids and opheliids. In traditional classifications scalibregmatids are grouped together with opheliids (Hartmann-Schröder, 1996). Within the Scalibregmatidae Kudenov and Blake (1978) and Blake (1981) distinguish three groups representing different body forms. A group with an arenicoliform body (e.g. *Scalibregma*), a group with a maggot-like body (e.g. *Polyphysia*) and *Scalibregmella* which has a slender and elongated body.

Summarizing these findings, uncertainties regarding the monophyly of opheliids and scalibregmatids are due to placement of *Travisia*, a taxon which closely resembles

scalibregmatid species (Bellan *et al*., 1990), as basal opheliid. Our analysis strongly supports a common ancestry of *Travisia* and scalibregmatids, whereas *Travisia* can be treated as ingroup taxon of Scalibregmatidae. This results supports the view of Blake (2000b), who in the same way hypothesizes a close relationship between these taxa. After transferring *Travisia* to the Scalibregmatidae it has to be proofed if the presence of lateral grooves and a strong ventral groove might support the monophyly of the remaining new combined Opheliidae. Monophyly of the remaining Opheliidae is also strongly supported in all of our analyses independent of the choice of method. A rugose epidermis can be assessed as an apomorphy for a taxon consisting of the Scalibregmatidae and *Travisia*.

Monophyly and the position of Ctenodrilidae

The first described ctenodrilid, *Ctenodrilus serratus*, has been originally included into the rhabdocoel Turbellaria by Schmidt (1857). Since then only a few specimen have been described in the genera *Ctenodrilus*, *Aphropharynx, Raphidrilus,* and *Raricirrus*. A new genus and species is mentioned in Rota *et al.* (2001), but is still waiting to be described. *Zeppelina* was synonymisized with the cirratulid taxon *Dodecaceria* by George and Petersen (1991). In exchange they include the cirratulid genus *Raricirrus* within the Ctenodrilidae. Fauvel (1927) and Day (1967) considered the ctenodrilids to be a part of the Cirratulidae, while Hartmann-Schröder (1971) retains them as a seperat drilomorph family. An examination of the nervous system of *Ctenodrilus serratus* by Gelder and Palmer (1976) reinforced the affinities with the Cirratulidae. In the cladistic analysis of Rouse and Fauchald (1997) they form a clade together with Fauveliopsidae, Poebiidae and Sternaspidae. In the maximum parsimony analysis of a 18S rDNA dataset by *Rota et al.* (2001) the new Ctenodrilid species forms a clade with *Arenicola* (Arenicolidae) and *Dodecaceria* (Cirratulidae), but lacks sufficient bootstrap support.

Surprisingly, the newly-discovered ctenodrilid species from Elba (Rota *et al*., 2001) branches off between the opheliids and a common ancestry of this taxa is strongly supported irrespective of the method used. This suggests, that in fact this new species proofes to be a juvenile or progenetic opheliid. Further biological data on this species are neccessary to confirm one of the two hypotheses. *Ctenodrilus serratus*, the other Ctenodrilid species included in the analysis, clusters together with the cirratulids and shows a close affinity to the two included sequences of *Dodecaceria*. Interestingly it has

been reported, that similarities between adult ctenodrilids and juvenile individuals of *Dodecaceria* often caused confusion concerning taxonomical problems (see Petersen and George, 1991). The systematic placement of the ctenodrilids as part of the Cirratulidae has a long tradition (Mesnil and Caullery, 1897; Fauvel, 1927; Day, 1967) and is herewith confirmed, contradicting Hartmann-Schröder (1996) and Rouse and Fauchald (1997) who treated them as a separate taxon outside the Cirratulidae. This result is validated through high support of Bayesian probabilities, Bootstrap- and Jackknifing-frequencies. Nevertheless, it still has to be tested wether all remaining ctenodrilid taxa have to be included into the Cirratulidae.

Position of Myzostomida

Myzostomida are marine worms associated with Echinoderms (Grygier, 2000). As host-specific symbionts (or parasites) they show a highly derived anatomy in their adult morphology (Eeckhaut *et al*., 2000). While many authors regard them as derived annelids (Nielsen, 2000) or polychaetes (Rouse and Fauchald, 1997), recent cladistic analyses of morphological and molecular data support the hypothesis that Myzostomida are not nested within annelids (Haszprunar, 1996; Eeckhaut *et al*., 2000; Zrzavy *et al*., 2001). Zrzavy *et al*. (2001) propose that Myzostomida are the sistergroup of the Cycliophora, in Eeckhaut *et al*. (2000) they are closely related with Plathelminthes, and Haszprunar (1996) favors a sistergroup relationship to a taxon consisting of sipunculids, clittelates and polychaetes. A close relationship to acanthocephalans is proposed by Mattei and Marchand (1987) on the basis of ultrastructural sperm cell similarities. As have not included most of the above mentioned possible myzostomid sistergroups we cannot reject any of the hypotheses. But the results of the ML-analysis (Fig.1) support the idea that Myzostomida are abberant polychaetes. This view is congruent with the results of Müller and Westheide (2000) on the nervous system of *Myzostoma cirriferum*, which exhibits several structures typical for the polychaete nervous system.

These results support our assumption that on the basis of a broader taxon sampling the phylogenetic position of controversly discussed taxa can be inferred by using 18S rDNA sequence data.

References

Abouheif, E., Zardoya, R., and Meyer, A. (1998). Limitations of metazoan 18S rRNA sequence data: Implications for reconstructing a phylogeny of the animal kingdom and inferring the reality of the Cambrian explosion. *J. Mol. Evol.* **47**: 394-405.

Bartolomaeus, T. (1998). Chaetogenesis in polychaetous Annelida – significance for annelid systematics and the position of the Pogonophora. *Zoology* **100**: 348-364.

Beesley, P.L., Ross, G.J.B., and Glasby, C.J. (2000). "Polychaetes & Allies: The Southern synthesis. Fauna of Australia. Vol. 4A Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula." CSIRO Publishing, Melbourne.

Bellan, G., Bellan-Santini, D., and Dauvin, J.-C. (1990). Phenetique et phylogenie des Opheliidae (Annelides Polychetes). *C. R. Acad. Sci. Paris, Series III,* **310**: 175-181.

Blake, J.A. (1981). The Scalibregmatidae (Annelida: Polychaeta) from South America and Antarctica collected chiefly during the cruises of the R/V Anton Bruun, R/V Hero and USNS Eltanin. *Proc. Biol. Soc. Wash.* **94**: 1131-1162.

Blake, J.A. (1996). Family Orbiniidae Hartman, 1942. *In* "Taxonomic Atlas of the Benthic Fauna of the Santa Maria Basin and Western Santa Barbara Channel, Vol. 6: The Annelida, Part 3: Polychaeta: Orbiniidae to Cossuridae (J.A. Blake, B. Hilbig, and P.H. Scott, Eds.), pp. 1-26, Santa Barbara Museum of Natural History, Santa Barbara, California.

Blake, J.A. (2000a). Family Opheliidae Malmgren, 1867. *In* "Taxonomic Atlas of the Benthic Fauna of the Santa Maria Basin and Western Santa Barbara Channel, Vol. 7: The Annelida, Part 4: Polychaeta: Flabelligeridae to Sternaspidae (J.A. Blake, B. Hilbig, and P.H. Scott, Eds.), pp. 145-168, Santa Barbara Museum of Natural History, Santa Barbara, California.

Blake, J.A. (2000b). Family Scalibregmatidae Malmgren, 1867. *In* "Taxonomic Atlas of the Benthic Fauna of the Santa Maria Basin and Western Santa Barbara Channel, Vol. 7: The Annelida, Part 4: Polychaeta: Flabelligeridae to Sternaspidae (J.A. Blake, B. Hilbig, and P.H. Scott, Eds.), pp. 129-144, Santa Barbara Museum of Natural History, Santa Barbara, California.

Blake, J.A. (2000c). A new genus and species of polychaete worm (Family Orbiniidae) from methane seeps in the Gulf of Mexico, with a review of the systematics and phylogenetic interrelationships of the genera of Orbiniidae. Cah. biol. Mar. 41: 435-449.

Bleidorn, C., Vogt, L., and Bartolomaeus, T. (2002). A contribution to polychaete phylogeny using 18S rDNA sequence data. *J. zool. Sys. Evol. Res.* ***handed in***.

Brown, S., Rouse, G.W., Hutchings, P., and Colgan, D., (1999). Assessing the usefulness of histone H3, U2 and snRNA and 28S rDNA in analyses of polychaete relationships. *Aust. J. Zool.* **47**: 499-516.

Buckley, T.R., Arensburger, P., Simon, C., and Chambers, G.K. (2002). Combined data, Bayesian inference, and the origin of the New Zealand Cicada genera. *Syst. Biol*. **51**, 4-18.

Dauvin, J.-C. and Bellan, G. (1994). Systematics, ecology and biogeographic relationships in the sub-family Travisinae (Polychaeta, Opheliidae). *Mem. Mus. natn. Hist. nat.* **162**: 169-184.

Day, J.H. (1967). "A Monograph on the Polychaeta of Southern Africa. Part 2. Sedentaria," British Museum (Natural History), London.

Eeckhaut, I., McHugh, D., Mardulyn, P., Tiedemann, R., Monteyne, D., Jangoux, M., and Milinkovitch, M.C. (2000). Myzostomida: a link between trochozoans and flatworms? *Proc. R. Soc. Lond. B* **267**: 1383-1392.

Eisig, J. (1914). Zur Systematik, Anatomie und Morphologie der Ariciiden nebst Beiträgen zur generellen Systematik. *Mitt. zool. Stn. Neapel* **21**: 153-600.

Farris, J.S. (1997). The future of phylogeny reconstruction. *Zool. Scr.* **26**: 303-311.

Fauchald, K. (1977). The polychaete worms. Definitions and keys to the orders, families and genera. *Nat. Hist. Mus. LA County Sci. Ser*. **28**: 1-188.

Fauchald, K., and Rouse, G.W., 1997. Polychaete systematics: past and present. *Zool. Scr.* **26**: 71-138.

Fauvel, P. (1927). Polychetes Sedentaires. Addenda aux Errantes, Archiannelides, Myzostomaires. *Faune de France* **5**: 1-488.

Felsenstein, J. (1985). Confidence limits on phylogenys: an approach using the bootstrap. Evolution **39**: 783-791.

Gelder, S.R., and Palmer, R. (1976). The nervous system of the marine polychaete *Ctenodrilus serratus* and its importance in the taxonomic position of the Ctenodrilidae. Trans. Amer. Micr. Soc. **95**: 156-164.

George, J.D., and Petersen, M.E. (1991). The validity of the genus *Zeppelina* (Vaillant) (Polychaeta: Ctenodrilidae). Ophelia Suppl. *5*: 89-100.

Giribet,G., Carranza,S., Baguna,J., Riutort,M., and Ribera,C. (1996). First molecular evidence for the existence of a Tardigrada + Arthropoda clade. *Mol. Biol Evol.* **13**: 76-84.

Grygier, M.J. (2000). Myzostomida. *In* "Polychaetes & Allies: The Southern synthesis. Fauna of Australia. Vol. 4A Polychaeta, Myzostomida, Pogonophora, Echiura, Sipuncula" (P.L. Beesley, G.J.B. Ross, and C.J. Glasby, Eds.), pp. 297-329, CSIRO Publishing, Melbourne.

Halanych, K.M., Feldman, R.A., and Vrijenhoek, R.C. (2001). Molecular evidence that *Sclerolinum brattstromi* is closely related to Vestimentiferans, not to frenulate Pogonophorans (Siboglinidae, Annelida). *Biol. Bull.* **201**: 65-75.

Hall, B.G. (2001). “Phylogenetic trees made easy: a how-to manual for molecular biologists,” Sinauer Associates, Inc, Sunderland, Massachusetts.

Hall, T.A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids. Symp. Ser.* **41**: 95-98.

Hartman, O. (1957). Orbiniidae, Apistobranchidae, Paraonidae and Longosomidae. *Allan Hancock Pacific Expeditions* **15**: 211-393.

Hartmann-Schröder, G. (1971). Annelida, Borstenwürmer, Polychaeta. *Tierwelt Deutschlands* **58**: 1-594.

Hartmann-Schröder, G. (1996). Annelida, Borstenwürmer, Polychaeta. 2., neubearbeitete Auflage. *Tierwelt Deutschlands* **58**: 1-648.

Haszprunar, G. (1996). The Mollusca: coelomate turbellarians or mesenchymate annelids? In "Origin and evolutionary radiation of the Mollusca" (J.D. Taylor, Eds.), pp. 1-28, Oxford Universtiy Press, Oxford.

Hausen, H. (2001). „Untersuchungen zur Phylogenie “spiomorpher” Polychaeten (Annelida),“ Logos Verlag, Berlin.

Hausen, H., and Bartolomaeus, T. (1998). Setal structure and chaetogenesis in Scolelepis squamata and Malacoceros fuliginosus (Spionida, Annelida). *Acta Zool.* **79**: 146-161.

Hessling, R., and Westheide, W. (2002). Are Echiura derived from a segmented ancestor? - Immunohistochemical analysis of the nervous system in developmental stages of *Bonellia viridis*. *J. Morphol.* **252**: 100-113.

Huelsenbeck, J.P., and Ronquist, F.R. (2001). MrBayes: Bayesian inference of phylogeny. *Biometrics* **17**: 754-755.

Kojima, S. (1998). Paraphyletic status of Polychaeta suggested by phylogenetic analysis based on the amino acid sequences of elongation factor 1-alpha. *Mol. Phyl. Evol.* **9**: 255-261.

Kudenov, J., and Blake, J.A. (1978). A review of the genera and species of the Scalibregmidae (Polychaeta) with descriptions of one new genus and three new species from Australia. *J. Nat. Hist.* **12**: 427-444.

Leaché, A.D., and Reeder, T.W. (2002). Molecular systematics of the eastern Fence Lizard (*Sceloporus undulatus*): a comparison of Parsimony, Likelihood, and Bayesian approaches. *Syst. Biol*. **51**, 44-68.

Mattei, X., and Marchand, B. (1987). Les spermatozoides des acanthocephales et des myzostomides. Ressemblances et consequences phyletiques. *C.R. Acad. Sci., Paris* **305**: 525-529.

Mau, B., Newton, M., and Larget, B. (1999). Bayesian phylogenetic inference via Markov chain Monte Carlo methods. *Biometrics* **55**: 1-12.

McHugh, D. (1997). Molecular evidence that echiurans and pogonophorans are derived annelids. *Proc. Natl. Acad. Sci. USA* **94**: 8006-8009.

McHugh, D. (1998). Phylogeny of the Annelida:Siddall et al. (1998) rebutted. *Cladistics* **15**: 85-89.

McHugh, D. (2000). Molecular phylogeny of the annelids. *Can. J. Zool.* **78**: 1873-1884.

Mesnil, F., and Caullery, M. (1897). Sur la position systematique du genre *Ctenodrilus* Clap.; ses affinities avec les Cirratuliens. *C.R. hebd. Seanc. Acad. Sci., Paris* **125**: 542-544.

Meyer, K., and Bartolomaeus, T. (1996). Ultrastructure and formation of the hooked setae in *Owenia fusiformis* delle Chiaje, 1842: Implications for annelid phylogeny. *Can. J. Zool.* **74**: 2143-2153.

Müller, M.C., and Westheide, W. (2000). Structure of the nervous system of *Myzostoma cirriferum* (Annelida) as revealed by immunohistochemistry and cLSM analyses. *J. Morphol.* **245**: 87-98.

Nielsen, C. (2000). "Animal Evolution," Oxford University Press, Oxford.

Page, R.D.M. (1996). TREEVIEW: An application to display phylogenetic trees on personal computers. *Comp. Appl. Biosci.* **12**: 357-358.

Petersen, M.E., and George, J.D. (1991). A new species of Raricirrus from Northern Europe, with notes on its biology and a discussion of the affinities of the genus (Polychaeta: Ctenodrilidae). *Ophelia Suppl.* **5**: 185-208.

Posada, D., and Crandall, K.A. (1998). Modeltest: testing the model of DNA substitution. *Bioinformatics* **14**: 917-918.

Posada, D., and Crandall, K.A. (2001): Selecting the best-fit model of nucleotid substitution. *Syst. Biol.* **50**: 580-601.

Purschke, G. (1997). Ultrastructure of the nuchal organs in polychaetes (Annelida) – new results and review. *Acta Zool.* **78**: 123-143.

Purschke, G., and Tzetlin, A.B. (1996). Dorsal ciliary folds in the polychaet foregut: structure, prevalence and phylogenetic significance. *Acta Zool.* **77**: 33-49.

Rannala, B., and Yang, Z.H. (1996). Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *J. Mol. Evol.* **43**: 304-311.

Rota, E., Martin, P., and Erseus, C. (2001). Soil-dwelling polychaetes: enigmatic as ever? Some hints on their phylogenetic relationships as suggested by a maximum parsimony analysis of 18S rRNA gene sequences. *Contributions to Zoology* **70**: 127-138.

Rouse, G.W. (1999). Trochophora concepts: ciliary bands and the evolution of larvae in spiralian Metazoa. *Biol. J. Lin. Soc.* **66**: 411-464.

Rouse, G.W., and Fauchald, K. (1995): The articulation of annelids. *Zool. Scr.* **24**: 269-301.

Rouse, G.W., and Fauchald, K. (1997): Cladistics and polychaetes. *Zool. Scr.* **26**: 139-204.

Rouse, G.W., and Fauchald, K. (1998). Recent views on the status, delineation and classification of the Annelida. *Amer. Zool.* **38**: 953-964.

Rouse, G.W., and Pleijel, F. (2001). Polychaetes. University Press, Oxford.

Santos, S.L. (1977). A new species of *Travisia* (Polychaeta, Opheliidae) from Tampa Bay, Florida. *Proc. Biol. Soc. Wash.* **89**: 559-564.

Schmidt, O. (1857). Zur Kenntnis der Turbellaria, Rhabdocoela und einiger anderer Würmer des Mittelmeeres. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften, Wien, Mathematisch-Naturwissenschaftliche Klasse* **23**: 347-366.

Storch, V. (1988). Integument. *In* "The Ultrastructure of Polychaeta. Microfauna Marina Vol. 4." (W. Westheide, and C.O. Hermans, Eds.), pp. 13-36, Gustav Fischer Verlag, Stuttgart.

Struck, T., Hessling, R., and Purschke, G. (2002). The phylogenetic position of the Aelosomatidae and Parergodrilidae, two enigmatic oligochaete-like taxa of the "Polychaeta", based on molecular data from 18S rDNA sequences. *J. zool. Sys. Evol. Res.* ***in press***.

Swofford, D.L. (2001). PAUP*. Phylogenetic analysis using parsimony, version 4.0b8. Sinauer, Sunderland, MA.

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512-526.

Thompson, J.D., Higgins, D.G, and Gibson, T.J. (1994): CLUSTAL W: improving the sensivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673-4680.

Westheide, W. (1997). The direction of evolution within the Polychaeta. *J. Nat. Hist.* **31**: 1-15.

Westheide, W., McHugh, D., Purschke, G., and Rouse, G. W. (1999). Systematization of the Annelida: different approaches. *Hydrobiologia* **402**: 291-307.

Whittingham, L.A., Slikas, B., Winkler, D.W., and Sheldon, F.H. (2002). Phylogeny of the tree swallow genus, Tachycineta (Aves: Hirudinidae), by Bayesian analysis of mitochondrial DNA sequences. *Mol. Phyl. Evol*. **22**: 430-441.

Zrzavy, J., Hypsa, V., and Tietz, D.F. (2001). Myzostomida are not annelids: Molecular and morphological support for a clade of animals with anterior sperm flagella. *Cladistics* **17**: 170-198.

Testing and Weighting Characters

L. Vogt

Zoomorphologie und Systematik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, E-Mail: lars.vogt@biologie.uni-bielefeld.de

eingereicht und akzeptiert in:

Organisms, Diversity & Evolution

Abstract

The justification of weighting characters in parsimony analyses varies tremendously. Some authors argue for weighting *a posteriori*, some for *a priori* and especially those authors that rely on a falsificationist approach to systematics argue for non-weighting. To find a decision, while following the falsificationist approach, one first has to investigate the necessary conditions for the possibility of phylogenetic research to establish an empirical science *sensu* Popper. A conception of *phylogenetic homology* together with the *criterion of identity* is proposed, which refers to the genealogical relations between individual organisms. From this conception a differentiation of the terms *character* and *character state* is proposed, constituting each character as a single epistemological argument for the reconstruction of a unique transformation event. Synapomorphy is distinguished from homology by referring to the relationship between species instead of individual organisms, thus constituting the logical linkage of the set of all synapomorphies as a subset of the set of all homologies. By examining the structure of characteristics during character analysis and hypothesizing specific types of transformations responsible of having caused them, a specific degree of severity is assigned to each identity test. It thus provides a specific degree of corroboration for every hypothesis that successfully passed this test. Since the *congruence criterion* together with the methodological principle of parsimony test hypotheses of synapomorphy against each other on grounds of their degree of corroboration gained from the identity test, these different degrees of corroboration determine the specific weights given to characters and character state transformations before the cladistic analysis. This provides a reasonable justification for an *a priori* weighting scheme within a falsificationist approach to phylogeny. It also demonstrates the indispensable necessity of its application.

[*a priori* weighting, congruence test, falsification, homology, identity test, synapomorphy]

Introduction

The weighting of phylogenetic characters within cladistic analyses, equally or differentially, is common and compelling procedure in systematics. The theoretical justifications for this procedure vary tremendously (Farris, 1969; Neff, 1986; Carpenter, 1988; Bryant, 1989; Goloboff, 1993; Chippindale and Wiens, 1994; Allard and Carpenter, 1996; Milinkovitch *et al.*, 1996; Kluge, 1997a; Björklund, 1999; Källersjö *et al.*, 1999; Wenzel, 1999; Broughton *et al.*, 2000; Lutzoni *et al.*, 2000). I will try to examine the conditions for the possibility of justifying a weighting scheme within a falsificationist approach to phylogeny.

The effect of applying Popper's falsificationist methodology on phylogenetic inferences has been and still is extensively discussed (e. g., Farris, 1970, 1979, 1983, 1995, 2000; Bock, 1973; Wiley, 1975; Kitts, 1977; Platnick and Gaffney, 1977, 1978a, 1978b; Cracraft, 1978; Platnick, 1979; Sober, 1993; Kluge and Wolf, 1993; Kluge, 1997, 1997a, 1998; Siddall and Kluge, 1997; Rieppel, 1999; Brower 2000; de Queiroz and Poe, 2001; Kluge, 2001). One central statement, however, remains vague. If one wants to demonstrate that phylogenetic research is an empirical science *sensu* Popper, one has to show that phylogenetic hypotheses are testable, which means that they have to be open to refutation by empirical evidence. For testing phylogenetic hypotheses those hypotheses and all their embodied definitions should be stated as clearly and unambiguously as possible, because only precisely formulated statements can be critically discussed and severely tested (Popper, 1983: 276, 1994: XV, 97-105, 211-218). This refers to the claim of simplicity of scientific hypotheses which is identical with their degree of falsifiability and their empirical content (Popper, 1994: 97-105). The number and the severity of independent tests a hypothesis passes establish its corroboration because the degree of falsifiability of a hypothesis is directly dependent on its testability. The testability of a hypothesis is, in its turn, identical with its empirical content (Popper, 1983: e.g., 230f, 244f, 1994: 213, 339-373). Thus, the degree of falsifiability sets the upper limit of corroboration (Farris, 1995).

Let us focus on three different kinds of phylogenetic hypotheses and discuss whether they are suitable to serve as falsifiable hypotheses: cladistic hypotheses (tree

hypotheses), synapomorphy hypotheses and hypotheses on phylogenetic homology (*sensu* Roth, 1984).

One has to examine whether these hypotheses are logically linked or independent of each other and what could serve as empirical evidence and relevant background knowledge for testing them empirically in a falsificationist approach.

Cladistic and synapomorphy hypotheses are statements about descent of organisms and species. To be able to test such statements one has to have an explanation of what descent means and species are. The theory of evolution gives such an explanation by describing different kinds of evolutionary events, which are: reproduction, heredity and speciation. One has to assume that such events take place and are part of the relevant background knowledge, if one wants to test concrete statements of descent. But, looking for empirical evidence for cladistic and synapomorphy hypotheses one has to concede that no empirical observation clearly and unambiguously indicates how organisms are evolutionary related to one another. Thus, neither a cladistic nor a synapomorphy hypothesis is directly testable on empirical evidence.

This paper represents an attempt to detect the empirical grounds for testing cladistic hypotheses. This is done by referring to the concept of homology and interpreting it from a falsificationist point of view. By examining its empirical testability and its logical linkage to hypotheses of synapomorphy, the conditions for empirically testing cladistic hypotheses are evaluated. The investigations result in a modification of the common interpretation of the application of Popper's methodology on phylogenetic research and of the procedure of parsimony analyses of cladistic characters. As a result, the necessity of *a priori* weighting of cladistic characters according to the results of the character analysis is revealed.

Characteristics

If one wants to test a hypothesis against empirical evidence, the relevant area of matter (domain of discourse) has to be defined beforehand. Therefore, I will use the term *"characteristic"* as corresponding to an observationally distinct unit of an organism. The set of all possible characteristics then constitutes the area of matter relevant to phylogenetic hypotheses. Such a unit can represent a morphological as well as a fossilized or molecular structure, a biochemical compound or a specific behavior. In

every case no connection to the theory of evolution shall be associated with the term “characteristic”.

To be able to use characteristics of organisms as empirical evidence for statements of their history one has to explain what makes some characteristics ‘historical’. Again, the theory of evolution helps. Reproduction together with heredity gives an explanation of the phenomenon of identical characteristics between parents and their offspring. Transformation by mutation is the type of event that explains why offspring sometimes yields new characteristics. Thus, the identity of characteristics can be explained by heredity and their difference by transformation. Therefore, one has to include the possibility of mutations in the assumed background knowledge. This together with the other types of events - reproduction, heredity and speciation - can be summed up by the term *descent with modification*. If one projects these types of events onto time one would expect to perceive organisms bearing characteristics with specific properties. These properties link certain characteristics of one organism with certain characteristics of other organisms by hypothesizing that they have an identical historical origin. This is the essence of the concept of homology.

Homology

A phylogenetic concept of homology should cover its theoretical definition and criteria of identification of homologous characteristics. The theoretical definition of homology must not contradict the assumed background knowledge. It should supply the theoretical foundation for empirical tests of hypotheses of homology and it should ascribe ‘historical characteristics’ to the organisms - properties which allow conclusions about the evolutionary history of those organisms. These characteristics would represent the empirical evidence against which a hypothesis of homology should be tested. The ‘historical characteristics’ are represented by what is called homologies and differ from the characteristics used in a purely comparative, not by any means evolutionary, approach. This historical interpretation of the concept of homology is necessary for the possibility of testing hypotheses about the evolutionary past of lines of descent of organisms.

A concept of homology should assist in uncovering the linkage of the historical with the materialistic thus organismic. This is attempted by means of the theory of *descent with modification* and one receives a theoretical definition of homology:

> Homologous characteristics are shared character states of at least two organisms that are derived from a singular transformation/mutation event in their common ancestor.

Remarks and conclusions:

(1) The here proposed definition of homology uses an *"all-or-none"* theoretical criterion, as Patterson (1988) calls it (concept of quality *sensu* Dover, 1987), for defining the term homology. This should be distinguished from what Dover (1987) calls the *concept of quantity, as in degree of similarity*.

(2) As a consequence, 'their common ancestor' refers to the individual ancestral organism in which the corresponding transformation took place. In that way, the proposed definition of homology differs from the most common definition by referring to the relationships of single individuals (genealogy) rather then of species. I utilize this interpretation of the concept of homology because I want to get conceptually as close as possible to the relevant empirical observations of cladistic inferences, which are, in my opinion, the studied empirical entities and thus properties of individual organisms.

(3) As evolution and 'continuing change' takes place within the lines of individuals and their offspring, and since single organisms are the biological entities in space and time which we study, any effort to reconstruct their evolution should apply the concept of homology to lines of individual organisms. Therefore, it is possible to denote characteristics that are only shared by individuals of a single species (intraspecific) as homologies, as long as they are derived from a singular mutation event in the common ancestor of those organisms.

(4) As hypotheses of homology refer to the pattern causing processes, it is obvious that a complete hypothesis of homology consists of two components:

> a) The distribution pattern of the two character states (see *Phylogenetic Characters*) involved, of which one represents the homologous characteristic which arouse from common ancestry.
> b) The transformation hypothesis consisting of the character state before and after the transformation together with the specific type of transformation-causing mutation event.

Therefore, the proposed concept of homology necessarily entails taxic as well as transformational homology (see Hawkins *et al.*, 1997; de Pinna, 1991; Patterson, 1982).

(5) The result of a mutation event (e.g., an insertion) might function as the substrate for a subsequent mutation event, where the derived character state of an evolutionary older character turns into the ancestral character state of an evolutionary younger one. This is a problem concerning character transformation, hierarchy and the methodologically claimed independence of phylogenetic characters and is relevant to the coding of phylogenetic characters to data matrices.

For example states Haszprunar (1998) that *"there are supraspecific homologous structures based on clearly orthologous genes (e.g., cerebral eyes in metazoans, limbs of arthropods and vertebrates) which are certainly not synapomorphies (e.g., Nilsson, 1996); Shubin et al., 1997)"*. When applying the proposed definition of homology one has, in principle, no problem interpreting this phenomenon without contradiction. The gene in its totality does not necessarily represent a homologous character state. It could also be a composite character state (see Fig. 6). One cannot preclude that it is the result of more than a single mutation event. In this case, it is possible that the gene shares a common origin for all Metazoa in the sense that it (partly) evolved in their common ancestor, which would not be equal to a total homologous property of such a characteristic. Subsequent mutations within the different clades of the Metazoa probably changed the function of that gene. These mutations then resulted in 'new' homologies (limbs of arthropods and limbs of vertebrates).

(6) With this definition it is possible to apply the term homology also for 'loss' character states (contradicting Haszprunar, 1998; de Pinna, 1991). 'Loss' character states, though, cannot be identified prior to the cladistic analysis; this can only be done when interpreting the character state distribution of a rooted most parsimonious tree.

(7) The definition explains the need for a conditional phrase (*"homologues as what?"*) for every homology statement (Bock, 1969, 1973; Patterson, 1982). This conditional phrase is needed to link the historical (the hypothesized mutation event) to the organismic (the observed characteristics). This is done by referring to the common ancestor of the two or more organisms under comparison and the single mutation which took place in this ancestral organism. The use of this conditional phrase, which refers to a specific transformation event, makes a hypothesis of homology empirically testable and the distinction of character and character state not only plausible but necessary (see also de Pinna, 1991; Hawkins *et al.*, 1997; contradicting: Bock, 1973; Wiley, 1981; Schoch, 1986; Ax, 1987; Patterson, 1988).

Phylogenetic Characters

Homology links identical characteristics of different organisms historically by referring to them as to the product of a transformation caused by a mutation event in a common ancestor, inherited by them via a line of reproduction. A transformation is recognizable only by perceiving a difference in a condition found before and after the transformation.
An unequivocal usage of the term *character* and *character state* could be the following:
What we perceive by empirical observations of organisms are only character states. As Hennig (1966: 89) states: *"[Character states] are "characters" in the sense that they distinguish their bearers from one another, but we must always be aware of the fact that "characters" that can be compared are basically only character conditions (...) produced by transformation."* As a consequence, a character, or better *phylogenetic character*, cannot itself be called homologous since it represents only the units of comparison. The term phylogenetic character somehow represents the 'position' within the organism where the mutation occurs and its corresponding transformation. A

phylogenetic character always includes the condition *before* and *after* a transformation. Therefore is a phylogenetic character always a phylogenetic hypothesis (Neff, 1986). The two conditions, the *character states*, are qualitatively different, and are distinguishable in *ancestral* and *derived*, of which only the derived can be denoted as being homologous when shared by at least two organisms. Since the definition of homology refers to a single transformation event, only one of the two character states of a single phylogenetic character represents the homologous character state of this character hypothesis. Therefore, a single phylogenetic character is only a single argument and as such can only provide evidence for the homology of the derived character state of the hypothesized transformation. Independent of this fact, it is possible that the ancestral character state of one phylogenetic character represents the derived and thus homologous character state of another phylogenetic character (see point 5 above). Its homology statement is, however, part of the argument of the latter and not the former phylogenetic character.

This conclusion differs from the common practice of using the term homology and is much stricter. Composite phylogenetic characters that also contain ancestral details, or that are ancestral all together, are often also called homologies (Wägele, 1996). But this practice would weaken the historical aspect of the proposed concept and would furthermore weaken the testability of hypotheses of homology and is therefore rejected (see *Complexity*). One would also have difficulties in stating a single conditional phrase for the homology statement of such a composite character state, which would represent the result of more than a single transformation.

Furthermore, the concept of *phylogenetic character* and *character state* advocated here uses the two terms as representing units of evolutionary processes (like it is used by e.g. Lloyd and Calder, 1991). It must be distinguished from a different concept that uses the two terms as representing (smallest) units of observation (like it is used by e.g. Giribet and Wheeler, 1999), which are called characteristics within this paper.
For the identification of the corresponding character states that belong to a single phylogenetic character, two of the three criteria of Remane (1952) should be applied: similarity of topographical position, which is the *criterion of position*, and similarity of ontogenetic constraints, the *criterion of continuity* (though the criterion of continuity is only applicable on morphological characteristics). They have to be applied to make it

possible to state a hypothesis of topographical correspondence (Rieppel, 1988) of two or more different character states, also called the hypothesis of ‘positional homology’ (Swofford *et al.*, 1996; Titus and Frost, 1996), ‘topographic homology’ (Jardine, 1966), ‘provisional homology’ (Giribet and Wheeler, 1999).

The terms ‘positional homology’, ‘topographic homology’ and ‘provisional homology’ are ambiguous, though, and therefore in a way unsuitable, because a hypothesis of a ‘positional/topographic/provisional homology’ is not a hypothesis of a homology (in this point agreeing with Rieppel (1988) and with Brower and Schawaroch (1996), but not with their arguments) according to the applied definition. A topographic correspondence merely provides the characteristics that shall be compared, which represent something like homolocalities. Therefore, the term “topographical correspondence” is preferred, which does not necessarily need an evolutionary framework – already Aristoteles stated hypotheses of topographical correspondence when giving characteristics of different organisms the same name. Only when inferring phylogenetic characters and testing phylogenetic hypotheses, topographical correspondences are interpreted from an evolutionary perspective. Thus, the possibility of characteristics asserted by a topographical correspondence stemming from a common ancestor is only in the light of the theory of evolution a necessary condition. Stating hypotheses of topographical correspondence does not premise an evolutionary framework.

In this sense a phylogenetic character represents a hypothesis of topographical correspondence consisting of two different character states of which one is interpreted as representing a hypothesis of homology, rendering this character a phylogenetic hypothesis.

Testing Homology: Identity

A criterion of identification of homologous characteristics has to fulfill the following conditions:

- It helps to distinguish between homologous and non-homologous characteristics.
- It does not contradict the assumed background knowledge descent with modification and it is deducible from it and the theoretical definition of homology.

- It gives the opportunity to severe tests of a hypothesis of homology, constituting phylogenetic inference as an empirical science.

Assuming that reproduction, heredity and transformation are the only types of evolutionary events taking place in time, one can conclude the following: if a character state X of organism A is transformed by a mutation and inherited by offspring of A without any subsequent transformation of X, and the offspring reproduces without any subsequent transformation of X, it holds true that for all descendants of A which inherited X, X is identical.
Characteristics of two or more organisms that originated in only a single mutation event in their common ancestor should be identical. Thus, a homologous characteristic is characterized by its identity over all characteristic bearing organisms. Characteristics of two organisms that do not have identical inheritable properties cannot completely have the same origin and can therefore not be explained by a single transformation respectively. Therefore, identity is the test criterion for every hypothesis of homology. This criterion forms the basis for the test of any hypothesis of homology. It is done *a priori* to the construction of a cladogram. As a consequence, hypotheses of homology are principally falsifiable.

Homologous character states have to be identical if they are the result of the same single mutation event. The term *identity* is unambiguously applicable with nucleotide sequence data since this type of data is one-dimensional (Woese, 1987) - at least as long as one does not consider any possible secondary structures of the molecule. If one regards morphological data, this does no longer apply. But one could still recognize morphological characteristics that share identical *'qualities'* - identical structural properties called *'homomorphies'* -, even if their superficial shape is not identical.

Some authors deny that hypotheses of homology could be subject of severe tests but could only be estimated (e.g. Haszprunar, 1998). They assign a relative probability of homology and hence distance themselves from the *"all-or-none"* theoretical criterion. It is true that there is no such thing as a decisive test which could result in a conclusive *yes* or *no*. But this holds true for any empirical hypothesis since there is no basis for verifying any kind of empirical statement (Popper, 1983). And still there is an empirical basis which helps to decide which of all the possible hypotheses of homology are to be

preferred in the light of the present empirical knowledge and this is done by the different *explanatory powers* of the competing hypotheses. The explanatory power of a hypothesis depends on the outcome and the number and severity of independent tests the hypothesis passed, and thus its degree of corroboration (Popper, 1983, 1994).

Though the importance of character analysis *a priori* the cladistic analysis has been stressed by many authors (e. g., Neff, 1986; Bryant, 1989; Wägele, 1994; Haszprunar, 1998), it is often overlooked that characters and character states which are used in the cladistic analysis are already hypotheses themselves and that they had been subject to tests *a priori* the cladistic analysis (Neff, 1986) - tests like that done by the criterion of identity.

The criterion of identity provides indeed the basis for a test (like similarity in Patterson, 1982; and Rieppel, 1988; contradicting de Pinna, 1991) which is, depending on the type of character, more or less severe. Most of all the hypotheses of homology that are theoretically possible in the light of the background knowledge fail this test when exposed to the experiences gained in the character analysis.

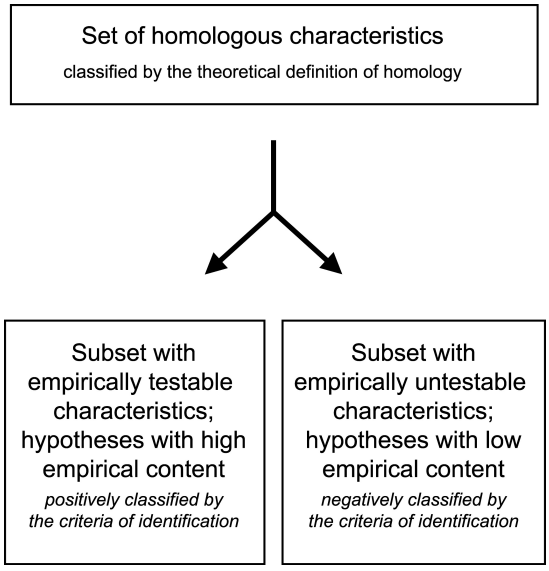

Fig. 1: The set of true homologous characteristics derived from the theoretical definition of homology, and the two subsets which are derived from the criteria of identification. The subsets depend on the empirical testability of the true homologous characteristics.

It is possible that not all homologies in the sense of the definition pass this empirical test. The test is, however, still necessary if homology should be more than a plain theoretical concept which is empirically empty.

One could think of a set of characteristics that are homologous in the sense of the definition. When applying the test subsets of this set are constituted (Fig. 1). The subsets consist on one hand of those hypotheses of homology that are of use for phylogenetic inference

because of their high empirical content (because of the existence of a certain amount of potential falsifiers) and on the other hand of those that are of no use because they cannot be tested (because they lack a sufficient amount of potential falsifiers). They, thus, consist of phylogenetically informative and non-informative characters respectively.

Synapomorphy

In which way is homology and synapomorphy logically linked to one another?

A synapomorphy is a shared character state of two taxa that is derived from a singular transformation event in their common ancestor, the stem species of the two taxa (Hennig, 1966; Ax, 1987). The two taxa in their turn constitute a sister taxon (adelphotaxon).

While homology is a concept that focuses on the organismic level and the history of reproductive lines of individual organisms, synapomorphy is a concept that is applicable on the species level, thus focusing on the history of lines of species.

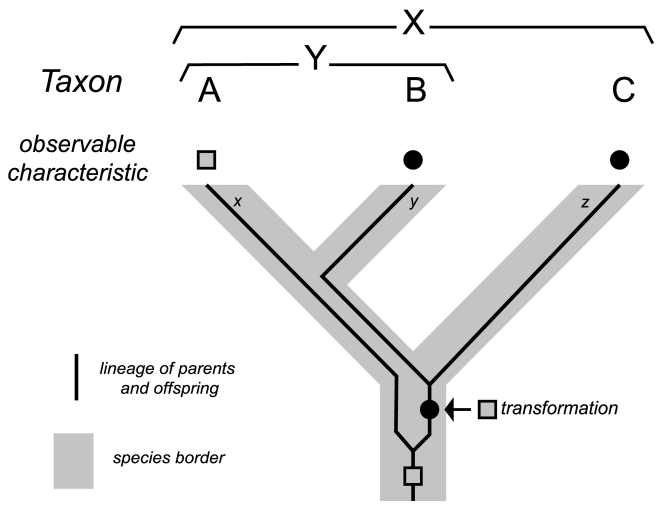

Fig. 2: This illustrates the evolutionary relationship of the taxa A, B and C, of which A and B represent sister taxa. *x*, *y* and *z* stand for three organisms with their individual lineage of descent (genealogy). Organism *y* and *z* share a characteristic '•', which derives from a transformation event in their common ancestor X. Thus, the characteristic '•' represents a homology. Since organism *x* does not have this characteristic and the characteristic never appeared within its line of descent, and since taxa A and B are sister taxa, characteristic '•' cannot represent a synapomorphy.

According to the definition of synapomorphy stated above, a homologous characteristic shared only by organisms of a single species (intraspecific) cannot represent a synapomorphy, since a synapomorphy must be shared by at least two species (interspecific). And even if a homologous characteristic is shared by two or more species, it does not necessarily have to represent a synapomorphy. One could think for instance of the following scenario which is

perfectly consistent with the theory of evolution (for an illustration see Fig. 2): There is a species X, which by geological activities was split into two populations. These two populations are potentially fertile, but, because of a topographic barrier, no gene flow happens between them. In one population a mutation occurs leading to the transformation of a characteristic (Fig. 2: from “square” to “circle”). In exactly the same population another transformation takes place splitting this population into two parts which are fertilely isolated, constituting a speciation. This speciation event gives rise to the two species C and Y. At this point in the scenario, one would thus observe species C with a single population. Its individuals bear the newly transformed characteristic (“circle”). Furthermore, one would observe species Y with two populations which are geographically separated from one another but potentially fertile. Individuals of one of the two populations bear the newly transformed characteristic (“circle”) and those of the other population the ancestral characteristic (“square”), though they both belong to the same species Y. In another step of this scenario, a speciation event takes place in species Y, producing species A and B. Consequently, the scenario ends with three species, of which species A and B represent an adelphotaxon. And the newly transformed characteristic (“circle”) is a homology between individuals of species B and C. But it does not represent a synapomorphy, because species B and C do not resemble an adelphotaxon.

One can thus conclude that a homology is not necessarily also a synapomorphy, but a synapomorphy is always a homology. The set of all synapomorphies is, therefore, a subset of the set of all homologies; and, logically speaking, homology, as it is defined in this paper, cannot be equivalent to synapomorphy (contradicting Patterson, 1982; and de Pinna, 1991; and Brower and Schawaroch, 1996).

In correspondence with the term “phylogenetic character”, which refers to a transformation event leading to a homologous character state, the term *“cladistic character”* is introduced. A cladistic character is a hypothesis of a transformation event which has led to a synapomorphic character state. In this sense, only cladistic characters bear information on the evolutionary relationship of species. They are a subset of the set of phylogenetic characters, which also bear information on the evolutionary relationship of individual organisms.

Testing Synapomorphy: Congruence

As a consequence, one is confronted with the question whether an empirical criterion exists to distinguish synapomorphies from non-synapomorphies.

If one considers the evolutionary events of reproduction, heredity and mutation together with speciation and projects them onto time, one can expect the distribution pattern of apomorphic character states between species to be structured hierarchically. This distribution pattern serves as evidence for nested groupings of species characterized by sharing specific apomorphic character states. Furthermore, those groupings have to be congruent with one another. Hence, synapomorphic character states are not randomly distributed. This criterion is called *congruence* (Patterson, 1988; de Pinna, 1991; Lipscomb, 1992; Kluge, 1997; sensu criterion of coincidence of Wagner, 1986).
Hypotheses of synapomorphy can with this criterion only be tested against other hypotheses of synapomorphy while applying the logical sentence of contradiction and the methodological sentence of parsimony. In a nutshell: if there are two different character states hypothesized as being synapomorphic and their distribution within the species classifies two contradicting sets, at least one of those hypotheses must be wrong in the light of the background knowledge; thus, at least one of them must be interpreted *ad hoc* as being homoplasious (Kluge, 1997, 1998) or homologous, but not synapomorphic. And in terms of testability, it can be said that the more parsimonious the hypothesized species relationships and character state optimizations are the better (Farris, 1970) (see below). Thus, only sets of hypotheses of synapomorphy are principally falsifiable at this level of inference – single hypotheses of synapomorphy cannot be tested by the congruence test.

Many authors call hypotheses of homology that pass this test ‘secondary homology’ (e.g., de Pinna, 1991; equivalent to Rieppel homology, 1988; see also Brower and Schawaroch corroborated homology, 1996); but interpreting the congruence criterion more strictly, it is, actually, not homology that is tested, but synapomorphy.

Only these tested hypothetical synapomorphies *"constitute [the] empirical evidence in phylogenetic systematics"* (Kluge, 1997, according to Hennig, 1966) and from these a cladogram is derived (Bock, 1973). The cladogram which results from a parsimony analysis represents the optimum fit for the hypotheses of synapomorphy within a frame of a hierarchical character state distribution in terms of parsimony, constituting the presently most corroborated cladistic hypothesis. This reveals the logical linkage between the three different types of phylogenetic hypotheses mentioned above: the hypotheses of homology and synapomorphy, and the cladistic hypotheses. Furthermore, it explains the central role of the concept of synapomorphy for cladistic analyses. The fact that the step from tested homologies to tested synapomorphies is not a trivial one is, for instance, illustrated by the lineage sorting and the gene-tree/species-tree problem (e.g., Page and Charleston, 1997).

For effectively performing the congruence test, the amount of hypotheses of homology that are no synapomorphies and that pass the first test must be minimal, so that one can reasonable test the logical consistency of sets of hypotheses of synapomorphy.
Since synapomorphies are always homologies, the identity test is also a test of hypotheses of synapomorphy. Only those synapomorphy hypotheses that successfully pass this test may take part in the congruence test.
The outcome of the cladistic analysis is not independent of the results of the test of identity, and therefore also not independent of the character analysis (see also Neff, 1986). This is due to the fact that the tested hypotheses of synapomorphy themselves, together with the methodological criterion of parsimony and the logical sentence of contradiction, are the basis for this analysis. Parsimony is needed, because without it there would be no methodological criterion which would help in choosing which of the alternative sets of hypotheses that are incongruent with one another should be preferred. If two sets of hypotheses of synapomorphy are incongruent, parsimony prefers the set of hypotheses that has the highest sum of *degrees of corroboration* gained in the identity test. That is the reason why the results of the test of identity have to be employed in the test of congruence.

Because of the methodological principle of parsimony (meant to include not necessarily only the cladistic method of Maximum Parsimony), there is no non-weighting of

characters in a cladistic analysis (contradicting e.g. Kluge, 1997). It also explains why one has no choice but to weight phylogenetic characters in a cladistic analysis, equally or differentially. The function of parsimony within this test is *not primarily* to minimize the requirements for *ad hoc* hypotheses of homoplasy for the most parsimonious hierarchy, but to maximize its degree of corroboration within the given setting of evidence, hypothesis and background knowledge (contradicting Kluge, 1997, 1997a; Farris, 1983) and to minimize the corroboration for all character transformations that have to be interpreted *ad hoc* as homoplasies. The total degree of corroboration of the cladistic hypothesis is not only dependent on the number of *ad hoc* hypotheses of homoplasy, but also on the results of the first *and* the second test.
Hence, for inferring the cladistic hypothesis with the maximum explanatory power it is necessary to include the degrees of corroboration of every single hypothesis of synapomorphy, because they constitute the degree of corroboration of the cladistic hypothesis. Minimizing the number of *ad hoc* hypotheses is not sufficient in that context.

This is the only but decisive theoretical reason for weighting characters and character state transformations differentially a priori to the cladistic analyses in a refutationist program of cladistic research. Only in the case of different degrees of corroboration of hypotheses of synapomorphy which have successfully passed the first test it is necessary to give them differential weights before they are subjected to the second test (in contradiction to Kluge, 1997, 1997a).

Thus, weighting characters is not a question of estimating some intrinsic quality of the character (Neff, 1986) in an essentialist manner, but a question of the degree of corroboration of hypotheses of synapomorphy gained in the character analysis.

It should be mentioned that none of those two tests are truly decisive (see also Patterson, 1988; Kluge, 1997, 1998), because not only synapomorphies/homologies but also convergences can be identical. As a consequence, identical convergences would pass the first test and could also be congruent with other hypotheses of synapomorphy, which would mean that they would also pass the second test.

If the identity test is not interpreted as a test of synapomorphy or homology (see de Pinna's (1991) similarity criterion), the congruence test together with parsimony would not constitute an empirical test in cladistic analyses. Because the congruence test and the principle of parsimony test hypotheses against hypotheses on the basis of their specific degrees of corroboration, they all together have to have passed an empirical test beforehand, as their degree of corroboration would, otherwise, equal zero. And no matter how many hypotheses of synapomorphy contradict a single other hypothesis - if they all have zero corroboration one has no basis for deciding which of the two sets of hypotheses is to be preferred. It is zero versus zero. This is due to the fact that the congruence test is no direct empirical test but a test against the consistency of all hypotheses in question. If this test shall indirectly be an empirical test, the tested hypotheses must have gained explanatory power beforehand. This explanatory power is obtained by successfully passing the identity test.

A supplementary objection remains to be stated: the quality of the hypotheses of topographical correspondence also influences the degree of corroboration of the involved hypotheses of synapomorphy which have passed the identity test, and it might also represent another factor that has to be taken into consideration when evaluating weights of phylogenetic characters and character state transformations.

How to Weight

> *"The wish to grade hypotheses according to the tests passed by them is legitimate: I do not know of any serious objection."* (Popper, 1983: 220)

Popper's hypothetico-deductive approach consists of three elements: background knowledge, hypothesis and empirical evidence (Fig. 3).

The background knowledge (Popper, 1983, 1994) is relevant knowledge that is accepted, while the hypothesis is tested. It may include initial conditions. The important point is that the background knowledge has to be consistent with the hypothesis

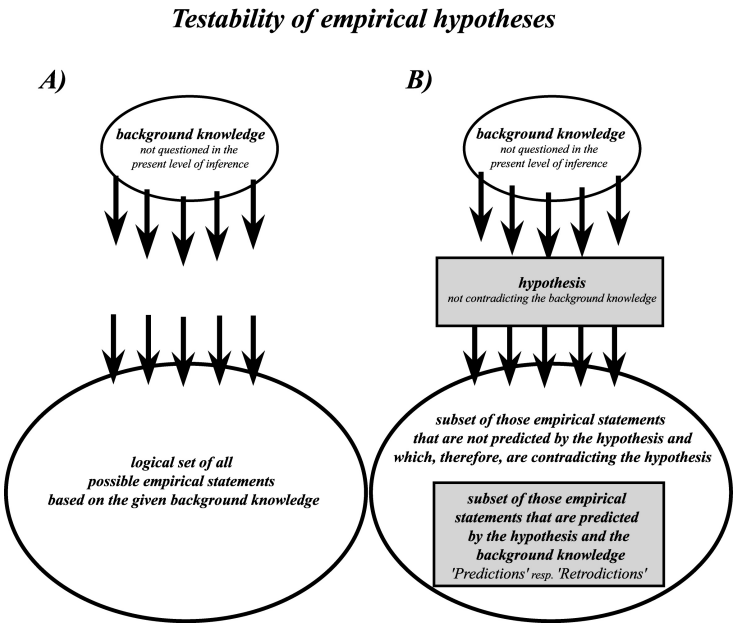

Fig. 3: A) The logical relationship of the background knowledge to its possible empirical statements. B) The function of the hypothesis within the hypothetico-deductive setting. For further explanation see text.

(Popper, 1983: 236). From this background knowledge one can deduce a logical set of all theoretically possible empirical statements that do not contradict the chosen background knowledge. The empirical hypothesis is the hypothesis one wants to test. It must not contradict the background knowledge. If one deduces the theoretically possible empirical statements that are consistent with the conjunction of background knowledge and hypothesis, one receives a subset of the former set. This subset includes all empirical statements that are predicted by the hypothesis (*'predictions'* or *'retrodictions'*). Those statements that only belong to the former set, and that thus contradict the hypothesis but not the background knowledge, constitute the set of potential falsifiers of the hypothesis within the given hypothetico-deductive setting. It is this particular set against which the hypothesis is tested. One can, therefore, say that the more empirical statements a hypothesis prohibits, i.e. the more it ventures, the more severely it is testable and the more it explains potentially. If at all, the testability (or falsifiability – according to which view one takes) of a hypothesis can, therefore, only be measured by the content of the set of potential falsifiers. And following Popper, the empirical content of a hypothesis can be equated to the degree of testability and the degree of falsifiability (*sensu* POPPER, 1983, 1994).

The degree of corroboration of a hypothesis is less dependent on the number of tests passed, which is equal to the total amount of supporting evidence, than on the severity of each test. The severity of the tests depends on the amount of accredited potential falsifiers, which in its turn depends on the degree of falsifiability, or, in other words, on

the degree of testability of the hypothesis which is directly dependent on its empirical content. Popper (1983: 245) proposes the convention that corroborability should equal testability and empirical content.
The empirical content is *"a measure of the class of its falsifiers"* (Popper, 1983: 231) and the hypothesis with the 'larger' class of falsifiers is the hypothesis with the larger empirical content. (Popper, 1994: 77, 211-218; 1983: 230, 244)

The degree of corroboration a hypothesis gains by successfully passing an empirical test can, therefore, be equated to the amount of potential falsifiers that are accredited by this test, and which constitute its severity.

Measuring the Class of Falsifiers of a Hypothesis

In principle, it is not possible to give an absolute and objective measure of the degree of falsifiability of a hypothesis or the severity of a test. However, in some special cases a relative measure of the amount of falsifiers accredited by a test of two alternative hypotheses can be given.

Hypothesis *x* has a higher degree of corroboration then hypothesis *y* if - and only if - the class of possible falsifiers of hypothesis *x* includes the possible falsifiers of hypothesis *y* as a true subclass (Popper, 1994: 80) and both hypotheses have successfully passed the same severe test.

Unfortunately, this concept is not applicable to hypotheses of homology and synapomorphy, as they represent historical and therefore singular statements, and as every such hypothesis has its own and individual objectively incommensurable class of possible falsifiers. *Thus, there is no strictly objective foundation for the introduction of a quantitative system to determine the severity of tests, the degrees of corroboration of hypotheses or an a priori weighting system of cladistic characters respectively.* In this sense, Kluge (1997a) is right when concluding that there is no theoretical basis for an *a priori* character weighting and that an exact value of the degree of corroboration of a hypothesis cannot be determined (Kluge, 1997).

Nevertheless, *a priori* weights have to be applied inevitably if one wants to perform a cladistic analysis on empirical grounds. Such a weighting is either an equal or a differential weighting. Due to the fact that hypotheses of synapomorphy are tested against one another within the cladistic analysis, their respective degree of corroboration represents the decisive criterion in the case of conflict for the cladistic analysis (see above). The degrees of corroboration of those hypotheses of synapomorphy which have passed the test of identity are, therefore, compared with one another quantitatively. This comparison is immanent to any weighting scheme– whether equal or differential. For this reason one has to weight!

So, we still have to ask ourselves, what one should do.

One should attempt to approximate the degrees of severity of each identity test.

One could categorize the possible falsifiers of a hypothesis into *classes of identity*, classes of identical units of observation. The resulting number of classes would be directly dependent on the number of theoretically possible different results of a single mutation event. Because hypotheses of synapomorphy refer to mutation events, the possible falsifiers of these hypotheses, and therefore also the units of observation which come into question, have to refer to the possible results of the hypothesized mutation event.

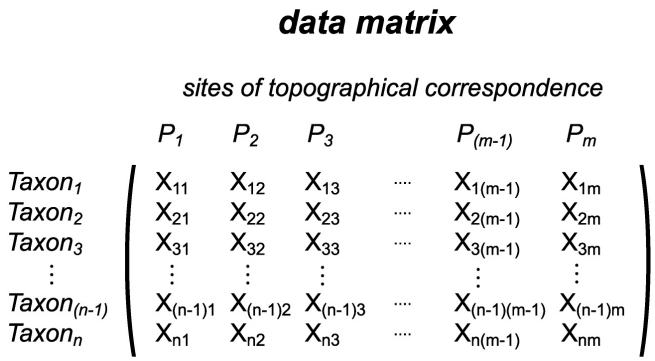

Fig. 4: A matrix for *n* taxa and *m* sites of topographical correspondence. For details see text.

For an 'empty' data matrix this would mean (Fig. 4) that, *potentially*, many different kinds of hypotheses of synapomorphy could be stated for every single position in a column of a matrix. Such a statement - for instance for site P_1, position X_{11} and an observable characteristic *y* - could look as follows: *"X_{11} is the result of a mutation event that has taken place in the last common ancestor of $Taxon_1$ and all the other taxa that share the same type of character state y for the site P_1 of topographical correspondence"*. The statement would be falsified according to the

criterion of identity if, and only if, $X_{11} \neq y$. Thus, the first test of synapomorphy is independent of the number and the sample of taxa used in the data matrix and the cladistic analysis. Falsifiers are all character states that are non-*y* and the number of different falsifiers is determined by the number of possible character states of the corresponding mutation.

An example for the classification into classes of identity can be taken from molecular data. For instance for a nucleotide substitution, every individual nucleotide is represented by its corresponding class of identity - for adenine A, for guanine G, for cytosine C and for thymine T. One would, thus, receive three classes of observationally different, possible falsifiers for a given sequence position, because there are four possible results from any nucleotide substitution event. One of them would stand for the hypothesis and the remaining ones would be the potential falsifiers in the sense of the criterion of identity. With such a classification one can, at least in theory, measure the classes of possible different results of a mutation event quantitatively, thus measuring the number of classes of different falsifying statements for every hypothesis of synapomorphy. These classes of identity would be true subclasses of the class of all possible falsifiers of a hypothesis.

If one wishes to follow the convenient phylogenetic interpretation of Popper's falsificationist approach and, therefore, refuses to consider process probabilities as part of the relevant background knowledge, one could, as a 'null hypothesis', ascribe every such class of identity the same weight in terms of empirical content, which would lead to an *equal weighting* of these classes. The severity of every test of identity could be quantified, which would in turn quantify the degrees of corroboration of every hypothesis of synapomorphy that successfully passed the first test. These weights should be included in the cladistic test where the most parsimonious character state distribution for a given phylogeny is inferred (Fig. 5 gives an overlook of the whole procedure). In the case of two incongruent hypotheses of synapomorphy, their degrees of corroboration from the first test give the basis to decide which hypothesis is to be preferred – i.e. which hypothesis has maximum explanatory power. The degrees of all the other hypotheses of synapomorphy and their most parsimonious character state

distribution must of course be taken into account, to receive the globally most parsimonious solution.

Transferring this concept to the example of the 'empty' data matrix would lead to the following result: when the matrix has been 'filled' with empirical content derived from the character analysis, every *position* of a column entails only a part of *one* hypothesis of a putative synapomorphy. This means that, within this step, all other possible hypotheses of synapomorphy are excluded from this special site of topographical correspondence and are, hence, falsified. When the character states are coded and filled into the data matrix the corresponding hypotheses of synapomorphy has therefore already passed the first test. The result of this test should be included in the data matrix by giving every cladistic character a weight corresponding to the severity of the successfully passed test (or, in applying a stepmartix, a weight for the corresponding transformation respectively).

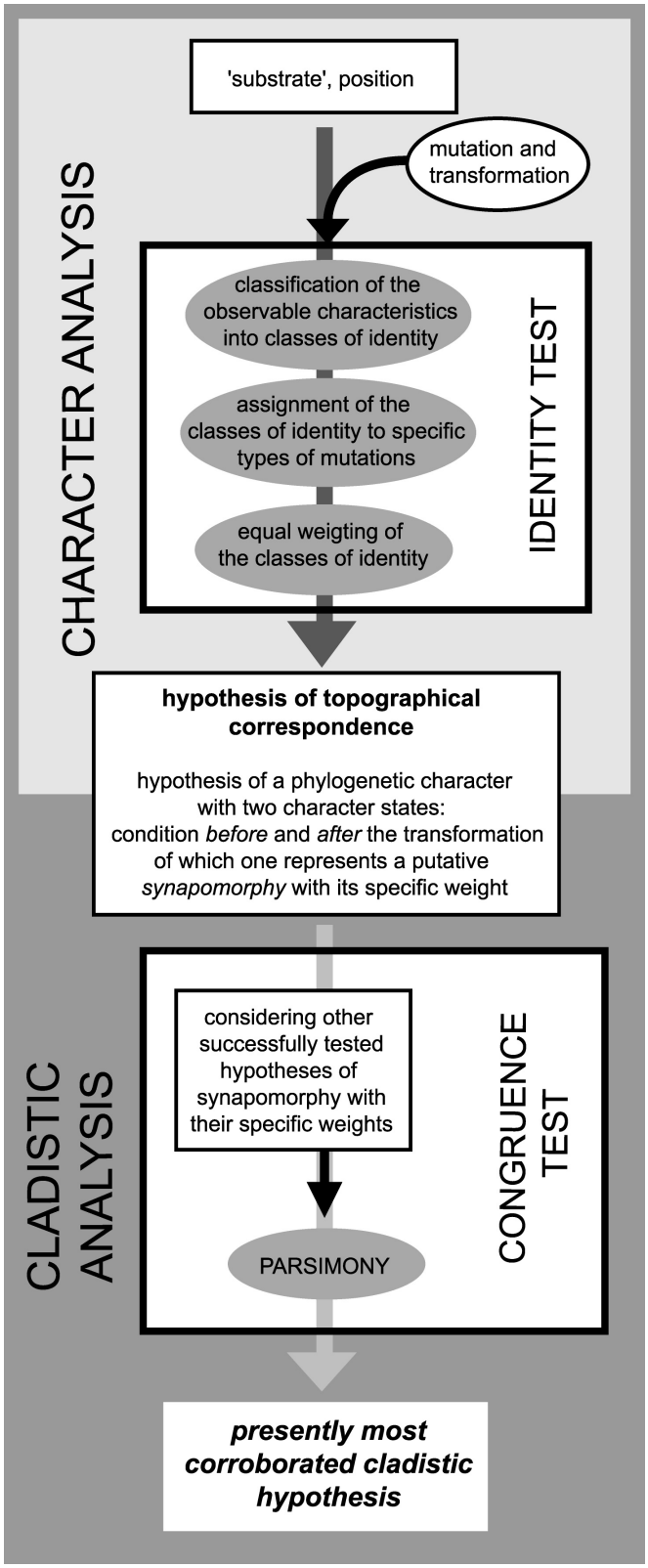

Fig. 5: Scheme of the procedure of cladistic research and differential *a priori* character weighting. For details see text.

Many authors postulate the reduction of the background

knowledge to descent with modification only to avoid other conditions relating to pattern and process. This facilitates the critical inference of other theories on the basis of the results of the phylogenetic inference without the problem of circular reasoning when, for instance, applying rates and patterns of character evolution (Kluge, 1997; see also Sluys, 1996). Giving each of the classes of identical falsifiers the same weight would correspond with this claim.

One can thus summarize: the severity of the first synapomorphy test is dependent on the character type, which is the type of hypothesized transformation event, and on the number of different possible results. Besides, the procedure also has consequences on the background knowledge one has to assume: the assumptions *'character types can be classified according to the type of mutation/transformation which caused them and the constraints resulting from the inference of the topographical correspondence'* and *'the classes of observationally identical potential character states have all the same information content and are therefore weighted equally'* have to be added to the relevant background knowledge. The former is already necessary for stating a hypothesis of topographical correspondence, especially when analyzing molecular data. The latter represents a convention or a null hypothesis respectively, which neglects the effect that the consideration of different process probabilities would have an effect on the outcome of the analysis. It is the crucial assumption of the entire weighting scheme that has been proposed and requires severe testing.
The proposed classification provides, however, at least a clear and unambiguous foundation for all kinds of different tests of *a priori* weighting systems and for discussing their advantages.

Complexity

Complexity is a term applied when interpreting characteristics, especially morphological characteristcs. However, complexity depends on the individual 'eye of the observer'. An event with a rather simple structure is often perceived as being very complex as long as one does not know its mechanics and causality. In this sense, complexity would be a concept extremely open to subjectivity, which belongs methodologically rather to description than to analysis.

If one wants to use the term complexity in the context of justifying weights of characters at all, it should be used in the sense of the complexity of the transformation event to which the severity of the first test is related rather than the complexity of the organismic structure of the characteristics (contradicting Neff, 1986; and Patterson, 1988). One should not simply equate structural complexity with the complexity of an underlying event. There is no theoretical nor methodological basis available yet for weighting characters simply according to their structural complexity within a refutationist program of phylogenetic inference.

Some authors (e. g., Wägele, 1995, 1996) state that the phylogenetic information content of a character state is higher when it is caused by a larger number of specific mutations, constituting a complex molecular character. This would represent an useful

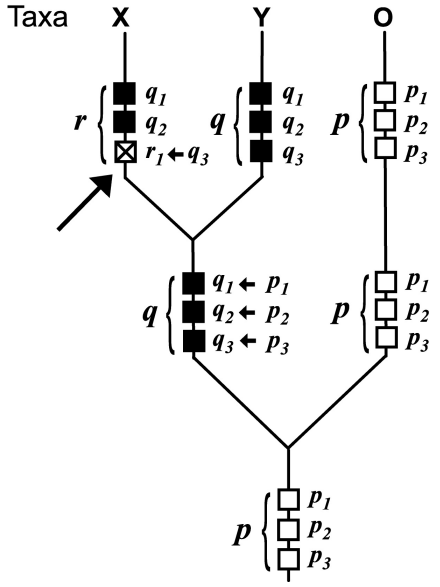

Fig. 6: Consequences and problems with the coding of results of multiple, independent transformation events as 'composite characters' or 'composite character states'. **X** and **Y** represent sister taxa, **O** the nearest related outgroup; ***p*** recodes three independent plesiomorphic character states which are represented by p_1-p_3; ***q*** recodes three independent synapomorphic character states which are represented by q_1-q_3 and which support the sister group relationship of the taxa **X** and **Y**; ***r*** recodes two synapomorphic character states which are represented by q_1 and q_2 and which support the sister group relationship of the taxa **X** and **Y**, and one autapomorphic character state r_1 which supports the monophyly of taxon **X**. When using the character states p_1-p_3, q_1-q_3 and r_1 in a cladistic analysis one would reconstruct the true phylogeny of **X**, **Y** and **O**. If one uses the recoded 'composite character states' ***p***, ***q*** and ***r***, one would not be able to reconstruct this relationship: Though the 'composite character state' ***r*** supports the true monophyletic condition of taxon **X,** it is only a single transformation (r_1) that, empirically, supports only this relationship and not three. This coding would, thus, lead to an artificially high weighting of ***r***. The 'composite character state' ***q*** would support a monophyletic condition of taxon **Y** without any empirical evidence and the sister group relationship of taxa **X** and **Y** is not supported at all by that type of character coding, in spite of the fact that empirical evidence (q_1 and q_2) does exist.

approach to character weighting as long as history, and thus evolution, ends after such a multitude of mutation events and all mutations take place in a single ancestor species only. However, if these two conditions are not fulfilled, it is difficult to treat such a ‘composite character state’ correctly. A single mutation event in only one of the positions where the former mutations occurred would already transform it into a new ‘composite character state’. Much of the phylogenetic information would be lost in this type of character state coding (Fig. 6). As a consequence, this new ‘composite character state’ has to be given the same weight as the old ‘composite character state’, since the former is in the same way the result of a multitude of mutation events as the latter, even though only a single mutation separates the apomorphic from its plesiomorphic condition. Hence, this newly transformed ‘composite character state’ would support only a single phylogenetic grouping - and with an artificially high weight -, though phylogenetic information for another grouping is available. Information on, for instance, a sister group relationship could thus get lost (Fig. 6). Furthermore, in the worst case, this coding could lead to the support of artificial groupings. This problem can be caused when plesiomorphic and apomorphic character states are combined and form a ‘composite character state’. And as the qualities ‘plesiomorphic’ and ‘apomorphic’ can only be assigned *a posteriori* to the cladistic analysis, such a ‘composite character state’ coding should be avoided whenever possible. The described problem affects morphological as well as molecular characters.

Final Comments

To demonstrate that phylogenetic research is an empirical science *sensu* Popper, the empirical basis for tests of cladistic hypotheses has to be revealed, and, thus, their logical linkage to hypotheses of synapomorphy and homology. It is necessary to go back to the concept of homology to get conceptually as close as possible to the relevant empirical observations.

The proposed conception of phylogenetic homology together with the criterion of identity represents a setting suitable for direct empirical testing within a falsificationist approach. This is due to the property of this conception being related to the individual history of the organisms that are studied and their corresponding genealogy. It utilizes the distribution pattern of identical and different characteristics of the investigated organisms to state hypotheses of the history of those characteristics and, therefore,

indirectly of the organisms carrying them. Thus, one concludes from the reconstructed history of the characteristics to the history of the organisms bearing those characteristics and from them to the history of the corresponding species. By examining the structure of characteristics and hypothesizing specific types of transformations responsible for having caused them, a specific degree of severity is assigned to each identity test and thus a specific degree of corroboration for every hypotheses that passes the test successfully.

Character analysis is, therefore, the moment within the procedure of phylogenetic research where a direct test on empirical observations is performed. This is the source from which cladistic hypotheses receive their empirical corroboration at the end of the analysis. This is why the potentially available amount of explanatory power for the presently most corroborated cladistic hypothesis is directly restricted by the quality and results of the character analysis.

Since hypotheses of homology do not represent cladistic hypotheses, the logical linkage of synapomorphy and homology has to be revealed. It is shown that they are not equivalent to each other, but that the former represents a subclass of the latter. Putative synapomorphies can, therefore, also be empirically tested by the criterion of identity. This is very important, because otherwise the congruence test would not constitute an empirical test. Only in the case where the hypotheses that are tested against each other have successfully passed an empirical test before, in which they gained corroboration, the congruence test is not an empirically empty test. And only in this case are cladistic hypotheses empirically testable and is phylogenetic research an empirical science.

This paper is also an attempt to evaluate the conditions for the possibility to justify a procedure within a falsificationist approach to phylogeny that is intimately linked to the circumstances stated above: the necessity to weight cladistic characters within a cladistic analysis.

And as the congruence criterion tests hypotheses of synapomorphy against each other on grounds of their degree of corroboration gained from the identity test, these different degrees of corroboration determine the specific weights given to phylogenetic characters and character state transformations before the cladistic analysis. This is the only reasonable justification for a weighting scheme and it also demonstrates the indispensable necessity of its application. There is no non-weighting but only *a priori* weighting of phylogenetic characters in a falsificationist approach to phylogeny.

A convention to classify cladistic characters in correspondence with the transformation type that has caused them and a convention to classify character states in respect to their identity, had to be set up. This is not problematic since it does not effect a possible weighting scheme directly. To be able to justify a weighting scheme, a proposal has to be stated to the effect that every class of possible identical character states has the same weight in terms of its explanatory power and empirical content. This is a problematic proposal, since it disregards any effect of different process probabilities of transformations and needs severe testing on empirical data. A first test is presented in Vogt (*Weighting Indels as Phylogenetic Markers of 18S rDNA Sequences in Diptera and Strepsiptera)*.
On the whole, the advantage of the proposed scheme lies, however, in its clearly and unequivocally stated assumptions. This enables one to discuss them and their alternatives as well as their impact on the methods of cladistic analysis.

As a consequence of the presented investigation, the background knowledge that has necessarily to be assumed in every cladistic inference that corresponds with a falsificationist approach can be summarized as follows:

1) logical sentence of contradiction
2) methodological criterion of parsimony
3) descent with modification, including knowledge about reproduction, heredity, speciation and mutation/transformation
4) determination of the relevant area of matter (domain of discourse), including the assumption of the observability of characteristics, the categorization of identical character states into classes of identity, the assignment of the classes of identity to specific types of mutations
5) the proposal to weight all classes of identity equally in terms of determining the severity of the identity test (if one wishes to follow the conventional interpretation of Popper's falsificationism for phylogenetic research)

When considering morphological data, the necessity of assumption 4), in particular, gives rise for well founded concern. The knowledge of the genetic linkage and of the

mechanisms of transformation of morphological characteristics is very small. This knowledge is, however, fundamental to being able to interpret the phylogenetic information content of morphological characteristics and thus to performing a more or less effective cladistic analysis. Another problem is the necessity of developing a methodology that allows a more objective description of morphological characteristics, so as to be able to test them more severely on identity. A cladistic analysis of morphological data is, therefore, only reasonable in those cases where the number of homoplasies after the character analysis is very low and the general phylogenetic information content thus comparably high. Even though this sounds like a serious blow, there are still a number of promising approaches to the possibility of evaluating the quality of morphological characteristics methodologically within the character analysis (e.g. Neff, 1986). This aspect will require further investigation.

ACKNOWLEDGMENTS

I want to thank A. G. Kluge and K. M. Kjer who read an earlier draft of the manuscript offering many valuable comments and M. Hendy and A. Dress for discussing some of my conclusions. I am particularly indebted to T. Bartolomaeus, P. Ax and W. Ahlrichs for discussing and criticizing my views of the concept of homology and the nature of the congruence test extensively. My conclusions concerning *a priori* weighting and the value of indels as phylogenetic markers were first presented at the XVIIIth meeting of the Willi Hennig Society, 1999, in Göttingen.

This study was supported by the Deutsche Forschungsgemeinschaft (BA 1520/4-1).

References

Allard, M. W. and Carpenter, J. M. (1996): On weighting and congruence. Cladistics 12: 183-198.

Ax, P. (1987): The phylogenetic system. John Wiley & Sons, Chichester, England.

Björklund, M. (1999): Are third positions really that bad? A test using vertebrate cytochrome b. Cladistics 15: 191-197.

Bock, W. J. (1969): Comparative morphology in systematics. *In* "Systematic Biology, Proceedings of an International Conference Conducted at the University of Michigan, Ann Arbour, Michigan, June 13-16, 1967", pp.411-448. National Academy of Science, Washington.

Bock, W. J. (1973): Philosophical Foundations of Classical Evolutionary Classification. Syst. Zool. 22: 375-392.

Broughton, R. E., Stanley, S. E., and Durrett, R. (2000): Quantification of homoplasy for nucleotide transitions and transversions and a reexamination of assumptions in weighted phylogenetic analysis. Syst. Biol. 49: 617-627.

Brower, A. V. Z. (2000): Evolution is not a necessary assumption of cladistics. Cladistics 16: 143-154.

Brower, A. V. Z., and Schawaroch, V. (1996): Three Steps of Homology Assessment. Cladistics 12: 265-272.

Bryant, H. N. (1989): An evaluation of cladistic and character analyses as hypothetico-deductive procedures, and the consequences for character weighting. Syst. Zool. 38: 214-227.

Carpenter, J. M. (1988): Choosing among multiple equally parsimonious cladograms. Cladistics 4: 291-296.

Chippindale, P. T. and Wiens, J. J. (1994): Weighting, partitioning, and combining characters in phylogenetic analysis. Syst. Biol. 43: 278-287.

Cracraft, J. (1978): Science, Philosophy, and Systematics. Syst. Zool. 27: 213-216.

de Pinna, M. C. C. (1991) : Concepts and tests of homology in the cladistic paradigm. Cladistics 7: 367-394.

De Queiroz, K. and Poe, S. (2001) : Philosophy and phylogenetic inference : A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. Syst. Biol. 50: 305-321.

Dover, G. A. (1987): Letter to the editor. Cell 51: 515.

Farris, J. S. (1969): A successive approcimations approach to character weighting. Syst. Zool. 18: 374-385.

Farris, J. S. (1970): Methods for computing Wagner trees. Syst. Zool. 19: 83-92.

Farris, J. S. (1979): The information content of the phylogenetic system. Syst. Zool. 28: 483-519.

Farris, J. S. (1983): The logical basis of phylogenetic analysis. *In* "Advances in Cladistics" (N. J. Platnick, and V. A. Funk, Eds), Vol. 2, pp. 7-36. Columbia University Press, New York.

Farris, J. S. (1995): Conjectures and refutations. Cladistics 11: 105-118.

Farris, J. S. (2000): Corroboration versus "strongest evidence". Cladistics 16: 385-393.

Giribet, G., and Wheeler, W. C. (1999): On Gaps. Mol. Phylogenet. Evol. 13: 132-143.

Goloboff, P. A. (1993): Estimating character weights during tree search. Cladistics 9: 83-91.

Haszprunar, G. (1998): Parsimony Analysis as a Specific Kind of Homology Estimation and the Implications for Character Weighting. Mol. Phylogenet. Evol. 9: 333-339.

Hawkins, J. A., Hughes, C. E., Scotland, R. W. (1997): Primary homology assessment, characters and character states. Cladistics 13: 275-283.

Hennig, W. (1966): Phylogenetic Systematics. University of Illinois Press, Urbana, Illinois.

Jardine, N. (1969): The observational and theoretical components of homology: a study on the morphology of the dermal skull-roofs of rhidpidistian fishes. Biol. J. Linn. Soc. 1: 327-361.

Källersjö, M., Albert, V. A., and Farris, J. S. (1999): Homoplasy increases phylogenetic structure. Cladistics 15: 91-93.

Kitts, D. B. (1977): Karl Popper, Verifiability, and Systematic Zoology. Syst. Zool. 26: 185-194.

Kluge, A. G., and Wolf, A. J. (1993): Cladistics: What's in a world? Cladistics 9: 183-199.

Kluge, A. G. (1997): Testability and the Refutation and Corroboration of Cladistic Hypotheses. Cladistics 13: 81-96.

Kluge, A. G. (1997a): Sophisticated falsification and research cycles: Consequences for differential character weighting in phylogenetic systematics. Zool. Scr. 26: 349-360.

Kluge, A. G. (1998): Total Evidence Or Taxonomic Congruence: Cladistics Or Consensus Classification. Cladistics 14: 151-158.

Kluge, A. G. (2001): Philosophical conjectures and their refutation. Syst. Biol. 50: 322-330.

Lipscomb, D. L. (1992): Parsimony, homology and the analysis of multistate characters. Cladistics 8: 45-65.

Lloyd, D. G., and Calder, V. L. (1991): Multi-residue gaps, a class of molecular characters with exceptional reliability for phylogenetic analyses. J. Evolution. Biol.4: 9-21.

Lutzoni, F., Wagner, P., Reeb, V., and Zoller, S. (2000): Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst. Biol. 49: 628-651.

Milinkovitch, M. C., LeDuc ,R. G., Adachi, J., Farnir, F., Georges, M., and Hasegawa, M. (1996): Effects of Character Weighting and Species Sampling on Phylogeny Reconstruction: A Case Study Based on DNA Sequence Data in Cetaceans. Genetics 144: 1817-1833.

Neff, N. A. (1986): A rational basis for a priori character weighting. Syst. Zool. 35: 110-123.

Nilsson, D. E. (1996): Eye – ancestry: Old genes for new eyes. Curr. Biol.6: 39-42.

Page, R. D. M., and Charleston, M. A. (1997): From Gene to Organismal Phylogeny: Reconciled Trees and the Gene Tree/Species Tree Problem. Mol. Phylogenet. Evol. 7: 231-240.

Patterson, C. (1982): Morphological characters and homology. *In* "Problems of phylogenetic reconstruction" (K. A. Joysey and A. E. Friday, Eds.), pp. 21-74. Academic Press, London.

Patterson, C. (1988): Homology in Classical and Molecular Biology. Mol. Biol. Evol. 5: 603-625.

Platnick, N. I. (1979): Philosophy and the transformation of cladistics. Syst. Zool. 28: 537-546.

Platnick, N. I., and Gaffney, E. S. (1977): Systematics: A Popperian Perspective. Syst. Zool. 26: 360-365.

Platnick, N. I., and Gaffney, E. S. (1978a): Evolutionary Biology: A Popperian Perspective. Syst. Zool. 27: 137-141.

Platnick, N. I., and Gaffney, E. S. (1978b): Systematics and the Popperian Paradigm. Syst. Zool. 27: 381-388.

Popper, K. R. (1983): Realism and the Aim of Science; From the Postscript to the Logic of Scientific Discovery. London and New York: Routledge,.

Popper, K. R. (1994): Logik der Forschung. (10., verbesserte und vermehrte Auflage von "Die Einheit der Gesellschaftswissenschaften; Band 4). Tübingen: J. C. B. Mohr (Paul Siebeck).

Remane, A. (1952): Die Grundlagen des Natürlichen Systems, der Vergleichenden Anatomie und der Phylogenetik. Leipzig 1952.

Rieppel, O. C. (1988): Fundamentals of Comparative Biology. Birkhäuser Verlag, Basel.

Rieppel, O. (1999): Einführung in die computergestützte Kladistik. Verlag Dr. Friedrich Pfeil, München.

Roth, V. L. (1984): On Homology. Biol. J. Linn. Soc 22: 13-29.

Schoch, R. M. (1986): Phylogeny reconstruction in paleontology. Van Nostrand-Rheingold, New York.

Shubin, N., Tabin, C., and Carroll, S. (1997): Fossils, genes and the evolution of animals limbs. Nature 388: 639-648.

Siddall, M. E., and Kluge, A. G. (1997): Probabilism and Phylogenetic Inference. Cladistics 13: 313-336.

Sluys, R. (1996): The notion of homology in current comparative biology. Z. Zool. Syst. Evol. 34: 145-152.

Sober, E. (1993): Experimental Tests of Phylogenetic Inference Methods. Syst. Biol. 42: 85-89.

Swofford, D. L., Olsen, G. J., Waddel, P. J., and Hillis, D. M. (1996): Phylogenetic inference. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz and B. K. Mable, Eds.) Second ed., pp. 407-514. Sinauer, Sunderland, Massachusetts.

Titus, T. A., and Frost, D. R. (1996): Molecular Homology Assessment and Phylogeny in the Lizard Family Opluridae (Squamata: Iguania). Mol. Phylogenet. Evol. 6: 49-62.

Vogt, L. (*Weighting Indels as Phylogenetic Markers of 18S rDNA Sequences in Diptera and Strepsiptera)* - handed in together with this manuscript

Wagner, G. P. (1986): The systems approach: an interface between development and population genetic aspects of evolution. *In* "Patterns and processes in the history of life" (D. M. Raup and D. Jablonski, Eds.), pp. 149-165. Springer, Berlin.

Wägele, J. W. (1994): Review of methodological problems of ´Computer cladistics` exemplified with a case study on isopod phylogeny (Crustacea: Isopoda). Z. Zool. Syst. Evol. 32: 81-107.

Wägele, J. W. (1995): On the information content of characters in comparative morphology and molecular systematics. Z. Zool. Syst. Evol. 33: 42-47.

Wägele, J. W. (1996): First principles of phylogenetic systematics, a basis for numerical methods used for morphological and molecular characters. Vie Milieu46 : 125-138.

Wenzel, J. W., and Siddall, M. E. (1999): Noise. Cladistics 15: 51-64.

Wiley, E. O. (1975): Karl R. Popper, Systematics, and Classification: A Reply To Walter Bock and Other Evolutionary Taxonomists. Syst. Zool. 24: 233-243.

Wiley, E. O. (1981): Phylogenetics. John Wiley & Sons, New York.

Woese, C. R. (1987): Bacterial evolution. Microbiol. Rev.51: 221-271.

Weighting Indels as Phylogenetic Markers of 18S rDNA Sequences in Diptera and Strepsiptera

L. Vogt

Zoomorphologie und Systematik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, E-Mail: lars.vogt@biologie.uni-bielefeld.de

eingereicht und akzeptiert in:

Organisms, Diversity & Evolution

Abstract

Opinions split when it comes to the significance and thus the weighting of indel characters as phylogenetic markers. This paper attempts to test the phylogenetic information content of indels and nucleotide substitutions by proposing an *a priori* weighting system of non protein coding genes. Theoretically, the system rests on a weighting scheme which is based on a falsificationist approach to cladistic inference. It provides insertions, deletions and nucleotide substitutions weights according to their specific number of identical classes of potential falsifiers, resulting in the following system: nucleotide substitutions weight = 3, deletions of *n* nucleotides weight = $(2^n - 1)$ and insertions of *n* nucleotides weight = $(5^n - 1)$. This weighting system and the utility of indels as phylogenetic markers are tested against a suitable data set of 18S rDNA sequences of Diptera and Strepsiptera taxa together with other metazoen species. The indels support the same clades as the nucleotide substitution data, and the application of the weighting system increases the corresponding consistency indices of the differentially weighted character types. As a consequence, applying the weighting system seems to be reasonable and indels appear to be good phylogenetic markers.

[18S rDNA, *a priori* weighting, Diptera, falsification, identity test, indel]

Introduction

In most phylogenetic analyses of molecular data sets it is customary that only nucleotide substitution positions are utilized for cladistic analyses (e.g. Aguinaldo *et al.*, 1997; Aleshin *et al.*, 1998; Zrzavý *et al.*, 1998; Canapa *et al.*, 2000). However, nucleotide substitution characters have the disadvantage of exhibiting frequent parallelisms and reversions due to their low information content (e.g., Lloyd and Calder, 1991; Wägele, 1996). Insertions and deletions are usually not used as characters, though their potential value as phylogenetic markers has been repeatedly pointed out by some authors (Hixson and Brown, 1986; Meyer *et al.*, 1986; Williams and Goodman, 1989; Giribet and Wheeler, 1999; Mitchison, 1999). This potential value is assigned due to the complex mutational mechanism which causes indels (e. g., Van Dijk *et al.*, 1999) as well as their comparably high immunity against reversals and parallelisms. This immunity is ascribed to the dependence of those processes on the position and the length of the indel and, as in the case of insertions, also on their specific nucleotide sequence (Lloyd and Calder, 1991). Excluding indels from the cladistic analysis could thus reduce the explanatory power of cladistic hypotheses since insertions and deletions represent phylogenetically significant historical information (Giribet and Wheeler, 1999). In a few analyses indels are included in the data matrix, but their usage is, however, diverse, reaching from coding them as a fifth character state (e.g., Swofford, 1998) to recoding them to absent-present characters (e.g., Baum *et al.*, 1998; Simmons and Ochoterena, 2000; Lutzoni *et al.*, 2000), and their weightings are comparably diverse.

In the following a weighting system will be proposed as well as the utilization of indels as phylogenetic markers will be tested with a suitable data set consisting of 18S rDNA sequences of 14 Diptera species and of several other Insecta and Metazoa taxa. Many other sequences serve as an outgroup. The inference focuses particularly on insertions and deletions that have taken place within the Diptera and the Strepsiptera clade. The data set seams to be suitable since it exhibits a comparably high amount of gaps within its alignment which indicates a multitude of indel events that must have taken place within this clades evolution. Moreover, the analysis of the nucleotide substitution data in itself provides a phylogeny with a high resolution and, when taking the

corresponding bootstrap frequencies into consideration, also stable internal nodes, thereby providing a suitable basis for the test.

Weighting Cladistic Characters of 18S rDNA Sequences

This paper proposes a weighting system, which weights all nucleotide substitutions equally, while weighting all insertions and deletions differently according to their length. It is based on the theoretical and methodological arguments presented in Vogt (*Testing and weighting characters*), following Popper's falsificationism. The basic idea is, that when the results of the character analysis are coded for the data matrix, the stated hypotheses of synapomorphy have already passed a first empirical test (identity test). By passing this test successfully the hypotheses gain corroboration, the degree of corroboration gained depending on the severity of the test (Vogt, *Testing and weighting characters*). This is an important aspect, as without this prior test the cladistic analysis, in which only the congruence test together with the methodological principle of parsimony is applied, would lack its empirical basis. The congruence test tests hypotheses against other hypotheses as to their consistency. Only if the tested hypotheses have already gained corroboration by passing a previous test successfully does the congruence test together with parsimony represent an empirical test. The only test that could serve as such an empirical test is the test of identity (comparable to the similarity test of Patterson, 1982; and Rieppel, 1988). And since the hypotheses are tested against one another in respect to their specific degrees of corroboration, one has to evaluate whether the identity test is equally severe on every hypothesis of synapomorphy (Vogt, *Testing and weighting characters)*. This is necessary, as the degree of corroboration of a hypothesis depends on the number and severity of independent tests passed by this hypothesis (Popper, 1983, 1994); whereas the severity of the test depends on the measure of the class of potential falsifiers of the specific hypothesis it employs. Only in case of the severity of the identity test being equal to every type of hypothesis of synapomorphy, i.e. when the measure of possible falsifiers is identical, one could undertake a cladistic analysis with equally weighted characters (Vogt, *Testing and weighting characters*). When considering this for sequence data, one has to take a look at the different types of cladistic characters as well as the range of character states corresponding with each of these types, both being entailed in sequences of non-protein coding genes.

Under the premise the alignment is correct, one can easily distinguish nucleotide substitution characters from indel characters. All columns in an alignment that have no gaps can be referred to as representing characters of the type ‘nucleotide substitution’. Indels are represented by those columns in an alignment that exhibit gaps. They differ from one another by their length of directly neighboring gap sites. To be able to distinguish gap columns into deletions and insertions, an outgroup comparison with an adequate outgroup has to be applied.

In a next step the number of different types of possible falsifiers for each character type is evaluated. Differential weights are given corresponding to the number of different types of possible falsifiers of every character type.

Nucleotide Substitution Characters

All nucleotide substitution characters belong to the same character type. A nucleotide substitution event can have four different results - adenine, guanine, thymine or cytosine. As a consequence, for every hypothesis of a synapomorphic nucleotide substitution there are four types of possible character states according to the four different nucleotide types. Thus, *three different classes of identical falsifiers* exist, all of which would falsify a hypothesis of a synapomorphic nucleotide substitution of a specific nucleotide in the identity test.

Deletion Characters

The problem with deletion events – as with insertion events – is that, theoretically, they have no concrete upper limit according to their possible nucleotide length. As a consequence, for a hypothesis of a synapomorphic deletion of a given length one would get an almost infinite number of different classes of identical falsifiers. The number would be independent of the actual length of the hypothesized deletion. This does not seam plausible and therefore, as a convention, the use of an operational approach to interpreting the alignment which is derived from the inference of the topographical correspondence is proposed. Only those positions of the alignment that potentially represent the character states of a single cladistic character serve as the basis for the classification of their corresponding character type. They also set a limit to what could potentially serve as a falsifier of this hypothesis within the alignment. In practice, according to the proposed approach, the position of the ‘window’ in the alignment which is considered when evaluating the amount of potential falsifiers for the

corresponding hypothesis is set by the longest uninterrupted row of gaps. Hence, if a deletion of *n* nucleotides is hypothesized, only those *n* corresponding positions of the alignment serve as empirical evidence and the source for potential falsifiers.

Considering deletions, there are two possible states for every alignment position in question: presence or absence of the result of a nucleotide deletion. This means that every position with a gap is understood as indicating the presence, and every position that has a nucleotide as indicating the absence of the result of a deletion. Thus, for a given hypothesis of a synapomorphic deletion of *n* nucleotides, one gets 2^n possible patterns of different combinations of absence and presence that could potentially be observed in the alignment. Therefore, there are *2^n – 1 different classes of identical falsifiers of such a hypothesis*.
This is why such a hypothesis is not only falsified by all non-deletion sites but also by every deletion which is smaller in length than the hypothesized one.

Within the suggested system, sites of multiple gaps are understood as single character states that have to be hypothesized most parsimoniously as the result of a single event rather than multiple independent events (contradicting Giribet and Wheeler, 1999). Though every sequence position represents an observationally distinguishable unit, it does not necessarily represent an evolutionary independent unit of mutational processes, hence does not necessarily represent single character states.

Insertion Characters

The problem with deletions also holds true for insertions, and the same operational approach is applied. Because, in an insertion, nucleotides are inserted into an existing sequence there are 5 possible states for every position: the four possible types of inserted nucleotides and the absence of any nucleotide, a gap. As a consequence, for a given hypothesis of a synapomorphic insertion of *n* nucleotides, one gets 5^n possible patterns of different combinations of those 5 states that could potentially be observed in the alignment. And, as a consequence, one receives *5^n – 1 different classes of identical falsifiers of that hypothesis*.

When this classification of classes of identical falsifiers is applied and one wishes to weight the received classes equally, which would correspond to the conventional

interpretation of Popper's falsificationist approach for phylogenetic research, as it disregards process probabilities, one receives the following differential character weights:

A)	**nucleotide substitutions:**	***3***
B)	**deletions of *n* nucleotides:**	$2^n - 1$
C)	**insertions of *n* nucleotides:**	$5^n - 1$

The proposed weighting system is only applicable to insertions that exhibit a specific quality. Due to the conditions set by the identity criterion, only directly neighboring positions of an insertion are considered, recoded and weighted as an insertion character of a specific length, that is identical throughout all sequences which posses the insertion (for details see discussion).

Materials and Methods

Species examined

178 18S rDNA sequences were taken from NCBI/GenBank via the internet. For full species names and GenBank accession numbers for the sequences used in the alignments, in the spectral analyses and the parsimony analyses see in the appendix.

Alignment and Cladistic Analysis

Two data sets were analyzed. A *large data set* consisting of all 178 sequences was aligned and analyzed with spectral analysis, parsimony jackknifing and parsimony analysis. The results of these analyses were used for determining the taxon composition of the *small data set* consisting of 14 Diptera and 4 Strepsiptera sequences and a smaller sample of 48 closely related outgroup sequences. This small data set was aligned and analyzed in the same way as the large one. In some cases, outgroup comparison of the small data set facilitated a differentiation of indel events into insertion and deletion events within the Diptera and Strepsiptera ingroup. With a *subset* of the small alignment consisting of the 18S rDNA sequences of Diptera species together with 4 Strepsiptera and 2 Hymenoptera species only, another spectral, parsimony jackknifing and

parsimony analysis was performed. The hypothesized insertion and deletion events were mapped onto the parsimony jackknifing tree and their degree of consistency in relation to the nucleotide substitution data and the effect of the specific weights was assessed.

Multiple alignments of the two data sets were performed using ClustalW (Thompson *et al.*, 1994) and corrected by hand. The alignment of the large data set consists of 4,007 positions and the alignment of the smaller data set of 3,857 positions.

For the spectral analysis of splitsupporting positions, the parsimony jackknifing analysis and the parsimony analysis of the nucleotide substitution characters only, sequence areas that contained indels or that could not be aligned unambiguously were excluded before writing the data matrix. From the large alignment 2,447 positions were excluded, while 1,560 positions remain. From the small alignment 2,212 positions were excluded and 1,645 positions remain.

Spectral analysis of splitsupporting positions was performed with PHYSID (Wägele and Rödding, 1998a). The results are presented according to Wägele and Rödding (1998a, 1998b), allowing 15% of noisy positions in every row and column of ingroup and outgroup sequences respectively. Only those splits with the highest number of splitsupporting positions are shown in the Figure 1.

Parsimony jackknifing analysis and parsimony analysis were performed with PAUP*4.0 (Swofford, 1998). Parsimony jackknifing analysis of the large data set was performed with 500 replicates, a deletion percentage of 50% and a heuristic search option with nearest-neighbor interchange. Analyses of the small data set were performed with 1,000 replicates, keeping the other parameters; and the subset also with 1,000 replicates and with tree bisection-reconnection heuristic search option. The parsimony analysis of the subset was performed under branch and bound search settings.

Results

The large alignment consists of 799 parsimony-informative positions, the small alignment of 713 and the subset of 574 parsimony-informative positions.

The results of the spectral analyses of the data sets show patterns of splitsupport that can hardly be explained by what one would expect as a pattern resulting from random processes. One has to assume that the data contains – at least for some cladistic hypotheses – relevant phylogenetic information. The spectral analysis of the large data set assigns a high degree of support to three split-groupings in particular. Besides the

two choanoflagellate sequences with 15, the Culicoida, a subgroup of the Diptera, with 14, and the Diptera themselves with 7 splitsupporting positions get a high degree of support. As a consequence, they or the corresponding group of the splits are supported as monophyletic groups. All other splits have only 3 or less supporting positions. The same analysis of the small data set leads to similar results with an even stronger signal. Not only the Culicoida (45 supporting positions) and the Diptera (32) receive the highest support in this analysis, but also the Strepsiptera (11) and Tipuloida (7) are supported as split-groupings (Fig. 1).

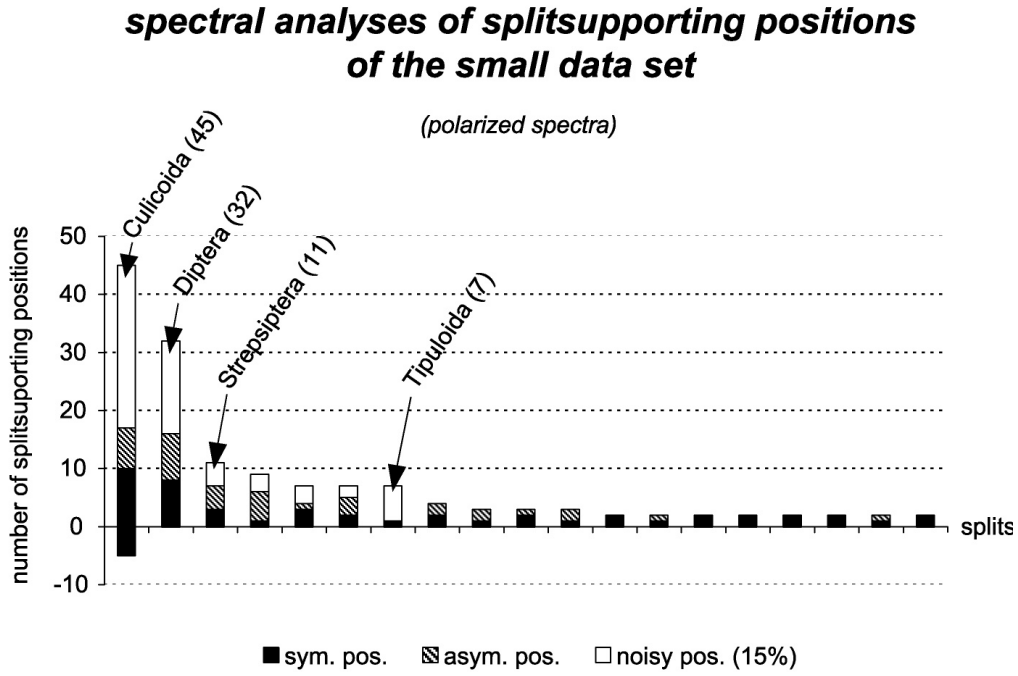

Fig. 1: Spectrum of conserved putative synapomorphies (splitsupporting positions) for groups of taxa for the small data set. The splits with the highest number of splitsupporting positions are represented out of over 2,800 splits of the complete spectrum.
Sym. pos.: symmetrical supporting positions; asym. pos.: asymmetrical supporting positions; noisy pos. (15%): splitsupporting positions with deviations in up to 15% of the sequences.

The parsimony jackknifing analyses of the three data sets assign jackknifing frequencies of 100.00 to the groups that are also highly supported by the results of the spectral analyses: Strepsiptera, Diptera, Culicoida and Tipuloida (Fig. 2 and 5).

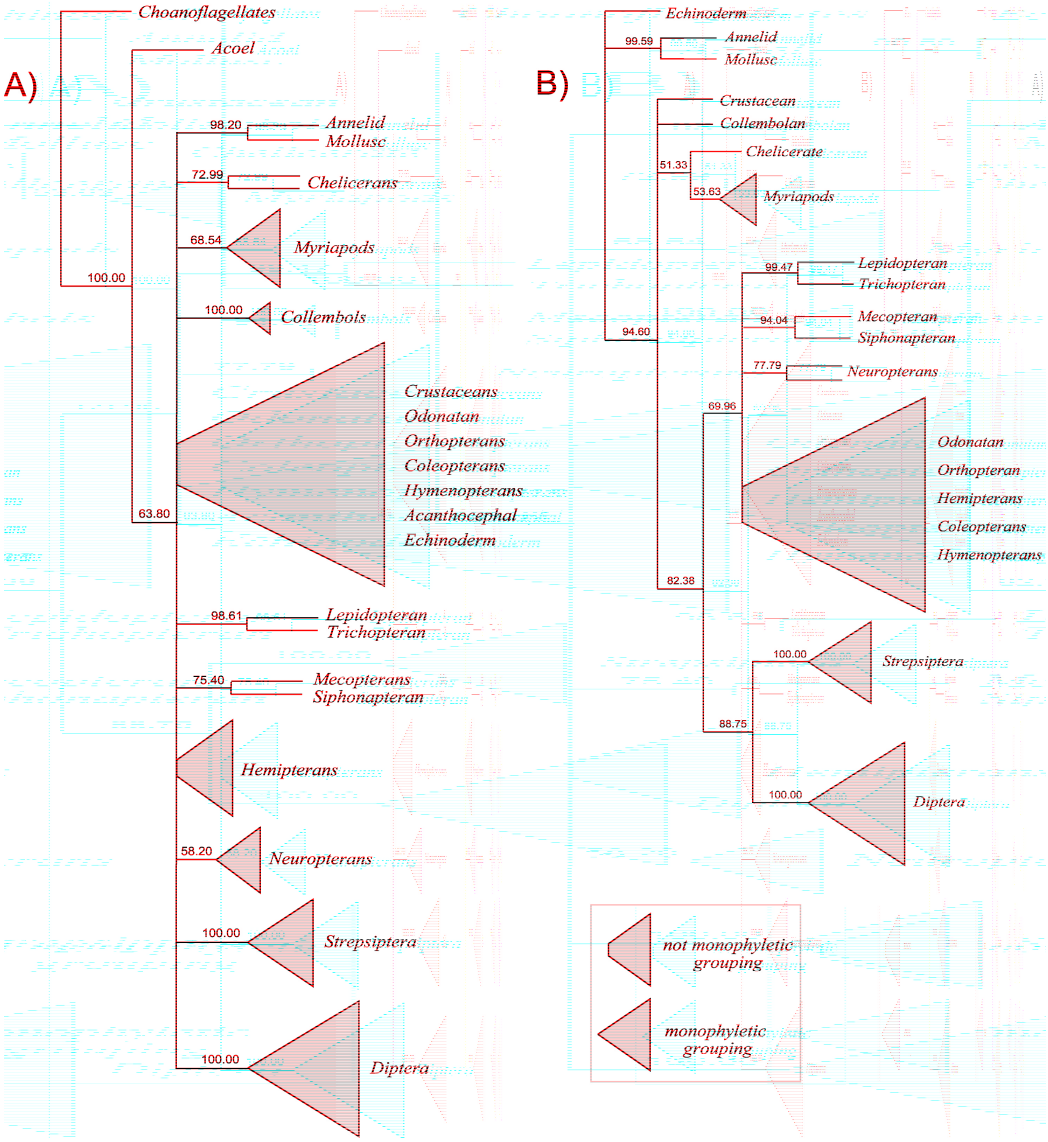

Fig. 2: A) 50% majority-rule consensus tree based on 500 maximum parsimony jackknifing replicates of the large 18S rDNA alignment using PAUP* 4.0. The jackknifing frequencies are labeled onto the corresponding branches. The tree was rooted *a posteriori* with the choanoflagellate sequence. B) 50% majority-rule consensus tree based on 1,000 maximum parsimony jackknifing replicates of the small 18S rDNA alignment using PAUP* 4.0. The jackknifing frequencies are labeled onto the corresponding branches. The tree was rooted *a posteriori* with the echinoderm sequence. For clarity of representation not all of the resolved nodes are shown.

In the following analysis, the Strepsiptera and the Diptera are hypothesized as monophyletic groups in relation to the other taxa of the data sets which represent their outgroup. This assumption is supported by the results of the spectral analyses and the parsimony jackknifing analyses.

Comparison of Indel Characters and Nucleotide Substitution Characters

Based on the assumption that Strepsiptera and Diptera represent groups of monophyletic origin it is possible to differentiate indel events, that took place within these clades, into insertions and deletions by applying an outgroup comparison. This comparison is

Table 1: Number of hypothesized indel events within the Diptera and Strepsiptera clade counted in the small alignment of 18S rDNA sequences

nucleotide length of the indels	inconsistent insertion characters	consistent insertion characters	inconsistent deletion characters	consistent deletion characters
1 nucleotide	5	25	9	17
2 nucleotides	-	5	3	3
3 nucleotides	-	-	-	2
4 nucleotides	-	2	-	-
5 nucleotides	-	2	-	-
15 nucleotides	-	1	-	-
Ó of all events	5	35	12	22

performed on the basis of the alignment of the small data set. All the other taxa of the alignment serve as the outgroup. As a result of this comparison a sum of 74 such indels were hypothesized - 40 insertions and 34 deletions. Insertions and deletions that could not be unambiguously hypothesized are not included in this statistic (Table 1).

From the subset a spectral analysis, a parsimony jackknifing analysis and a parsimony analysis were performed. The number of splitsupporting positions, the jackknifing indices and the number of putative apomorphies of the maximally parsimonious tree are drawn onto the resulting tree for every monophylum. Those insertions and deletions that are congruent with this tree are mapped onto it (Fig. 5). The number of congruent and incongruent insertions and deletions are summarized in Table 1 and illustrated in Figure 3.

23% of all hypothesized indels are incongruent; as are 35% of all deletions and 12.5% of all insertions. It is interesting to note that the observed rate of inconsistency decreases with the length of the hypothesized indel event.

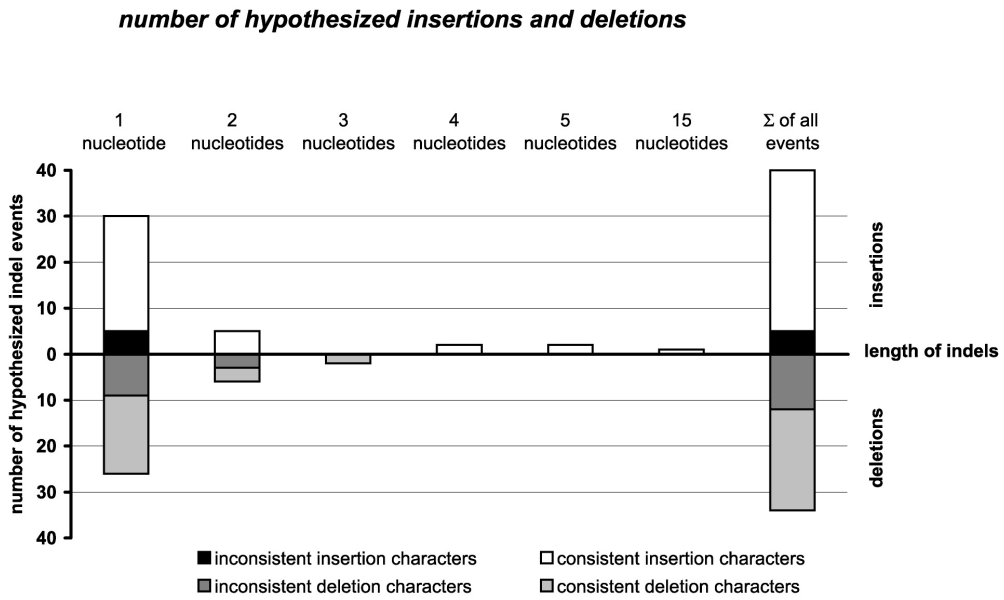

Fig. 3: Number of hypothesized insertion and deletion events. Insertions and deletions are differentiated with regard to their nucleotide length and to whether they are consistent with the most parsimonious tree obtained from the analysis of the nucleotide substitution data of the subset.

The ensemble consistency index (CI) for the sum of all the hypothesized insertion and deletion events was calculated on the basis of the tree shown in Figure 5. In addition, the CI was calculated for every character type and compared to the CI of only the nucleotide substitution data which was obtained by the parsimony analysis. Furthermore, the CIs for the sum of all indels, the sum of all insertions and the sum of all deletions was determined separately and compared to the corresponding modified CIs obtained by the application of the weighting system proposed in this paper (Table 2 and Fig. 4).

Table 2: Weighted and unweighted ensemble consistency indices (CI) of insertion, deletion and nucleotide substitution events in dependence of their nucleotide length. The CIs were calculated on the basis of a most parsimonious tree obtained from the subset and insertion and deletion events hypothesized by outgroup comparison on the basis of the small alignment of 18S rDNA sequences. The 15 nucleotide long insertion is not included in the calculation of the CI of the sum of all indels, since its weight with $3.05x10^{10}$ is extraordinarily high and would suppress the effects of all the other indels of this comparison.

Character state type	Nucleotide length	CI of unweighted data	CI of weighted data
Insertions	1	0.8571	-
	2	1	-
	4	1	-
	5	1	-
	15	1	-
Ó (without 15 nucleotide insertion)	-	0.8864	0.9974
Ó	-	0.8889	1
Deletions	1	0.7027	-
	2	0.6667	-
	3	1	-
Ó		0.7083	0.7436
Ó of all indels (without 15 nucleotide insertion)	-	0.7957	0.9949
Nucleotide substitutions	-	0.6305	-

All determined CIs of the different indel character types are higher than 0.65, the CI of the nucleotide substitution characters. The lowest CI of the indel characters is found with deletions, with a CI of 0.71, compared to a CI of 0.89 of the insertions. All insertions of a length higher than one nucleotide posses a CI of 1.00.

When applying the proposed weighting system, the calculated CIs of the now weighted indels are constantly higher than those of the unweighted indels. This holds true for the insertions, the deletions and all indels together.

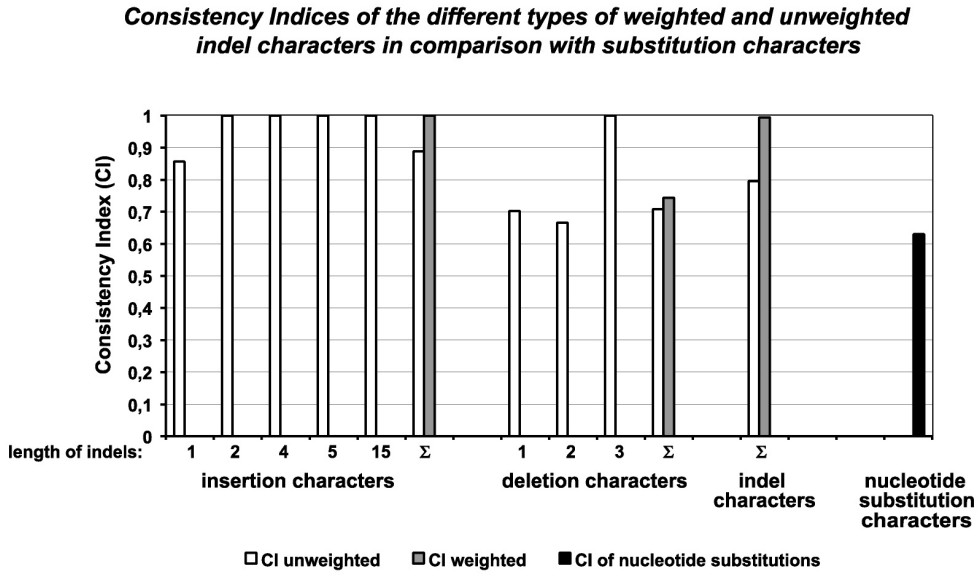

Fig. 4: Weighted and unweighted ensemble consistency indices (CI) of the different types of insertion and deletion characters. The CIs were calculated on the basis of the most parsimonious tree obtained from the analysis of the nucleotide substitution data of the subset. The CI of the nucleotide substitution characters was calculated by this parsimony analysis with PAUP* 4.0 using the branch and bound search option. For the calculation of the weighted CIs the weighing system proposed in this paper was applied.

The distribution of the congruent indel characters is clustered. Those groups that gain high support from the substitution data (high jackknifing frequencies of 100.00 and high splitsupports) also exhibit a high number of observed indels (Fig. 3).

Table 3: The distribution of the incongruent indels on contradicting cladistic hypotheses. Incongruent insertions and deletions and their nucleotide lengths are differentiated, and the minimum number of required steps are inferred for their distribution on the most parsimonious tree obtained from the analysis of the subset. The total weight of every observed set of incongruent character states that support a contradicting cladistic hypothesis are summed up. The incongruent insertion and deletion events were hypothesized by outgroup comparison on the basis of the alignment of the small data set.

Insertion of a single nucleotide	Deletion of a single nucleotide	Deletion of two nucleotides	Number of cladistic hypotheses supported that way	Minimum steps of the character states on the most parsimonious tree	Σ of the weights of the indel events
-	1	-	2	3	1
1	-	-	5	2	4
-	1	1	2	2	4
-	-	1	1	2	3
-	2	-	1	2	2
-	1	-	3	2	1

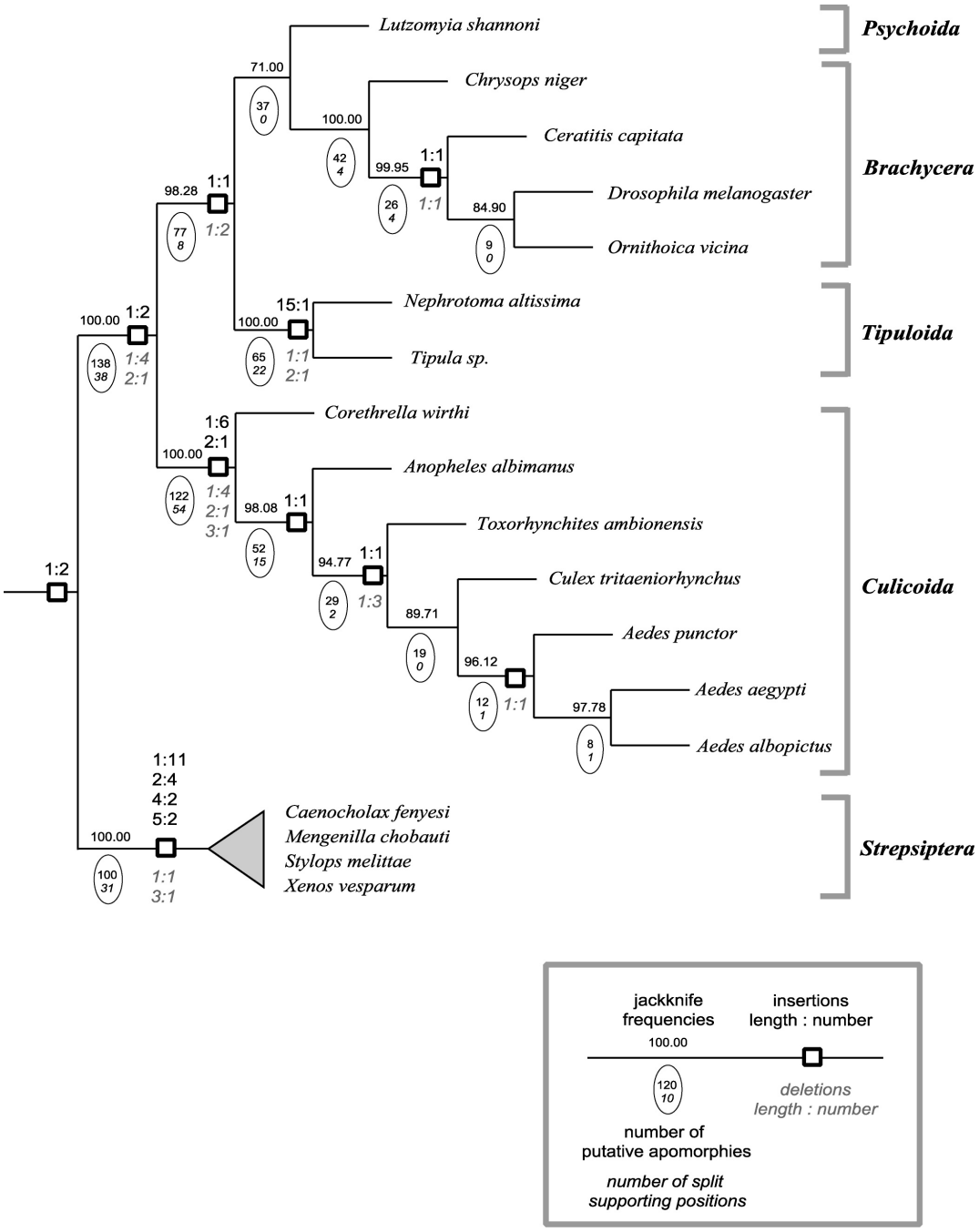

Fig. 5: 50% majority-rule consensus tree based on 1,000 maximum parsimony jackknifing replicates of the subset of the small 18S rDNA alignment using PAUP* 4.0. Jackknifing frequencies, the number of putative apomorphies which were calculated by the branch and bound search option and the number of splitsupporting positions, which were computed using PHYSID, are labeled onto the corresponding branches. The hypothesized insertion and deletion events that are congruent with this tree are also drawn onto the corresponding branches in dependence of their specific nucleotide length. The relationships within the Strepsiptera are not shown. The tree is rooted *a posteriori* with Hymenoptera sequences.

The cladistic distribution of the incongruent indels in comparison to the congruent ones is quite homogeneous (Table 3). There are 14 different contradicting cladistic hypotheses that are supported by the 17 incongruent indels. None of those 14 hypotheses is supported by more than two indels. And none of those contradicting hypotheses is supported by a higher sum of weights than 4.

All the different parameters that were calculated on the basis of this data set correlate very well with each other and show a clear pattern of strongly and less strongly supported cladistic relationships. This also applies to the indel characters. The distribution of the consistent indels in comparison to the inconsistent indels can hardly be explained by the result one would expect from a plain random process.

Discussion

Since testing the utility of indel characters is the aim of this paper, distance and maximum likelihood methods of cladistic analysis were not applied. The underlying evolutionary processes are too diverse and have, so far, not been well enough understood to be incorporated in a stochastical model of evolution. This and the problems of assessing hypotheses of topographical correspondence in highly variable regions of the alignment that have a substantial number of alignment gaps is also stated as a general argument against the use of insertions and deletions as informative phylogenetic characters in cladistic analyses (Swofford *et al.*, 1996). In spite of this argument, Van Dijk *et al.* (1999) use deletions by modifying the maximum likelihood procedure of Kishino *et al.* (1990) by not allowing the occurrence of reversals (back mutations) of deletion events. Moreover, Mitchison (1999) uses insertions and deletions in a probabilistic approach of combining alignment and cladistic analysis by means of sampling.

However, if one wants to test the proposed weighting scheme which rests on a refutationist approach to phylogeny, these two types of cladistic methods - distance and Maximum Likelihood - seam to be problematic. This speaks in favor of choosing maximum parsimony as the method of cladistic analysis in this paper. Maximum parsimony allows the combined use of all kinds of different data (insertions and deletions, and even morphological and other data) and is in principle consistent with a refutationist approach.

Some papers present results employing indels as cladistic characters in parsimony analyses. In addition to the substitution data, the indels are coded as absent-present characters and are analyzed in an equally weighted parsimony analysis (e.g., Baum *et al.*, 1998; and for a more sophisticated coding see also Barriel, 1994; and Simmons and Ochoterena, 2000). Giving all types of characters equal weights, these procedures have the implicit assumption that the phylogenetic information content of insertion, deletion and nucleotide substitution characters is the same without giving any empirical or methodological reasons.

Lloyd and Calder (1991) apply the same coding and discuss the process probabilities of indel events and claim to be able to evaluate the reliability of such character types in respect to their length, position and frequency. They also receive remarkably high consistency indices for the indel characters they utilized in their study.

Gatesy *et al.* (1993), Wheeler (1995) and Giribet and Wheeler (1999) use insertion and deletion characters in parsimony analyses and give them weights proportional to the costs assigned during their alignment. This procedure is based on the application of different models of sequence evolution in the alignment. Lutzoni *et al.* (2000) presented an interesting procedure of recoding and weighting of gaps. Here, unambiguously aligned sites are weighted by a step matrix which is calculated from relative frequencies of each possible transformation, and ambiguously aligned sites undergo a sophisticated method of recoding and "optimal weighting" resulting in a single character for each ambiguous region with its own step matrix. But they give insertions of a specific length the same weight as deletions of the same length.

Another procedure that has been suggested is to code the gaps as fifth character state (e.g., Swofford, 1998; Titus and Frost, 1996). However, since it is more parsimonious to hypothesize that one indel event with more than one nucleotide has taken place rather than several such events with only a single nucleotide independently, this fifth-character-state-coding neglects the dependence of those gap positions that are direct neighbors in an alignment. This procedure would therefore result in an artificial weighting of gaps relative to the number of sites (Barriel, 1994; Simmons and Ochoterena, 2000). Furthermore, insertion and deletion events are not differentiated.

Thus, all these analyses either apply weighting schemes that represent methodological proposals that are not consistent with a falsificationist approach, neglect the evolutionary dependence of directly neighboring gap positions within an alignment or

ignore the different information content of indels by not discriminating insertions and deletions.

One problem that complicates the application of indels as phylogenetic markers concerns insertion characters in particular:

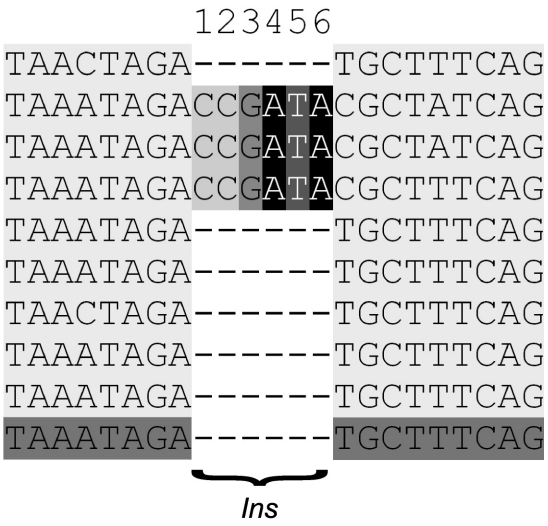

Fig. 6: Example of an identical insertion elements of 6 nucleotides present in three of ten sequences. Darker sequence represents the outgroup sequence. *Ins* = Insertion

As long as insertions are all together identical in their nucleotide sequence it is unproblematic to state a hypothesis of synapomorphy in the light of the assumed background knowledge. Figure 6 shows an example of such a clearly recognizable, all together identical insertion of the length of 6 nucleotides in three of ten sequences. Given this number of nucleotides there are potentially 15,624 different classes of identical falsifiers. Hence, if one would like to weight this cladistic character (*insertion of six nucleotides present or absent*) it would get the weight of 15,624 in comparison to 3 for any nucleotide substitution character.

However, when the nucleotide sequence of the insertion is not identical within the taxa under consideration, the hypothesis of a synapomorphic insertion is, following the identity criterion (Vogt, *Testing and weighting characters*), already falsified (Fig. 7). The perceivable empirical evidence does not represent a *single* insertion event only. In some cases one might still think of combined events of an insertion followed by some independent substitutions as an alternative hypothesis (Fig. 7: *A* and *B*). But this is not always possible (Fig. 7: *C*) and thus confronts one with problems of continuity. Furthermore, the empirical evidence would be explained by *ad hoc* hypotheses of subsequent events, which probably followed the insertion and thus might have caused the data. However, the more *ad hoc* hypotheses are endeavored the weaker is the explanatory power of the stated hypothesis, and thus the smaller is the information content of the corresponding phylogenetic character.

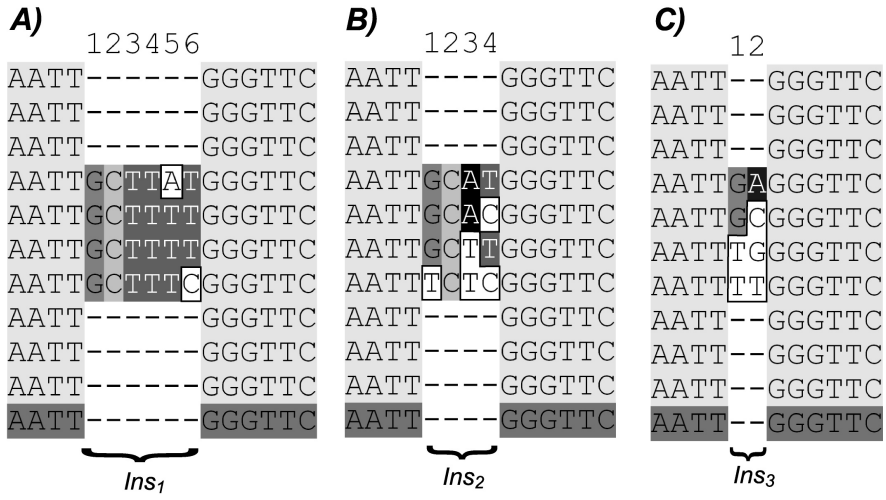

Fig. 7: Three examples of non-identical insertions. Darker sequences represent outgroup sequences. ***A)*** For this alignment one could hypothesize an insertion event, *Ins*$_1$, followed by two independent nucleotide substitutions. ***B)*** Here one could hypothesize an insertion event, *Ins*$_2$, followed by subsequent independent nucleotide substitutions. The nucleotide sequence of the original insertion is not unambiguously reconstructable. ***C)*** What actually happened here is not unequivocally interpretable and hypotheses of specific mutation events gain only weak empirical support. *Ins*$_1$ = Insertion of 6 nucleotides; *Ins*$_2$ = Insertion of 4 nucleotides; *Ins*$_3$ = Insertion(s) of 2 nucleotides

At the moment there is no theoretical foundation to weight these non-identical insertions completely. That is why only the identical neighboring positions within the insertions are appropriate for the proposed weighting system.

One is still confronted by an irresolvable problem: a large insertion could subsequently be followed by a large deletion which deletes all the previously inserted nucleotides or even more. If this deletion occurs in one line of descent after speciation events have taken place, the here proposed approach unfortunately would not be immune against an artificial coding of the results of such a combined event. This problem of possible reversals, however, is not limited to insertion and deletion characters only but affects almost any cladistic character.

The severity of the described test depends on the quality of the data set. Especially the available 18S rDNA sequences of Diptera and Strepsiptera species seam to provide a proper basis for such a test, since they exhibit an extensive amount of gaps when aligned with other insects and metazoa taxa. This phenomenon can be explained by assuming that many insertion and deletion events took place within these clades. Furthermore, their nucleotide substitution data also seams to provide a comparably high

informative data set as far as the spectral and the parsimony-jackknifing analyses are concerned.
The hypothesis of a monophyletic position of the Diptera and of the Strepsiptera in relation to all other Metazoa taxa of the data set is also highly corroborated by the nucleotide substitution data, as shown by the spectral and parsimony-jackknife analyses. All this enables one to choose an adequate sample of outgroup taxa for an outgroup comparison to differentiate the indel events within the Strepsiptera and within the Diptera into insertions and deletions; and the distribution of those events in relation to the result of the analysis of the nucleotide substitution data allows one to compare the patterns of cladistic distribution of the different types of characters.

The result of this test does not contradict the hypotheses that are tested. The cladistic distribution of the insertions and deletions reflect the conditions inferred by the nucleotide substitution data, since most of the observed indels are congruent with the most parsimonious tree (Fig. 5). Those insertions and deletions that are incongruent with this tree exhibit a rather coincidental distribution. Thus, there is no contradicting cladistic hypothesis that is supported by more than two indel events. In comparison, the consistent indels show a clear hierarchical distribution and some of the cladistic hypotheses are supported by 5 or even more hypothesized indel events. These findings are also supported by the obtained ensemble consistency indices (CI) of the different types of insertions and deletions. The CI is taken as the basis of comparison, as it counts the requirements for *ad hoc* hypotheses of homoplasy. The higher the CI the fewer *ad hoc* hypotheses are required to explain the data (Kluge, 1997). All of the CIs of the indels exhibit a higher value than the CI of the nucleotide substitution data. This is in accordance with the results of the inference of Lloyd and Calder (1991). It, thus, seams as if some insertions and deletions represent phylogenetically highly informative character types. Therefore, they seam to be comparably good phylogenetic markers.

Furthermore, the suggested weighting system withstood the test. None of the 14 contradicting cladistic hypotheses that are supported by indels are corroborated by a higher weight than 4. Of the most parsimonious clades in Figure 5 there are 8 cladistic hypotheses with a sum of weights each exceeding the value of 4 by far.
The sum of weights of all consistent indel character states is strikingly higher than the sum of weights of all the inconsistent indel character states (Table 4). The ratio of

consistent to inconsistent indel characters changes from 57:17 (22.97% inconsistency rate) when weighted equally to 3,878:19 (0.005% inconsistency rate) when weighted according to the proposed weighting system.
The application of the weighting system also leads to higher CIs. Looking at the composition of the inconsistent indels it is obvious that none of the insertion and deletion character state types that receive a higher weight than 4 are represented. This also supports the suggested weighting system.

So far, the proposed *a priori* differential weighting system for nucleotide sequence data of non protein coding genes all passed the empirical test. However, it still needs further testing on suitable empirical data, since it should always be the aim of a scientist to try to falsify rather than to try to verify a hypothesis (Popper, 1983, 1994); and this hypothesis withstood only a first test. Especially the concrete quantification which leads, in the case of long insertions, to tremendously high weights, is open to critique and resembles a methodological proposal. This proposal rests on the conventional interpretation of Popper's falsificationsim in phylogenetic inference, which disregards the necessity of taking different process probabilities for different types of transformations into consideration when analyzing the data. Whether this is the only possible and right interpretation of falsificationism is yet open to discussion (de Queiroz and Poe, 2001; Faith and Trueman, 2001; Kluge, 2001). By following this proposal, one does not consider the actual pattern of the nucleotides of given insertions of a specific length because one does not consider their process probabilities. Especially when comparing two insertions with e.g. AGGCCCGCGATAGT and AAAAAAAAAAAATA it seems counterintuitive to weight them equally since we know that AT rich insertions evolve more frequently then other insertions. If one does not want to follow the conventional interpretation of Popper's falsificationism and wants to include assumptions of process probabilities in the relevant background knowledge, the weighting of every class of potential falsifiers should rest on the process improbability of the corresponding transformation type.
Anyway, whether taking process probabilities into account or not, it still seams reasonable to record that the phylogenetic information content of indels tends to increase with their length and decrease from insertions to deletions and to nucleotide substitutions. Thus, the application of such relative weights should be considered,

especially in cases where the nucleotide substitution data alone does not give a strong support for any cladistic hypothesis. Also in terms of a total evidence approach all available empirical data should be considered (Kluge and Wolf, 1993) and a maximally corroborated cladistic hypothesis inferred. If alignment gaps are the result of a particular indel mutation event, then they inevitably bear phylogenetic information. Ignoring this type of cladistic character would equal ignoring empirical evidence and this could lead to cladistic hypotheses which are not maximally corroborated and therefore less explanatory (Giribet and Wheeler, 1999).

ACKNOWLEDGMENTS

I want to thank A. G. Kluge, K. M. Kjer and F. Lutzoni who read an earlier draft of the manuscript offering many valuable comments and M. Hendy and A. Dress for discussing some of my conclusions. I am particularly indebted to T. Bartolomaeus, P. Ax and W. Ahlrichs for discussing and criticizing my views of the concept of homology and the nature of the congruence test extensively. My conclusions concerning *a priori* weighting and the value of indels as phylogenetic markers were first presented at the XVIIIth meeting of the Willi Hennig Society, 1999, in Göttingen.

A special thanks goes to D. Bhattacharya who introduced me to the methods of molecular phylogenetics some years ago.

This study was supported by the Deutsche Forschungsgemeinschaft (BA 1520/4-1).

References

Aguinaldo, A. M. A., Turbeville, J. M., Linford, L. S., Rivera, M. C., Garey, J. R., Raff, R. A., and Lake, J. A. (1997): Evidence for a clade of nematodes, arthropods and other moulting animals. Nature 387: 489-493.

Aleshin, V. V., Milyutina, I. A., Kedrova, O. S., Vladychenskaya, N. S., and Petrov, N. B. (1998): Phylogeny of Nematoda and Cephalorhyncha derived from 18S rDNA. J. Mol. Evol. 47: 597-605.

Barriel, V. (1994) : Phylogénies moléculaires et insertion-délétions de nucléotides. C. R. Acad. Sci. Paris, Sciences de la vie/Life sciences 317 : 693-701.

Baum, D. A., Small, R. L., and Wendel, J. F. (1998): Biogeography and Floral Evolution of Baobabs (Adansonia, Bombacaceae) as Inferred from Multiple Data Sets. Syst. Biol. 47: 181-207.

Canapa, A., Barucca, M., Marinelli, A., and Olmo, E. (2000): Molecular data from the 16S rRNA gene for the phylogeny of Pectinidae (Molusca: Bivalvia). J. Mol. Evol. 50: 93-97.

De Queiroz, K. and Poe, S. (2001) : Philosophy and phylogenetic inference : A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. Syst. Biol. 50: 305-321.

Faith, D. P. and Trueman, J. W. H. (2001): Towards an inclusive philosophy for phylogenetic inference. Syst. Biol. 50: 331-350.

Gatesy, J., DeSalle, R., and Wheeler, W. (1993): Alignment-Ambiguous Nucleotide Sites and the Exclusion of Systematic Data. Mol. Phylogenet. Evol. 2: 152-157.

Giribet, G., and Wheeler, W. C. (1999): On Gaps. Mol. Phylogenet. Evol. 13: 132-143.

Hixson, J. E., and Brown, W. M. (1986): A comparison of the small ribosomal RNA genes from the mitochondrial DNA of the great apes and humans: sequence, structure, evolution, and phylogenetic implications. Mol. Biol. Evol. 3: 1-18.

Kishino, H., Miyata, T., and Hasegawa, M. (1990): Maximum Likelihood Inference of Protein Phylogeny and the Origin of Chloroplasts. J. Mol. Evol. 31: 151-160.

Kluge, A. G., and Wolf, A. J. (1993): Cladistics: What's in a world? Cladistics 9: 183-199.

Kluge, A. G. (1997): Testability and the Refutation and Corroboration of Cladistic Hypotheses. Cladistics 13: 81-96.

Kluge, A. G. (2001): Philosophical conjectures and their refutation. Syst. Biol. 50: 322-330.

Lloyd, D. G., and Calder, V. L. (1991): Multi-residue gaps, a class of molecular characters with exceptional reliability for phylogenetic analyses. J. Evolution. Biol.4: 9-21.

Lutzoni, F., Wagner, P., Reeb, V., and Zoller, S. (2000): Integrating ambiguously aligned regions of DNA sequences in phylogenetic analyses without violating positional homology. Syst. Biol. 49: 628-651.

Meyer, T. E., Cusanovich, M. A., and Kamen, M. D. (1986): Evidence against use of bacterial amino acid sequence data for construction of all-inclusive phylogenetic trees. P. Natl. Acad. Sci. USA83: 217-220.

Mitchison, G. J. (1999): A Probabilistic Treatment of Phylogeny and Secuence Alignment. J. Mol. Evol. 49: 11-22.

Patterson, C. (1982): Morphological characters and homology. *In* "Problems of phylogenetic reconstruction" (K. A. Joysey and A. E. Friday, Eds.), pp. 21-74. Academic Press, London.

Popper, K. R. (1983): Realism and the Aim of Science; From the Postscript to the Logic of Scientific Discovery. London and New York: Routledge,.

Popper, K. R. (1994): Logik der Forschung. (10., berbesserte und vermehrte Auflage von "Die Einheit der Gesellschaftswissenschaften; Band 4). Tübingen: J. C. B. Mohr (Paul Siebeck).

Rieppel, O. C. (1988): Fundamentals of Comparative Biology. Birkhäuser Verlag, Basel.

Simmons, M. P. and Ochoterena, H. (2000): Gaps as characters in sequence-based phylogenetic analyses. Syst. Biol. 49: 369-381.

Swofford, D. L. (1998): *PAUP4.0*: phylogenetic analysis using parsimony and other methods (beta version). Sinauer, Sunderland, Massachusetts.

Swofford, D. L., Olsen, G. J., Waddel, P. J., and Hillis, D. M. (1996): Phylogenetic inference. *In* "Molecular Systematics" (D. M. Hillis, C. Moritz and B. K. Mable, Eds.) Second ed., pp. 407-514. Sinauer, Sunderland, Massachusetts.

Thompson, J. D., Higgins, D. G., Gibson, T. J. (1994): CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22: 4673-4680.

Titus, T. A., and Frost, D. R. (1996): Molecular Homology Assessment and Phylogeny in the Lizard Family Opluridae (Squamata: Iguania). Mol. Phylogenet. Evol. 6: 49-62.

Van Dijk, M. A. M., Paradis, E., Catzeflis, F., and De Jong, W. W. (1999): The Virtue of Gaps: Xenarthran (Edentate) Monophyly Supported by a Unique Deletion in á A-Crystallin. Syst. Biol. 48: 94-106.

Vogt, L. (*Testing and weighting characters*) – handed in together with this manuscript.

Wägele, J. W. (1996): First principles of phylogenetic systematics, a basis for numerical methods used for morphological and molecular characters. Vie Milieu46 : 125-138.

Wägele, J. W., and Rödding, F. (1998a): Origin and Phylogeny of Metazoans as Reconstructed with rDNA Sequences. Prog. Molec. Subcell. Biol. 21: 45-70.

Wägele, J. W., and Rödding, F. (1998b): A Priori Estimation of Phylogenetic Information Conserved in Aligned Sequences. Mol. Phylogenet. Evol. 9: 358-365.

Wheeler, W. C. (1995): Sequence Alignment, Parameter Sensitivity, and the Phylogenetic Analysis of Molecular Data. Syst. Biol. 44: 321-331.

Williams, S. A., and Goodman, M. (1989): A statistical test that supports a human/chimpanzee clade based on noncoding DNA sequence data. Mol. Biol. Evol. 6: 325-330.

Zrzavý, J., Mihulka, S., Kepka, P., and Bezdek, A. (1998): Phylogeny of the Metazoa based on morphological and 18S ribosomal evidence. Cladistics 14: 249-285.

Appendix

The alignment and data matrix in which the sequences were included is assigned by a '+' in the corresponding columns:

Species name	GenBank accession number	Token	Large alignment	Small alignment	Subset of the small alignment
1. Choanoflagellata					
Diaphanoeca grandis	L10824	cho1	+	-	-
Ichthyophonus hoferi	U25637	cho2	+	-	-
2. Metazoa					
2. 1 Acanthocephala					
Corynosoma enhydri	AF001837	Acanthoce	+	-	-
2. 2 Acoela					
Paratomella rubra	AF102892	Acoel	+	-	-
2. 3 Annelida					
Stylaria sp.	U95946	Annelid	+	+	-
2. 4 Chelicerata					
Carcinoscorpius rotundicaudatus	U91491	Chel1	+	-	-
Stylocellus sp.	U91485	Chel2	+	+	-
2. 5 Crustacea					
Derocheilocaris typicus	L81937	Crus1	+	-	-
Rutiderma sp.	L81942	Crus2	+	+	-
2. 6 Echinodermata					
Amphipholis squamata	X97156	Echinoderm	+	+	-
2. 7 Insecta					
Coleoptera					
Amblytelus curtus	AF012484	1Coleop	+	-	-
Aptinus displosor	AF012480	2Coleop	+	+	-
Blethisa multipunctata aurata	AF002803	3Coleop	+	+	-
Ceroglossus chilensis	AF012509	4Coleop	+	+	-
Copelatus chevrolati renovatus	AF012524	5Coleop	+	-	-
Dyschirius sphaericollis	AF002798	6Coleop	+	+	-
Mecodema fulgidum	AF012501	7Coleop	+	-	-
Monolobus ovalipennis	AF012505	8Coleop	+	-	-
Morion aridus	AF002783	9Coleop	+	+	-
Notiophilus semiopacus	AF002804	10Coleop	+	+	-
Patrobus longicornis	AF002786	11Coleop	+	-	-
Scarites subterraneus	AF002795	12Coleop	+	-	-
Suphis inflatus	AF012523	13Coleop	+	-	-
Tenebrio molitor	X07801	14Coleop	+	-	-
Trechus sp.	AF002793	15Coleop	+	-	-
Collembola					
Crossodonthina koreana	Z36893	1Collembol	+	-	-
Hypogastrura dolsana	Z26765	2Collembol	+	-	-
Podura aquatica	AF005452	3Collembol	+	+	-
Diptera					
Brachycera					
Ceratitis capitata	AF096450	1Bra	+	+	+
Chrysops niger	AF073889	2Bra	+	+	+
Drosophila melanogaster	M21017	3Bra	+	+	+
Ornithoica vicina	AF073888	4Bra	+	+	+
Culicoida					

Aedes aegypti	U65375	1Cul	+	+	+
Aedes albopictus	X57172	2Cul	+	+	+
Aedes punctor	U48378	3Cul	+	+	+
Anopheles albimanus	L78065	4Cul	+	+	+
Corethrella wirthi	U49736	5Cul	+	+	+
Culex tritaeniorhynchus	U48385	6Cul	+	+	+
Toxorhynchites ambionensis	U48377	7Cul	+	+	+
Psychoida					
Lutzomyia shannoni	U48382	1Psycho	+	+	+
Tipuloida					
Nephrotoma altissima	U48379	1Tipu	+	+	+
Tipula sp.	X89496	2Tipu	+	+	+
Hemiptera					
Acyrthosiphon pisum	U27819	1Hemip	+	-	-
Aonidiella aurantii	U06475	2Hemip	+	+	-
Lygus hesperus	U06476	3Hemip	+	+	-
Mindarus kinseyi	U27821	4Hemip	+	-	-
Okanagana utahensis	U06478	5Hemip	+	-	-
Pealius kelloggii	U06479	6Hemip	+	-	-
Philaenus spumarius	U06480	7Hemip	+	-	-
Prokelisia marginata	U09207	8Hemip	+	+	-
Rhaphigaster nebulosa	X89495	9Hemip	+	-	-
Schizaphis graminum	U27826	10Hemip	+	-	-
Spissistilus festinus	U06477	11Hemip	+	-	-
Trioza eugeniae	U06482	12Hemip	+	+	-
Hymenoptera					
Agonum extensicolle	AF002775	1H	+	-	-
Amara apricaria	AF002774	2H	+	+	-
Amarotypus edwardsi	AF012506	3H	+	-	-
Amblytelus curtus	AF012484	4H	+	-	-
Antarctonomus complanatus	AF012504	5H	+	+	-
Apotomus rufithorax	AF012497	6H	+	-	-
Aptinus displosor	AF012480	7H	+	-	-
Arthropterus sp.	AF012516	8H	+	-	-
Asaphidion curtum	AF002792	9H	+	-	-
Batesiana hilaris	AF012489	10H	+	-	-
Bembidion levettei	AF002791	11H	+	+	-
Bembidion mexicanum	AF012490	12H	+	-	-
Blethisa multipunctata aurata	AF002803	13H	+	+	-
Brachinus armiger	AF012479	14H	+	-	-
Brachinus hirsutus	AF012478	15H	+	-	-
Broscosoma relictum	AF012502	16H	+	-	-
Calosoma scrutator	AF002800	17H	+	+	-
Calybe laetula	AF002772	18H	+	-	-
Carabus nemoralis	AF012507	19H	+	-	-
Carenum interruptum	AF012491	20H	+	+	+
Catapiesis brasiliensis	AF012476	21H	+	-	-
Ceroglossus chilensis	AF012509	22H	+	+	-
Chlaenius ruficauda	AF012473	23H	+	-	-
Cicindela sedecimpunctata	AF012518	24H	+	-	-
Clambus arnetti	AF012526	25H	+	-	-
Clinidium calcaratum	AF012521	26H	+	-	-
Clivina ferrea	AF002796	27H	+	-	-
Cnemalobus sulciferus	AF012474	28H	+	-	-
Copelatus chevrolati renovatus	AF012524	29H	+	-	-
Creobius eydouxi	AF012498	30H	+	-	-
Cychrus italicus	AF012510	31H	+	-	-
Cymbionotum pictulum	AF012496	32H	+	+	+
Cymbionotum semelederi	AF012495	33H	+	-	-

Cymindis punctigera	AF002773	34H	+	-	-
Diplochaetus planatus	AF002789	35H	+	-	-
Diplous californicus	AF002785	36H	+	-	-
Discoderus cordicollis	AF012472	37H	+	+	-
Dynastes granti	AF002809	38H	+	-	-
Dyschirius sphaericollis	AF002798	39H	+	+	-
Elaphrus californicus	AF012514	40H	+	-	-
Elaphrus clairvillei	AF002802	41H	+	-	-
Galerita lecontei lecontei	AF002780	42H	+	-	-
Gehringia olympica	AF012512	43H	+	-	-
Hydroscapha natans	AF012525	44H	+	-	-
Laccocenus ambiguus	AF012486	45H	+	+	-
Leistus ferruginosus	AF002806	46H	+	+	-
Leptothorax acervorum	X89492	47H	+	-	-
Loricera foveata	AF012503	48H	+	-	-
Loricera pilicornis pilicornis	AF002799	49H	+	-	-
Loxandrus n. sp. nr. amplithorax	AF002778	50H	+	-	-
Mecodema fulgidum	AF012501	51H	+	-	-
Mecyclothorax vulcans	AF012482	52H	+	-	-
Melisodera picipennis	AF012481	53H	+	-	-
Merizodus angusticollis	AF012487	54H	+	+	-
Metius sp.	AF012475	55H	+	+	-
Metrius contractus	AF012515	56H	+	-	-
Monolobus ovalipennis	AF012505	57H	+	-	-
Nebria hudsonica	AF002805	58H	+	-	-
*Notiophilus semiopa*cus	AF002804	59H	+	+	-
Omoglymmius hamatus	AF012520	60H	+	-	-
Omophron obliteratum	AF012513	61H	+	-	-
Omus californicus	AF012519	62H	+	-	-
Oopterus sp.	AF012488	63H	+	-	-
Opisthius richardsoni	AF012511	64H	+	-	-
Oregus aereus	AF012500	65H	+	-	-
Pachyteles striola	AF012517	66H	+	-	-
Pamborus guerinii	AF012508	67H	+	-	-
Pasimachus atronitens	AF002794	68H	+	-	-
Patrobus longicornis	AF002786	69H	+	-	-
Pericompsus laetulus	AF002790	70H	+	-	-
Pheropsophus aequinoctialis	AF012477	71H	+	-	-
Polistes dominulus	X77785	72H	+	-	-
Promecoderus sp.	AF012499	73H	+	-	-
Promecognathus crassus	AF012492	74H	+	-	-
Pseudaptinus rufulus	AF002781	75H	+	-	-
Pseudomorpha sp.	AF002782	76H	+	+	-
Psydrus piceus	AF002784	77H	+	-	-
Pterostichus melanarius	AF002779	78H	+	-	-
Scaphinotus petersi catalinae	AF002801	79H	+	-	-
Scarites subterraneus	AF002795	80H	+	-	-
Schizogenius falli	AF002797	81H	+	-	-
Siagona europaea	AF012493	82H	+	-	-
Siagona jennisoni	AF012494	83H	+	-	-
Sloaneana tasmaniae	AF002788	84H	+	-	-
Suphis inflatus	AF012523	85H	+	-	-
Systolosoma lateritium	AF012522	86H	+	-	-
Tetragonoderus latipennis	AF012471	87H	+	-	-
Trachypachus gibbsii	AF002808	88H	+	+	-
Trachypachus holmbergi	AF002807	89H	+	-	-
Tropopterus sp.	AF012483	90H	+	-	-
Xanthopyga cacti	AF002810	91H	+	-	-
Zolus helmsi	AF002787	92H	+	-	-

Lepidoptera					
Galleria mellonella	X89491	1Lepidop	+	+	-
Mecoptera					
Boreus sp.	X89487	1Mecopt	+	-	-
Panorpa germanica	X89493	2Mecopt	+	+	-
Neuropteroida					
Anisochrysa carnea	X89482	1Neurop	+	-	-
Oliarces clara	AF012527	2Neurop	+	-	-
Phaeostigma notata	X89494	3Neurop	+	+	-
Sialis sp.	X89497	4Neurop	+	+	-
Odonata					
Aeschna cyanea	X89481	1Odona	+	+	-
Orthoptera					
Carausius morosus	X89488	1Orthop	+	+	-
Forficula sp.	X89490	2Orthop	+	-	-
Siphonaptera					
Archaeopsylla erinacea	X89486	1Siphonap	+	+	-
Strepsiptera					
Caenocholax fenyesi	U65160	1Strep	+	+	+
Crawfordia sp.	U65163	2Strep	+	-	-
Elenchus japonica	U65162	3Strep	+	-	-
Mengenilla chobauti	X89441	4Strep	+	+	+
Stylops melittae	X89440	5Strep	+	+	+
Triozocera mexicana	U65159	6Strep	+	-	-
Xenos pecki	U65164	7Strep	+	-	-
Xenos vesparum	X77784	8Strep	+	-	-
Xenos vesparum	X74763	8aStrep	+	+	+
Trichoptera					
Hydropsyche sp.	X89483	1Trichop	+	+	-
2. 8 Mollusca					
Atrina pectinata	X90961	Mollusca	+	+	-
2. 9 Myriapoda					
Clinopodes poseidonis	AF000777	Myr1	+	-	-
Craterostigmus tasmanianus	AF000774	Myr2	+	+	-
Cryptops trisulcatus	AF000775	Myr3	+	-	-
Cylindroiulus punctatus	AF005448	Myr4	+	-	-
Lithobius variegatus	AF000773	Myr5	+	-	-
Polydesmus coriaceus	AF005449	Myr6	+	+	-
Pseudohimantarium mediterraneum	AF000778	Myr7	+	-	-
Scolopendra cingulata	U29493	Myr8	+	+	-
Scutigera coleoptrata	AF000772	Myr9	+	-	-
Theatops erythrocephala	AF000776	Myr10	+	-	-

Table of Hypothesized Indels

Table of putative indel events that might have taken place within the dipteran and strepsipteran lineage. These events are hypothesized after the application of an outgroup comparison on the basis of the small alignment. The table contains the position of each hypothesized insertion and deletion within the small alignment, its specific nucleotide length and the phylogenetic grouping (clade) that it supports. Additionally it is indicated wether the supported phylogenetic group is congruent with the cladogram received from the maximum parsimony analysis of the nucleotide substitution characters of the subset.

insertion with corres-ponding nucleotide length	deletion with corres-ponding nucleotide length	congruent with the cladogram from the analysis of subset	alignment position	supported grouping
1	-	Yes	80	Strep
5	-	Yes	218-222	Strep
4	-	Yes	225-228	Strep
-	1	No	279	Cul + Psycho
1	-	Yes	289	Strep
1	-	Yes	315	Strep
-	1	Yes	316	Diptera
4	-	Yes	346-349	Strep
1	-	Yes	367	Strep
-	2	Yes	368-369	Cul
1	-	Yes	373	Strep
15	-	Yes	451-465	Tipu
2	-	Yes	510-511	Strep
1	-	No	516	Bra + Cul
-	1	Yes	530	Bra + Psycho + Tipu
2	-	Yes	531-532	Strep
5	-	Yes	722-726	Strep
-	2	No	727-728	Diptera except Psycho
2	-	Yes	800-801	Strep
-	3	Yes	807-809	Strep
-	1	Yes	1,166	Cul
-	1	Yes	1.175	Diptera
-	2	Yes	1,181-1,182	Diptera
1	-	Yes	1,212	Cul
2	-	Yes	1,212-1,213	Strep
1	-	Yes	1,216	Diptera + Strep
1	-	No	1,225	Psycho + 2Bra
1	-	Yes	1,229	Bra + Psycho + Tipu
-	1	Yes	1,236	Strep

-	1	Yes	1,288	Tipu
-	1	Yes	1,290	Cul
-	1	Yes	1,293	Cul except 4Cul und 5Cul
1	-	Yes	1,294	Strep
-	2	Yes	1,331-1,332	Tipu
-	1	Yes	1,609	Cul
-	1	Yes	1,610	1Cul + 2Cul + 3Cul
-	1	Yes	1,632	Cul except 4Cul und 5Cul
-	2	No	1,633-1,634	Cul except 4Cul
-	1	No	1,637	1Cul + 2Cul + 3Cul + 7Cul
1	-	Yes	1,663	Cul
1	-	Yes	1,664	Cul except 4Cul und 5Cul
1	-	No	1,666	1Cul + 3Cul
2	-	Yes	1,667-1,668	Cul
1	-	Yes	1,680	Cul
1	-	Yes	1,767	Cul
1	-	Yes	1,767	Strep
1	-	Yes	1,816	Diptera + Strep
1	-	Yes	2,181	Cul
1	-	Yes	2,183	Cul except 5Cul
1	-	Yes	2,264	1Bra + 3Bra + 4Bra
-	1	Yes	2,457	Cul except 4Cul und 5Cul
1	-	Yes	2,463	Strep
-	3	Yes	2,501-2,503	Cul
1	-	Yes	2,761	Cul
-	1	Yes	3,077	Diptera
-	1	Yes	3,085	Diptera
1	-	No	3,098	2Bra + 1Cul + 2Cul
1	-	Yes	3,099	Strep
-	1	Yes	3,114	1Bra + 3Bra + 4Bra
-	1	No	3,115	Bra + Tipu
1	-	Yes	3,344	Diptera
1	-	Yes	3,354	Diptera
-	1	No	3,360	Diptera except Psycho
-	1	No	3,366	Diptera except 4Cul und 5Cul
-	1	Yes	3,374	Bra + Psycho + Tipu
-	2	No	3,375-3,376	Psycho + Tipu
1	-	Yes	3,425	Strep
1	-	Yes	3,431	Strep
1	-	No	3,439	Bra + 2Tipu
-	1	No	3,521	Cul except 4Cul
-	1	No	3,522	Tipu + Cul except 4Cul
-	1	Yes	3,557	Cul
-	1	No	3,608	Diptera except 3Bra
-	1	No	3,702	Cul + Psycho

Poster

Titel:

Zur Logik des Gewichtens phylogenetischer Merkmale

Vorgestellt: Jahresversammlung der Deutschen Zoologischen Gesellschaft (DZG) in Osnabrück, Juni 04.-08. 2001.

Abstract publiziert in: (2001) Zoology 104 (Suppl. IV) 74

Autor: Lars Vogt
Abt. Zoomorphologie und Systematik, 33501 Bielefeld, Germany;
email: lars.vogt@biologie.uni-bielefeld.de

Abstract:

Weighting characters in cladistic analyses is a central problem in almost every attempt to reconstruct the evolution of known species. It emerges whenever two or more putative homologous characters show an inconsistent distribution among taxa. One has to decide which hypotheses of homology should be prefered. The preference or equivalence of some hypotheses of homology against others is translated to a weighting system and influences the outcome of the cladistic analysis. The logical basis for evaluating a weighting system within a falsificationist framework is the degree of corroboration of each homology hypothesis. Homology is a historical concept which refers to character transformation processes that took place in the past and left traces behind which we can observe in the present. Thus the first intuitive idea could be that the degree of corroboration is somehow linked to process probabilities. This idea is tested with simulated data. Simulated data has the benefit that the "true" phylogeny and the "true" process probabilities are known. The program ROSE (Stoye J *et al* 1998 Bioinformatics 14:157) was used to generate several datasets under varying process probability settings. These datasets were analysed using different weights and the most parsimonious tree with its tree statistics as well as a bootstrap analysis with 500 replicates were performed. The results show a strong correlation between the applied process probabilities and the "best" weighting scheme. The "best" weights are those that correspond relative to the reciprocal process probabilities. But some conditions have to be followed to apply these weights: There must exist destinguishable types of processes with different process probabilities, and character states must be unambiguously assignable to such a specific type of process. This holds true especially for the differential weighting of distinguishable classes of transformations (e.g. transitions and transversions). Whereas the differential weighting of characters (sequence positions) seems to be rather unproblematic.

Financial support by the Deutsche Forschungsgemeinschaft (Ba 1520/4-1).

Process Probabilities and the Weighting of Characters in Systematics

Following the falsificationist program of phylogenetic research

L. Vogt

Zoomorphologie und Systematik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, E-Mail: lars.vogt@biologie.uni-bielefeld.de

eingereicht in:

Organisms, Diversity & Evolution

Abstract

This paper represents an attempt to reveal the principle relevance of process probabilities of evolutionary events for cladistic analyses and for the evaluation and the weighting of cladistic characters when applying the falsificationist approach to phylogenetic research. To bring some light into the controversial discussion about the interpretation of Popper's falsificationism in phylogenetic methodology, the concepts of falsification and corroboration are distinguished. While the application of the former is straight forward, leading to the identity test within the character analysis and the congruence test within the cladistic analysis, the application of the latter is not that clear. The degree of corroboration provides the basis for choosing the best cladistic hypotheses presently available and does therefore relate to the methodological principle of general parsimony. Degrees of corroboration are needed within cladistic analysis whenever character distributions contradict the hypotheses of their synapomorphy. Characters gain corroboration by successfully passing the identity test. This is expressed by the application of character weights. When considering Popper's formula for the calculation of degrees of corroboration, the only meaningful interpretation for phylogenetic inference is the application of process probabilities. Thereby exhibiting the following relation: The more improbable the evolutionary event that caused the character, the higher its degree of corroboration, i.e. the higher its relative character weight. Thus, if phylogenetic research wants to meet the requirements of an empirical science *sensu* Popper, character weights have to be applied in cladistic analysis. And there is no justification within the falsificationist approach for the application of the so called non-weighting maximum parsimony method.

[character weighting, corroboration, evolutionary process probabilities, falsification, maximum parsimony, phylogenetic methodology]

Introduction

The evaluation of the consequences and possibilities of a refutationist approach to phylogenetic inference represents an extensive and still ongoing discussion, within which different and partially contradicting positions are advocated (e. g., Farris, 1970, 2000; Bock, 1973; Kitts, 1977; Platnick and Gaffney, 1977; Brower 2000). The only common ground among these divergent positions, so it seems, is the reference to Popper's falsificationism. Nevertheless, by referring to his methodology, different authors intend to justify the choice of different specific methods of cladistic analysis while discrediting all others. This holds true especially when parsimony and likelihood methods are compared and phylogenetic characters are analyzed and weighted for the cladistic analysis (e.g. Siddall and Kluge, 1997; Farris, 1999; Siddall, 2001). In particular the interpretation of Popper's concept of corroboration for phylogenetic inferences is a matter of recent controversy (de Queiroz and Poe, 2001; Faith and Trueman, 2001; Farris *et al.*, 2001; Kluge, 2001).

By logically distinguishing *falsification* and *corroboration* and considering their application separately in phylogenetic inference, some of the points that are made within the controversial and sometimes confusing debate can be analyzed and evaluated methodologically. The choice and justification of a specific weighting scheme within a falsificationist approach seems to represent a promising phylogenetic problem, on which the benefits and shortcomings of the popperian falsificationism can be evaluated for phylogenetic methodology.

Popper's Falsificationism

The central role within Popper's approach plays the hypothetico-deductive setting consisting of a temporarily accepted relevant background knowledge, an empirical hypothesis and empirical observations (Popper, 1994). The setting represents the basis on which an empirical test of the questioned hypothesis is performed. The severity of such a test determines the degree of corroboration a hypothesis gains when passing the test successfully. Thereby depends the severity of the test on the amount of possible falsifiers of the hypothesis which an empirical test potentially accredits (Popper, 1983, 1994). The degree of corroboration at its turn provides the basis for choosing among

many possible hypotheses the best hypothesis presently available in the light of the existing empirical knowledge. Therefore, any attempt to justify a weighting scheme for cladistic characters within a falsificationist approach should rest on the estimation of different degrees of corroboration of the corresponding character hypotheses. However, Popper's approach consists of two aspects which should be differentiated: *Falsification* and *Corroboration*.

Falsification

For an empirical hypothesis to be scientific, Popper claims that it must be fallible in principle, i.e. that it can fail on our experience. This means, in terms of the hypothetico-deductive setting, that the chosen background knowledge has to predict the theoretical possibility of empirical observations which contradict the hypothesis, so that the hypothesis would be falsified by the occurrence of this type of empirical evidence. These observations would serve as potential falsifiers of the hypothesis (Popper, 1983, 1994). If one wants to apply this principle on phylogenetic research, one has to identify the steps of testing and the mechanisms of falsification within the procedure of phylogenetic inference. In my opinion, there are at least two such steps and mechanisms, the first comprised in the character analysis, the second in the cladistic analysis.

Falsification and Character Analysis

Every biological interpretation of a perception/observation represents a hypothesis. The explanatory power of such a hypothesis depends on its consistency with other accepted theories, with the empirical evidence it is based upon, and with newly obtained empirical evidence. During character analysis traits are analyzed, described and compared to one another. As a result of this procedure, traits are coded for the data matrix as being identical or different to one another in correspondence to the results of the perceptual judgments. In the data matrix identically coded traits are, within the cladistic analysis, interpreted as representing putative synapomorphies. Following this logic, traits that are not coded identically cannot represent synapomorphies in principle. Therefore, one can conclude: Traits of two organisms of two different species are falsified in representing synapomorphies/homologies in case they exhibit no identity. This holds true, because if those traits would represent the result of a single transformation of a trait in a common ancestor of those two species and they would,

somehow, be potentially recognizable as such, they would have to be identical. If this is the case, they can be employed as serving as empirical evidence and, thus, as an argument for the reconstruction of the phylogeny of the corresponding species (for details see Vogt, 2002a, 2002b). Therefore, during character analysis, the set of purely theoretically possible hypotheses of synapomorphy is confronted with the results of the study and observations of the corresponding real entities, which represent the relevant empirical evidence against which those theoretical hypotheses are tested. The principle possibility of performing such a test against empirical evidence gives those hypotheses the status of empirical hypotheses, and provides, in case of successfully passing this test, an explanation for the empirical evidence, i.e. the observed identity of the corresponding traits among members of different species. I call this test that is embedded within the character analysis the *identity test* (Vogt, 2002a).

Falsification and Cladistic Analysis

Another step of possible falsification within phylogenetic research is found within the cladistic analysis. Here, the putative synapomorphies that result from the identity test and which are coded as identical character states within the data matrix are tested against one another in respect to some constrains deduced from relevant background knowledge, which is descent with modification. That is to say, sets of hypotheses of synapomorphy are tested for their consistency with those constraints: If the distribution pattern of at least two hypotheses of synapomorphy code for contradicting groups of species, which violate the requirement of a nested hierarchy of monophyletic taxa, the total set of hypotheses is falsified. In principle, not all of these hypotheses of synapomorphy can represent true synapomorphies – at least one of them has to represent a homoplasy. But the test does not tell which hypothesis of synapomorphy has been falsified – it only tells us that the set as a whole is incongruent. Therefore is the set of hypotheses as a whole falsified. One does not know which of the contradicting hypotheses of synapomorphy represent homoplasies. This test is known as the *congruence test* (Patterson, 1988; de Pinna, 1991; Kluge, 1997; Vogt, 2002a; *sensu* criterion of coincidence of Wagner, 1986).

Corroboration

Corroboration is another aspect of Popper's approach (Popper, 1983, 1994), which accounts for a measure of the quality of an empirical hypothesis, indicating the severity

of the empirical tests this hypothesis successfully passed without any case of falsification. Thus, the degree of corroboration depends on the severity of empirical tests, which is measured by the amount of falsifiers the test *potentially* accredited.

Following this conception, the general epistemological principle of parsimony, i.e. the economy of thinking, which is the method of preferring the hypothesis with minimum expense of explanation, can be interpreted within a falsificationist framework as the methodological convention to chose the hypothesis with the highest degree of corroboration as the best presently available hypothesis.

This has to be translated into phylogenetic methodology: During character analysis, every hypothesis of synapomorphy which successfully passes the identity test gains corroboration in dependence on the severity of its specific test. Those hypotheses that are falsified by the test are discarded and have no influence on the subsequent cladistic analysis, i.e. the corresponding traits are either not considered within further steps of phylogenetic inference or are coded differently into the data matrix. Since all hypotheses that pass the identity test gain some degree of corroboration, the cladistic analysis would be quite simple as long as the congruence test is successfully passed by the whole set of these hypotheses; i.e. as long as there exists no logical conflict in the interpretation of the empirical evidence. If there exists such conflict, one has to chose the best set of all possible alternative subsets of hypotheses of synapomorphy that are congruent to another, leaving a residue of traits which are then reinterpreted *ad hoc* as representing homoplasies. If, and *only if* all hypotheses of synapomorphy gain the same degree of corroboration by their specific identity tests - which would mean that the identity test is equally severe for each hypothesis -, parsimony would imply to chose that congruent set of hypotheses of synapomorphy that minimizes the requirement of *ad hoc* reinterpretations of homoplasy (as Kluge claims, 1997, 1997a). If this is not the case, following a falsificationist approach one has to weight the hypotheses in dependence to their specific degrees of corroboration gained in each specific identity test. And the best set of hypotheses of synapomorphy of all alternative sets would be the set of congruent hypotheses with the highest sum of degrees of corroboration (Vogt, 2002a; contradicting Kluge, 1997, 1997a).

Furthermore, one should consider that the congruence test tests hypotheses against one another (one has to keep in mind, that the congruence test does not falsify any hypothesis of synapomorphy but only sets of such hypotheses). This test can only resemble an empirical test *sensu* Popper, in case the hypotheses that are tested against

one another represent empirical hypotheses that gained at least some corroboration beforehand. Otherwise, the hypotheses would merely be tested against the imagination of their inventor, which, if at all, might present a fruitful basis for an entertaining story but not the methodological grounds of scientific reasoning. Therefore depends the weight, a character is given in the cladistic analysis, directly on the severity of the identity test, which the character has passed successfully within the character analysis (Vogt, 2002a). According to this view, not weighting characters, as Kluge (1997, 1997a) demands, would imply that the corresponding hypotheses have no explanatory power at all. And performing a cladistic analysis with empirically empty hypotheses, that are tested against each others on congruence, would result in a most congruent, but also empirically empty cladistic hypothesis. This would equal a procedure of coding for a data matrix without previously performing a character analysis and, thus, without performing identity tests. Such a procedure would be unacceptable.

Following this, one can conclude: If the severity of the performed identity tests varies, one should account for this by giving corresponding weights (Vogt, 2002a). This confronts one with the necessity of determining the basis of evaluating the severity of the identity tests.

Weighting and Corroboration

What serves as an objective measure of the severity of an identity test? For morphological traits Neff (1986) suggests to take the structural complexity into account when weighting characters (see also Patterson, 1988; Wägele, 1995, 1996), giving more complex structured traits a higher weight. Besides the problems of coding such traits in a data matrix, examples of complex structured traits that exhibit a high intraspecific and interspecific variability (e.g. variable patterns in the shells of species of the genus *Conus* (Meinhardt, 1996)) contradict the generality of this principle, since they represent complex structured traits that bear little phylogenetic information. Therefore, this argumentation does not seem to provide a proper basis for weighting characters.

On the other hand, taking process probabilities/frequencies of the corresponding types of evolutionary events as the basis for weighting cladistic characters represents an ongoing discussion and especially advocates of a falsificationist approach reject their consideration. They neglect that those probabilities should be considered, because of the problems which occur by modeling evolution. Some authors even state that considering

process probabilities would unnecessarily add to background knowledge which would decrease the testability of hypotheses and, thus, would represent an inductive and therefore refutable procedure (Kluge, 1997), because induction would be inconsistent with popperian logic.

This paper represents an attempt to reveal the principle relevance of process probabilities of evolutionary events for cladistic analyses and for the evaluation and the weighting of cladistic characters when applying a falsificationist approach.

Process Probability and Reconstructing Phylogeny

I want to begin my argumentation by describing a possible scenario under the assumption of descent with modification for molecular data:

Let there be an organism carrying an individual guanine G_{y1} at a specific sequence position within its genome (Fig. 1, top right side). Having guanine at this sequence position is subsequently inherited to all the offspring belonging to a line of descent of this organism via mechanisms of DNA replication, cell cleavage, reproduction etc. As a consequence, an evolutionary line of individual G's is constituted (G_{y1}, G_{y2}, G_{y3}, G_{y4}),

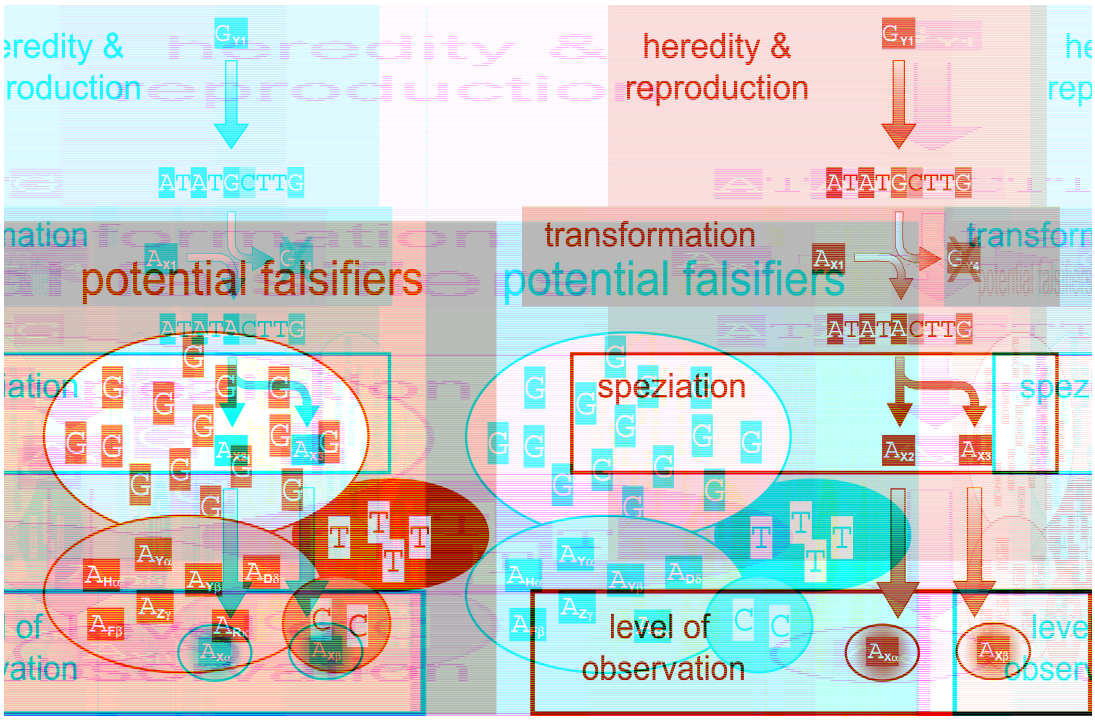

Fig. 1: Classes of identity. An example of the dependence of the severity of the identity test from the process probability of the corresponding type of transformation. For more details see text.

that can be classified to a specific class of guanines G_y. At some point a transformation takes place, changing G_{y4} to A_{x1}. A line of individual adenines is initiated, determining a specific class of adenines A_x. Speciation events take place and all species that come into being posses A_x at this specific sequence position. These A_x's can be traced back to A_1 which initiated this specific ontological historical class.
Now, if an observer tries to investigate the phylogeny of those species and she considers this specific sequence position, she first performs the identity test on this trait. As logical falsifiers for the hypothesis (trait at this position of the sequence originated in the specific transformation G_y to A_x) serve all guanines, all thymines, all cythosines *and* all adenines that do not belong to the class A_x (Fig. 1, left side). However, since no one can distinguish adenines belonging to A_x from adenines not belonging to A_x, the identity test accredits only those potential falsifiers that are not adenine. If there are only a no/few adenines of the type *non-x*, the identity test would accredit all/almost all logical falsifiers.

Thus, one is confronted with the question, how one can evaluate the ratio of effectively accredited potential falsifiers to all logical falsifiers.

For the case described above, the only possible interpretation of this problem I can think of is the following: Since identical nucleotides are classified to sets, one receives four different sets. To evaluate the relative amount of accredited potential falsifiers of the hypothesis, the relative size of the three sets of potential falsifiers (guanines, thymines and cythosines) in comparison to the set of adenines has to be determined. As the hypothesis is a historical hypothesis, referring to an event and thus to a process, the size of the sets can best be interpreted as depending on the corresponding process probabilities. If the size of the sets corresponds to their underlying process probabilities, each set should consequently be interpreted as consisting of an infinite number of elements. Therefore, the relative size of a set, and therewith the amount of accredited potential falsifiers, depends on the relative cardinality of the sets. When all four possible sets together equal the general probability of a nucleotide substitution, the concrete cardinality of each set can be interpreted as directly depending on the corresponding process probability in the moment of the mutation. The relative amount of accredited falsifiers would equal $1\text{-}p_{(specific\ transformation)}$. Thus, in the case discussed above, the amount of potential falsifiers accredited by the identity test depends on the process

probability of transforming guanine to adenine in the moment of transformation. The more improbable this specific transformation, the higher the relative cardinality of the three falsifying identity classes, the more severe the identity test respectively. This relation results from the human inability to distinguish different ontological historical entities which do not exhibit different qualities. Following this, the degree of corroboration of a hypothesis of synapomorphy which successfully passed the identity test should strongly depend on the *improbability* of the trait-causing transformation process. This corresponds largely with Popper's interpretation of the calculus of probability and with the falsifiability of a hypothesis as depending on its improbability (Popper, 1983, 1994).

It should be added that it is only the propensity interpretation of the calculus of probability which should be applied (Mahner and Bunge, 1997; Popper, 1983) when considering those evolutionary process probabilities. This is an important point that has to be considered when applying assumptions of process probabilities in cladistic analyses.

Process Probability and Corroboration

I want to confront my point of view of the relevance of evolutionary process probabilities in phylogenetic inference within a falsificationist approach with the formalism, which Popper himself suggested. He proposes a formula to evaluate the degree of corroboration $C_{(h,\ e,\ b)}$ provided to a hypothesis h by empirical evidence e in the light of the relevant background knowledge b (Popper, 1994):

$$C_{(h,\ e,\ b)} = \frac{p_{(e,\ hb)} - p_{(e,\ b)}}{p_{(e,\ hb)} - p_{(eh,\ b)} + p_{(e,\ b)}} \quad ;$$

where for instance $p_{(e,\ b)}$ is read as the probability of e given b. From this formula one receives that any e supports h in the case when $p_{(e,\ hb)} > p_{(e,\ b)}$, obtaining a positive $C_{(h,\ e,\ b)}$ (Popper, 1983; Farris, 2000).

Considering this formula when evaluating the effect of process probabilities on the severity of the identity test, one could compare two contradicting hypotheses of

synapomorphy. Trait x and trait y shall represent two contradicting hypotheses of synapomorphy. Both hypotheses successfully passed the identity test. Now, one wants to evaluate their corresponding degree of explanatory power to be able to decide which hypothesis to prefer. Let there be:

e_x = There are several organisms, each bearing an identical trait x, and these organisms belong to several species;

h_x = All x represent synapomorphies and are as such the result of a single transformation event that took place in the last common ancestor of the species, whose organisms bear x;

b_x = descent with modification; a transformation to trait x is less probable than a transformation to trait y;

The corresponding applies to e_y, h_y and b_y.

By considering the formula above, one receives for hypothesis x and y the following:

$p_{x(e,\,hb)} = p_{y(e,\,hb)}$, because, if h and b are both given, either transformation – to trait x and to trait y – took place only once, since those traits represent synapomorphies. Therefore, do h_xb_x and h_yb_y not differ from one another in respect to their effect on the probability to observe e_x or e_y.

$p_{x(e,\,b)} < p_{y(e,\,b)}$, because the probability to observe e_x or to observe e_y, if only the background knowledge is given (which is identical for both cases), is only dependent on the process probability of the transformation causing the corresponding trait. Since this probability is smaller for trait x than for trait y, one receives this inequality.

$p_{x(eh,\,b)} > p_{y(eh,\,b)}$, because the smaller the process probability, the higher the probability that the corresponding transformation happens only once.

Thus, considering the formula and the inequalities stated above, one receives the relative degree of corroboration of hypothesis *x* and *y* as: $C_{(hx, ex, bx)} > C_{(hy, ey, by)}$.
Therefore, one has to conclude that - from a falsificationist point of view and a propensity interpretation of the calculus of probability - process probabilities are intimately linked to the severity of the identity test and, therewith, to the explanatory power of a tested hypothesis of synapomorphy.
And one has to conclude that the quality of a trait in its function as serving as phylogenetically relevant evidence strongly depends on the process probability of its specific type of causing transformation.
When considering the formula for the measure of the amount of accredited falsifiers proposed above ($1\text{-}p_{(specific\ \ transformation)}$), one could, in case the relevant process probabilities are known, propose a formula representing an easy to handle approximation to the evaluation of relative degrees of corroboration of hypotheses of synapomorphy that have passed the identity test successfully:

$C_{Hyp} = (1\text{-}á)\ (1\text{-}â)$; C_{Hyp} *representing the degree of corroboration of a hypothesis of synapomorphy;* ***á*** *representing the process probability of a specific type of transformation (e.g. nucleotide substitution);* ***â*** *representing the process probability of a specific result of a transformation type.*

I want to stress that this formula represents only a rough approximation to give a vague impression of the relation of process probability and character weight - I do not recommend its application.

Process Probability and Phylogenetic Information Content

To know the relevant process probabilities and to be able to apply them for the estimation of proper character weights, the process probabilities have to be unambiguously assignable to the observable empirical evidence, thereby giving an explanation for their cause. What this means for cladistic analyses can best be illustrated by another example. Considering the process of *unequal crossing over*, while neglecting any effects of selection, such an event can, when occurring during meiosis, result in two possible lines of descent, one representing the result of a deletion of, e.g. 7 nucleotides,

and the other of an insertion of the same length. In this case, the process probabilities for receiving a deletion of this length and for receiving an insertion of the same length, have to be equal, since they represent the results of the same single event (Fig. 2, left). However, if one would compare deletions resulting from 5 independent unequal crossing over events without considering their 'neighboring' base composition, one could not distinguish them from one another; while, with the corresponding insertions, this might still be possible (Fig. 2, right). Therefore, one could claim, that insertion characters of a specific nucleotide length bear more phylogenetic information than deletion characters of the same length (Vogt, 2002b).

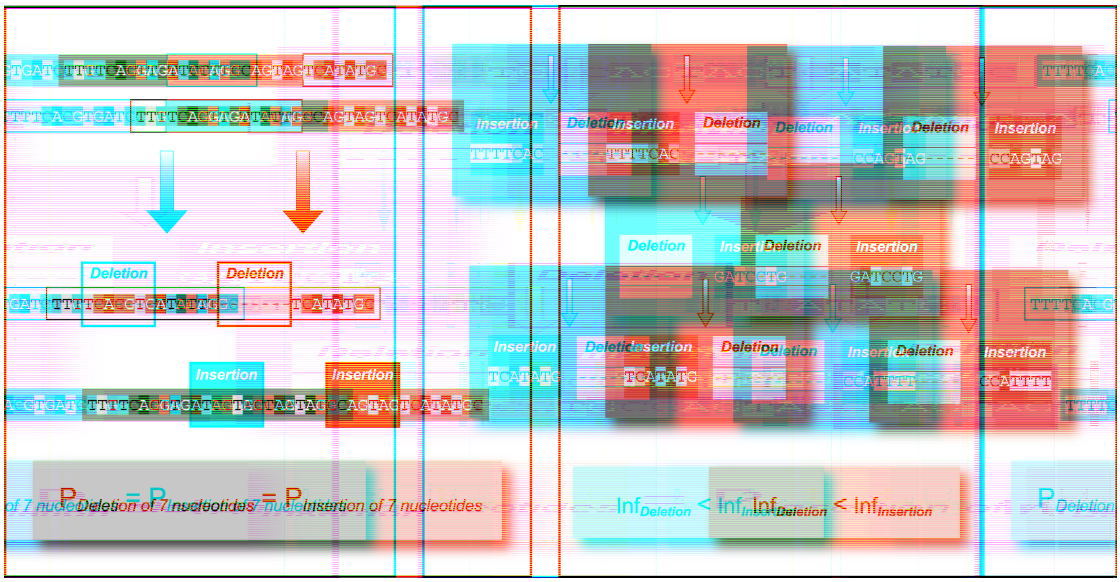

Fig. 2, *Left*: Unequal crossing over. A single event of unequal crossing over during meiosis can lead to two lines of descent, one exhibiting the result of a deletion of, in this case, 7 nucleotides, and the other of an insertion of the same length. Since both are caused by the same single evolutionary event, their corresponding process probabilities are identical. ***Right*: Comparison of the results of five such events of unequal crossing over.** When comparing the same facts in the case of 5 independent events of unequal crossing over, the deletions, according to their nature, cannot be distinguished from one another, while the insertion can, according to their specific base composition. Therefore, insertions of a specific length bear a higher phylogenetic information content (Inf) than deletions of the same length.

This specific example in mind, one has to ask oneself what is the source for the higher amount of phylogenetic information content of insertions in comparison to deletions and how is this interpreted within a falsificationist framework? Thus, what makes this insertion more valuable than the deletion of the same length in cladistic analyses?

Again, this can best be consistently explained within a falsificationist framework in terms of process probabilities and the possibility to assign empirical evidence most unambiguously to its causing type of transformation. The probability to receive a deletion of a specific length is much higher than of an insertion of the same length *and* a

specific base composition. Thus, the information content of a trait seems to be strongly dependent on the process probability of its causing type of transformation, thereby exhibiting the relation: the more improbable the types of transformation that are known to be able to cause a specific trait, the higher its information content after its corresponding hypothesis successfully passed the identity test. In this way does the information content of the distribution pattern of an identical trait equal the degree of corroboration of the corresponding hypothesis of synapomorphy.

Simulation

To test the relationship of process probabilities and weighting of cladistic characters within the procedure of parsimony analysis, a simulation study is performed. The simulation study and its results are presented as electronic appendix in the electronic supplements of this journal (http://www.senckenberg.uni-frankfurt.de/odes/).

Estimating Process Probabilities

The results of the simulation study together with the results of the theoretical analysis show that reasonable weighting is intimately linked to the process probabilities of the transformation-causing evolutionary events. Considering the inevitable necessity of weighting the empirical evidence in cladistic analyses (Vogt, 2002a) – whether equally or differentially -, the evaluation of the relevant process probabilities, as well as the assignation of process types to corresponding types of empirical evidence, represent key problems within every cladistic inference. Especially estimating the process probabilities proves to be the most problematic step. Since the application of the calculus of probability within a falsificationist approach premises a propensity interpretation (Mahner and Bunge, 1997), the process probabilities cannot be estimated by simply counting the frequencies of specific types of mismatches of pair wise aligned sequences. Such a procedure assumes that the observed sample represents the entire class of all past and future transformations and thus, that the sample is in equilibrium. The correctness of this assumption cannot be tested for real data and is therefore problematic. And even in case the general mutation probabilities would be known (by e.g. consideration of laws of biochemistry and physics), the probability of a specific transformation is not directly deducible from them, since selection forces influence

transformation probabilities as well. And the effects of different selection pressures and their impact on the phylogenetic information content of characters is yet not well enough understood to be systematically incorporated into a weighting scheme.

Thus, one is confronted with lots of open questions: How can one evaluate process probabilities empirically without knowing the corresponding phylogeny of the sequences? Are experimentally obtained process probabilities universal and, thus, applicable on every empirical data? What role plays selection and can selection be estimated/approximated and systematically incorporated into a weighting scheme or is this in principle impossible?

All those questions concern the problem of evaluating the results of historical processes that necessarily represent singularities, and the difficulties that emerge when one wants to generalize them into classes of processes within the analysis, to be able to handle them and interpret them in respect to our evolutionary background knowledge. But one should be aware of not getting caught into the trap of radical skepticism by discarding every statement and method that is not 100% validated, ending up in having no explanation at all for the observed distribution of identical traits among members of different species, because nothing bears up against such an impracticable standard.
However, since systematicists cannot bring history into the lab and force it to answer repeatedly questions about the past, one has to heuristically find one's way to always better approximations thereby applying methods, that might still fail under certain conditions but which tend to be more truth-keeping than all other methods. There is a need of methods that meet the dynamics of historical events and that overcome the boundaries of mono-causal reasoning.

However, though estimating transformation probabilities correctly is linked to many methodological and epistemological problems, there is no convincing argument within a refutationist approach to disregard process probabilities or even prohibit their consideration when weighting cladistic characters. On the contrary, they have to be approximated somehow, if one wants to justify scientific hypotheses about biological history at all.
To evaluate process probabilities, it might be useful to apply statistical methods and evolutionary models. And even if process probabilities cannot be statistically

approximated, as long as empirical evidence can be assigned to the occurrence of a specific type of transformation, the possible number of results of this type of transformation can be considered to give a first, but weak, approximation of the corresponding probabilities: the higher the number of possible different results is that a specific type of transformation potentially has, the more improbable is it to receive a specific result of this type of transformation, and the higher is its phylogenetic information content and, thus, its empirical weight (in accordance with Vogt, 2002a).

Conclusion

While the application of the concept of falsification in systematics is straight forward, the application of the concept of corroboration raises methodological challenges. It is the weighting of characters where the concept of corroboration has to provide the methodological basis for. However, Popper' falsificationism gives no convincing argument against the employment of process probabilities in phylogenetic research. On the contrary, only the application of process probabilities makes a meaningful interpretation of Popper's formula of the degree of corroboration within phylogenetic methodology possible. And as weighting is inevitably necessary if cladistic hypotheses shall represent empirical hypotheses *sensu* Popper, weights have to be applied. And since these weights are intimately linked to the corresponding transformation probabilities, *non-weighting* phylogenetic evidence would resemble an epistemologically inconsistent method and, therefore, does not provide a reasonable alternative.

Therefore, in the light of this study, the application of the specific type of maximum parsimony method that principally applies equal weights without giving an empirical justification for this weighting scheme, cannot not be justified with Popper's falsificationism; moreover, it is inconsistent with a falsificationist approach to phylogenetic inference (contradicting Kluge, 1997). There is no way to avoid the application of character weights in phylogenetic inferences.

References

Bock, W. J. (1973): Philosophical foundations of classical evolutionary classification. Syst. Zool. 22: 375-392.

Brower, A. V. Z. (2000): Evolution is not a necessary assumption of cladistics. Cladistics 16: 143-154.

de Pinna, M. C. C. (1991) : Concepts and tests of homology in the cladistic paradigm. Cladistics 7: 367-394.

de Queiroz, K. and Poe, S. (2001) : Philosophy and phylogenetic inference : A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. Syst. Biol. 50: 305-321.

Faith, D. P. and Trueman, J. W. H. (2001): Towards an inclusive philosophy for phylogenetic inference. Syst. Biol. 50(3): 331-350.

Farris, J. S. (1970): Methods for computing Wagner trees. Syst. Zool. 19: 83-92.

Farris, J. S. (1999): Likelihood and inconsistency. Cladistics 15(2): 199-204.

Farris, J. S. (2000): Corroboration versus "strongest evidence". Cladistics 16: 385-393.

Farris, J. S., Kluge, A.G. and Carpenter, J. M. (2001): Popper and likelihood versus "Popper*". Syst. Biol. 50(3): 438-444.

Kitts, D. B. (1977): Karl Popper, verifiability, and systematic zoology. Syst. Zool. 26: 185-194.

Kluge, A. G. (1997): Testability and the Refutation and Corroboration of Cladistic Hypotheses. Cladistics 13: 81-96.

Kluge, A. G. (1997a): Sophisticated falsification and research cycles: Consequences for differential character weighting in phylogenetic systematics. Zool. Scr. 26: 349-360.

Kluge, A. G. (2001): Philosophical conjectures and their refutation. Syst. Biol. 50(3): 322-330.

Mahner, M. and Bunge, M. (1997): Foundations of biophilosophy. Springer, Berlin.

Meinhardt, H. (1996): Models of biological pattern formation: common mechanism in plant and animal development. Int. J. Dev. Biol. 40, 123-134.

Neff, N. A. (1986): A rational basis for a priori character weighting. Syst. Zool. 35(1): 110-123.

Patterson, C. (1988): Homology in Classical and Molecular Biology. Mol. Biol. Evol. 5: 603-625.

Platnick, N. I., and Gaffney, E. S. (1977): Systematics: A Popperian Perspective. Syst. Zool. 26: 360-365.

Popper, K. R. (1983): Realism and the Aim of Science; From the Postscript to the Logic of Scientific Discovery. London and New York: Routledge,.

Popper, K. R. (1994): Logik der Forschung. (10., verbesserte und vermehrte Auflage von "Die Einheit der Gesellschaftswissenschaften; Band 4). Tübingen: J. C. B. Mohr (Paul Siebeck).

Siddall, M. E. and Kluge, A. G. (1997): Probabilism and phylogenetic inference. Cladistics 13: 313-336.

Siddall, M. E. (2001): Philosophy and phylogenetic inference: A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. Cladistics 17, 395-399.

Vogt, L. (2002a): Testing and weighting characters. O. D. E. 2:

Vogt, L. (2002b): Weighting of indels. O. D. E. 2:

Wägele, J. W. (1995): On the information content of characters in comparative morphology and molecular systematics. Z. Zool. Syst. Evol. 33: 42-47.

Wägele, J. W. (1996): First principles of phylogenetic systematics, a basis for numerical methods used for morphological and molecular characters. Vie Milieu46 : 125-138.

Wagner, G. P. (1986): The systems approach: an interface between development and population genetic aspects of evolution. *In* "Patterns and processes in the history of life" (D. M. Raup and D. Jablonski, Eds.), pp. 149-165. Springer, Berlin.

Electronic Appendix

Simulation Study to “Process Probabilities and the Weighting of Characters in Systematics”

To test the relationship of process probabilities and the weighting of cladistic characters within the procedure of parsimony analysis and, thus, the effect of the degree of corroboration of hypotheses of synapomorphy gained by successfully passing the identity test (Vogt, 2002) on the power of the analysis, a simulation study is performed. Different data sets with varying parameters are generated. To reduce the amount of variable parameters and to simplify the evaluation of the generated data, only nucleotide substitutions are allowed to 'evolve' (i.e. no indels). Additionally to the data sets, each corresponding 'evolutionary history' and the 'true' alignment is logged. This enables one to perform cladistic analyses for each data set under varying character weights and to compare the results with the 'true phylogeny' and the corresponding evolutionary parameters. Since the initial starting sequence, the 'true' ancestral sequence of the generated sequences, is previously known, it is used as the outgroup in the cladistic analysis, to root the resulting tree correctly. For the cladistic analyses, the maximum parsimony method is applied, since this method allows the application of different weights. Furthermore, maximum parsimony is advocated by those authors that claim to follow a refutationist approach but who propagate equal weighting as the only weighting scheme which is consistent with popperian falsificationism (e.g. Siddall and Kluge, 1997; Farris, 1999).

Performing this simulation has neither the intention to confirm a specific weighting scheme nor to test the parsimony method of cladistic analysis. It only attempts to test whether process probabilities have to be considered when interpreting the products of an evolving system which evolves under specific process probabilities. Moreover, this simulation study is performed to investigate how those probabilities are linked to an 'optimal' weighting scheme.

Methods

For the simulation of evolutionary processes and the generation of the different data sets, the program ROSE (Stoye *et al.*, 1998) was employed, thereby deactivating the function of the program that allows also indels to evolve. For the cladistic analyses PAUP*4 (Swofford, 1998) was used under parsimony settings.

Data sets

1. unconstrained generation of the sequences, thus evolving on a symmetrical tree – in sum 68 data sets

1.1 four classes of different values of relatedness (see Stoye *et al.*, 1998): 2000; 1000; 500 and 250

1.1.1 specific relative transformation probabilities of transition to transversion (Ts:Tv) for each class of relatedness - Ts:Tv with the values 15:1; 12:1; 9:1; 6:1; 5:1; 4:1; 3:1; 2:1; 1:1; 1:2; mixed

1.1.2 mutation probability of specific sequence positions for each class of relatedness; from the first sequence position onward and repeating until the end of the sequence - with the values 0.9, 0.7, 0.5; 0.9, 0.6, 0.3; 0.3, 0.2, 0.1; 0.9, 0.1; 0.2, 0.1; 0.9, 0.8

2. constrained generation of the sequences, thus evolving on a given asymmetrical tree – in sum 34 data sets

2.1 two classes with a different general mutation probability on every sequence position: 0.9 and 2.0

1.1.3 specific relative transformation probabilities of transition to transversion for both classes - Ts:Tv with the values15:1; 12:1; 9:1; 6:1; 5:1; 4:1; 3:1; 2:1; 1:1; 1:2; mixed

2.2 mutation probability of specific sequence positions; from the first sequence position onward and repeating until the end of the sequence - with the values 0.9, 0.7, 0.5; 0.9, 0.6, 0.3; 0.3, 0.2, 0.1; 0.9, 0.1; 0.2, 0.1; 0.9, 0.8; and 9.0, 7.0, 5.0; 9.0, 6.0, 3.0; 3.0, 2.0, 1.0 9.0, 1.0; 2.0, 1.0; 9.0, 8.0

Analyses

All 102 data sets were each analyzed applying different relative weights:

Each Ts:Tv data set with varying specific step matrices, relatively weighting Ts to Tv:

2:1; 3:2; 1:1; 2:3; 1:2; 1:3; 1:4; 1:5; 1:6; 1:9; 1:12; 1:15; mixed.

All other data sets each with varying specific character weights:

2 3 6; 3 6 2; 6 2 3; 6 3 2; 4 5 6; 5 6 4; 6 4 5; 6 5 4; 1 2; 2 1; 1 9; 9 1; 8 9; 9 8; 3 1; 1 3; 1 1.

For all these settings, parsimony analyses under hsearch/NNI default settings and parsimony bootstrapping analyses with 500 pseudoreplicates were performed. The results, together with the ensemble consistency index (CI) and the ensemble rescaled consistency index (RC) were logged and compared to the ‘true’ tree. In cases where the resulting tree topologies differed from the ‘true’ phylogeny, the number of branches that one had to swap to obtain the correct topology were counted and logged as long as it did not exceed the number five. The bootstrap frequencies were ordered to classes and the corresponding nodes for every such class were counted, as well as the number of collapsed nodes.

For every data set that was analyzed with different weights, the results are compared to the true phylogeny. The results are evaluated and classified into relatively good and relatively bad results for the analyses of each data set. The criterion for relatively good results is the comparably better fit of their tree topologies with the 'true' tree, which is investigated by counting the number of branches that would have to be swapped to receive the 'true' tree, by the resolution of the phylogeny (number of collapsed nodes) and by the bootstrap frequencies of the nodes. Thereby, the fit has the highest significance, followed by the number of collapsed nodes and by the bootstrap frequencies.

Results

A first and not very surprising result is, that the quality of the cladistic reconstruction depends on the evolutionary rate of the employed sequences. The higher the relatedness parameter is set - which equals increasing the evolutionary rate - the worse is the resolution and the worse is the bootstrap support of the results of the corresponding bootstrap parsimony analyses (exemplified in Fig. 1 for the analysis of the data sets described under 1.1.1 in *Simulation – Material and Methods*). For the data sets that were generated with a relatedness parameter of 250, almost all relationships are resolved by the analyses, and the obtained bootstrap frequencies are all very high. This changes continuously to the worse when analyzing the data sets that were simulated with a higher relatedness value. Many, and in some analyses all of the nodes collapsed into unresolved polytomies and only few nodes, if at all, show bootstrap frequencies of 1.00 to 0.85. This result coincides with the results from the analyses of the other data sets – the higher the rate of transformation, the worse the results of the analyses.

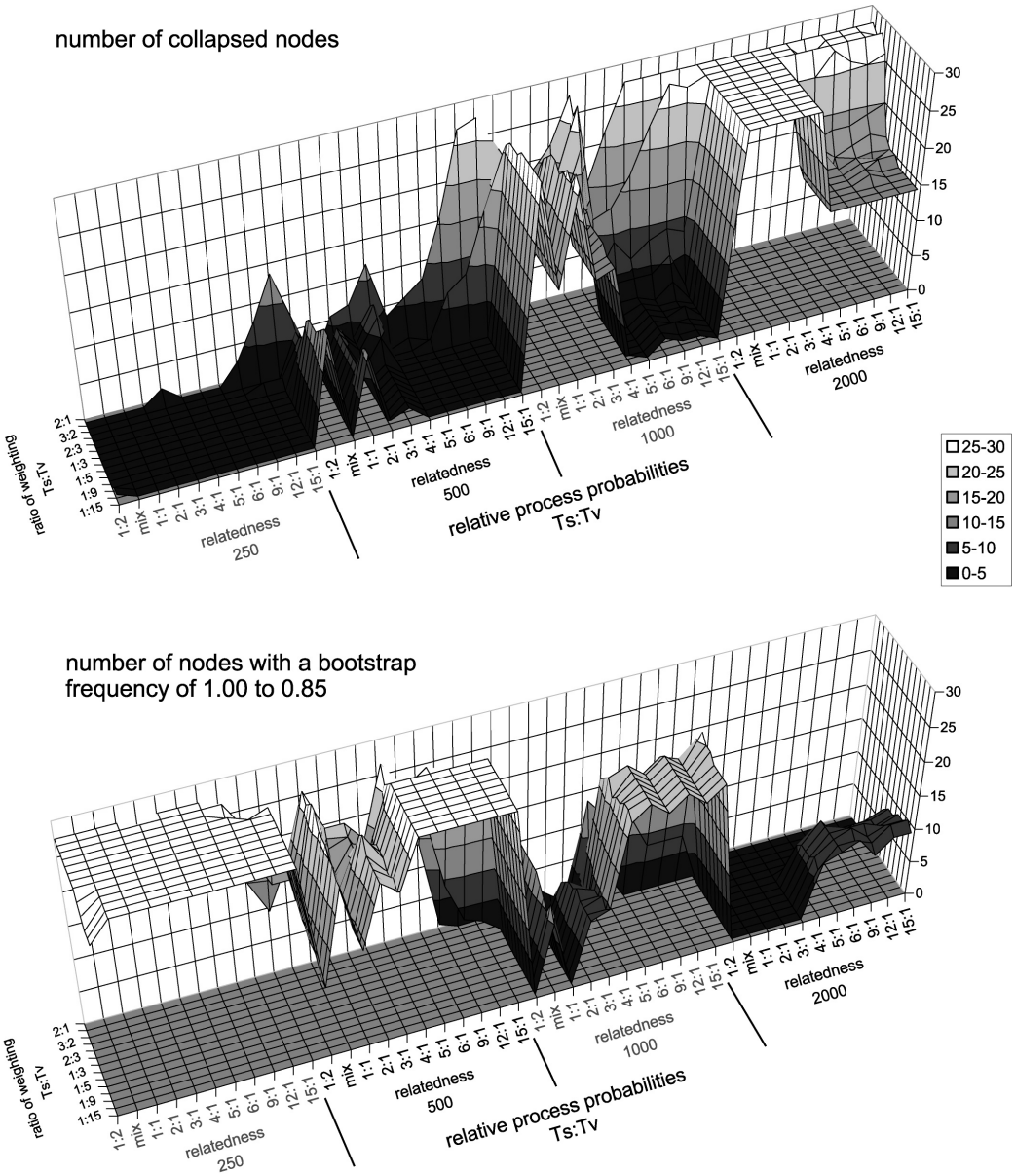

Fig. 1 Results of the bootstrap-analyses of the data set 1.1.1 (see *Simulation – Methods*).*Top:* Number of to polytomies collapsed nodes for the analyses of the data sets of varying general mutation rates and different relative rates of Ts:Tv. The number of collapsed nodes increases with the mutation rate and thus, the resolution decreases. *Bottom:* Number of nodes with a bootstrap frequency of 1.00 to 0.85 for the analyses of the same data sets. The number of nodes with such a high bootstrap frequency decreases with an increasing mutation rate.

Each simulated data set has been analyzed according to different relative weights. Figure 2 illustrates the outcome of the comparison of the results for all analyses of all generated data sets. The comparably better results are indicated by dark squares. One has to be aware of the fact, that the dark squares represent the outcome of a comparison of the results of different analyses of a single data set, all mapped into a single row. Thus, they indicate results that are similarly good and comparably better than the others a similar and absolute quality of results.

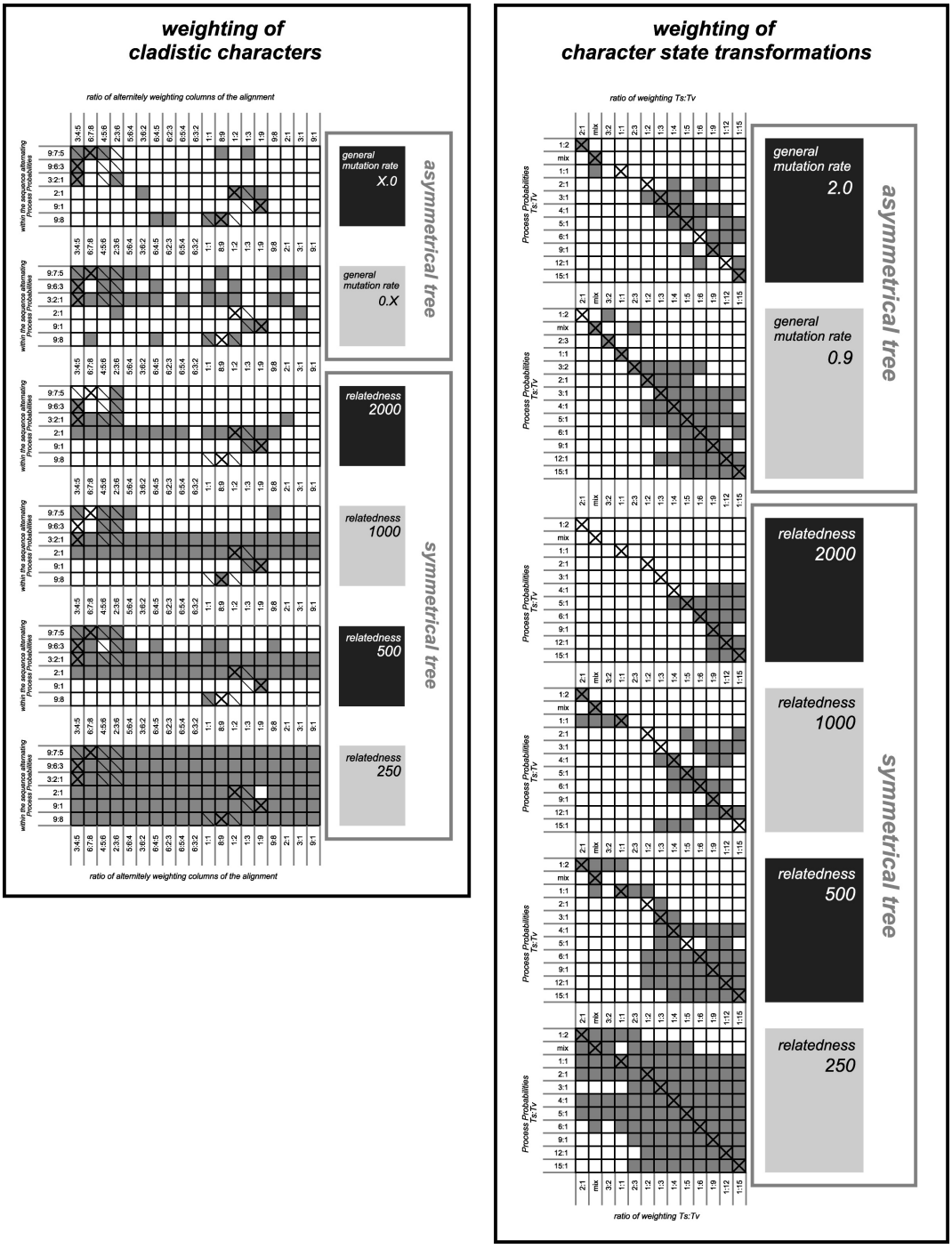

Fig. 2 Results of the different cladistic analyses of all generated data sets. The different data sets are ordered according to whether they were generated on a given asymmetrical tree or on a symmetrical tree, according to the mutation rate indicated by the different values for relatedness and the general mutation rate, and according to the relative mutation rates of sequence positions (left side) or the ratio of transitions (Ts) to transversions (Tv) (right side). Every row illustrates the results of the analyses of a single data set. Every column represents the results of analyses that applied a specific weighting scheme in the parsimony analysis. The results from the analyses of each data set that applied the predicted optimal weights are indicated by crossed squares, those that applied semi-optimal weights by back-slashed squares. The dark squares represent those analyses that obtained better results in comparison to the other analyses of the same data set (for details see text). Therefore, analyses that are represented by dark squares indicate results of comparably similar quality for every single row separately. They are not absolutely comparable between different rows. The dark squares indicate only, that the analyses of a data set yield better results when applying the corresponding weights than those indicated by the bright squares.

Additionally, in correspondence to the formula for the measure of the amount of accredited falsifiers, the analyses that applied the predicted optimal relative weights on the corresponding data set are represented by crossed squares, and semi-optimal by a backslash. One would expect, that the application of optimal and semi-optimal weights would more frequently lead to comparably better results than these are obtained in general.

To get a measure to test this expectation with the results of the simulation study, a *general* ratio of better-results-to-all-results has been calculated for each data set. Example: For the data set of 'weighted cladistic characters' (see Fig. 2), relatedness 2000, alternating process probabilities of 3:2:1, the results of 19 analyses are listed - each having applied different weights -, of which 5 represent comparably better results; this gives rise to the general ratio of 0.2632.

A *specific* and a *semi-specific* ratio has been calculated, which are the ratios of predicted optimal and semi-optimal weights, that have been applied and that lead to relatively better results, compared to the sum of analyses that applied optimal and semi-optimal weights. Example: For the same data set as in the example above, there are 3 analyses listed with 1 optimal and 2 semi-optimal weights, of which all 3 perform relatively good, giving rise to the ratio 1.00 for the optimal and semi-optimal analyses and 1.00 for the optimal analysis only.

In this way, 2-3 ratios were calculated, evaluating the results of the different analyses for each data set. For corroborating the prediction, one would, in average, expect to receive a high specific ratio, a smaller semi-specific ratio and an even smaller general ratio. On this account, the mean value from the three types of ratios was calculated for all data sets with varying transition-transversion mutation rates and for all data sets with varying mutation rates for sequence positions (Table 1). In both cases, the mean general

Table 1: The mean general, mean semi-optimal and mean optimal ratio obtained for the two types of data sets. For details see text.

	weighting of cladistic characters	weighting of character state transformations
mean general ratio	0.4269	0.3450
mean semi-optimal ratio	0.7708	-
mean optimal ratio	0.8611	0.7647

ratio is smaller than the mean specific and mean semi-specific ratio. The mean specific ratios exhibit the highest value. The results, thus, corroborate the prediction.

Furthermore, the simulation study reveals the weakness of the ensemble consistency index (CI) and the ensemble rescaled consistency index (RC) in their ability to indicate the quality of the results of cladistic analyses. There seems to be no strong correlation between the distribution of better results to the distribution of results obtaining comparably high values for the CI and RC (results not illustrated).

Table 2: Ratios of wrong clades that received a bootstrap support of at least 0.5, at least 0.85, and at least 0.95 to the total number of nodes. These ratios were calculated for both types of data sets and for all data sets together.

ratios of clades with specific bootstrap frequencies to total number of nodes	weighting of cladistic characters	weighting of character state transformations	all data sets together
0.5	0.006975	0.007970	0.007533
0.85	0.002959	0.002034	0.002441
0.95	0.002114	0.000706	0.001325

The received bootstrap frequencies seem to be more informative, since only few nodes with a bootstrap support of at least 0.5 represent a wrong grouping in comparison to the true phylogeny (Table 2). Though there are some wrong clades with a support of even 0.95-1.00, these are extremely rare.

Discussion

The results of the simulation study obviously corroborate the hypothesis, that the application of weights, which correspond to the relevant process probabilities that generated the data, results in a better reconstruction of the history of the sequences.

Asymmetry of Transition-Transversion

A problem when weighting nucleotide substitution characters concerns the nature of transitions and transversions. A nucleotide mismatch in a pair of sequences of two purines can be explained by a single or a multitude of transitions, but it can also be explained by two or more transversions only. A mismatch of a purine and a pyrimidine, though, cannot be explained by referring only to transitions – at least always one transversion can be assumed as well. This could affect the prediction of optimal weights and their reasonable application within the cladistic analysis using the method of

maximum parsimony, since the correlation of the empirical data to the process type that could have caused the data is not unambiguously possible – there exists an asymmetry in the mapping of the two processes causing these specific transformations and their corresponding observable mismatches between two sequences. Particularly, in case of a relatively high transformation rate of transversions, the application of 'optimal' weights could be problematic.

Let the relative probability for a transition be $P_{Ts} = 0.4$, it follows $P_{Tv} = 0.6$. Transitions are, as described above, only responsible for mismatches between pyrimidines or between purines (type A mismatch). Transversions – especially in the case of a high general mutation rate that rises the chance of multiple hits – are responsible for mismatches of type A as well as for mismatches between pyrimidines and purines (type B mismatch). This leads to a relative frequency of observable mismatches of type A that, with a value of $F_{type\ A}$ 0.4, exceeds what can be explained by transitions only. However, the maximum parsimony method would most parsimoniously explain these mismatches by transition events. In case the optimal weights would be applied (relative weighting of Ts:Tv with 3:2), some mismatches that have been caused by two or more transversions would be incorrectly interpreted as the result of a single transition, receiving a wrong weight. Yet, as long as the relative process probability of transitions is higher than of transversions, the effect of this 'asymmetry' seems to be negligible. The method of maximum likelihood takes this asymmetry into consideration.

References

Farris, J. S. (1999): Likelihood and inconsistency. Cladistics 15(2): 199-204.

Siddall, M. E. and Kluge, A. G. (1997): Probabilism and phylogenetic inference. Cladistics 13: 313-336.

Stoye, J., Evers, D., Meyer, F. (1998): Rose: generating sequence families. Bioinformatics 14(2): 157-163.

Swofford, D. L. (1998): *PAUP4.0*: phylogenetic analysis using parsimony and other methods (beta version). Sinauer, Sunderland, Massachusetts.

Vogt, L. (2002): Testing and weighting characters. O. D. E. 2:

Vortrag

Titel:

Tatort Evolution

Zeichentheorie und Phylogenetik

Gehalten: Gemeinsame Jahrestagung der Paläontologischen Gesellschaft und der Gesellschaft für Biologische Systematik (GfBS) in Oldenburg, Sept 17.- 21. 2001.

Abstract publiziert in: (2001) Terra Nostra 01/6, 122.

Autor: Lars Vogt
Abt. Zoomorphologie und Systematik, 33501 Bielefeld, Germany;
email: lars.vogt@biologie.uni-bielefeld.de

Abstract:

Der Versuch der Rekonstruktion einer Phylogenie weist in vielerlei Hinsicht deutliche Parallelen zu der Arbeit eines Detektiven oder Ermittlers auf, der den Tatort auch erst nach Ablauf der Tat betritt und dann aus dem, was an Spuren vorhanden ist, die Tat zu rekonstruieren versucht. Dabei geht es um Spurenaufnahme und -dokumentation, Spurenaus- und Bewertung und deren möglichst schlüssige Interpretation. Ziel ist die Ermittlung empirischer Indizien und von Beweismaterial, um den Tathergang rekonstruieren zu können und einen möglichen Täter zu identifizieren.

Vielleicht nicht ganz so fernsehtauglich, aber doch in vielen Schritten recht ähnlich, stellt sich die Tätigkeit des Phylogenetikers dar. Auch hier spielen die Spuren der Geschichte - ihr Auffinden, Auswerten und Interpretieren - die zentrale Rolle in der „Ermittlungsarbeit“. Dabei können diese Spuren als bestimmte Typen eines allgemeiner gefassten Zeichenkonzepts verstanden werden.

In dem Vortrag wird der Versuch unternommen, Methoden und Konzepte einer von C. S. Peirce ausgehenden allgemeinen Zeichentheorie (Semiotik) auf die Phylogenetik anzuwenden. Dabei wird gezeigt, dass sowohl die Phylogenie selbst, als auch die phylogenetische Rekonstruktion, als Zeichenprozesse (Semiose) verstanden werden können. Diese Zeichenprozesse zeichnen sich durch Interpretationsfolgen aus, wobei jede Interpretation ein „neues“ Zeichen erzeugt. Die Interpretation evolutionärer Vorgänge und ihrer Rekonstruktion als Zeichenprozesse eröffnet eine neue Perspektive für die Bewertung allgemeiner Problemstellungen in der phylogenetischen Systematik.

Gefördert durch die Deutsche Forschungsgemeinschaft (Ba 1520/4-1).

Signs and Phylogeny

A Semiotic Approach to Systematics

L. Vogt

Zoomorphologie und Systematik, Fakultät für Biologie, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, E-Mail: lars.vogt@biologie.uni-bielefeld.de

eingereicht in:

Biology and Philosophy

Abstract

An introduction to the triadic conception of *sign* developed by C. S. Peirce and his semiotic epistemology is given, and their application on phylogenetic methodology is presented. Many semiotic processes go on at the level of individual organisms as well as populations and species. Some of the signs involved in these processes, the inheritable organismic traits, are not only signs *of* phylogeny but also putative signs *to* phylogeny. If they exhibit specific properties, they are indices to the phylogenetic past of the corresponding species. As I understand the concept of indexicality as the epistemological correspondent of ontological causality, the indexical strength of a phylogenetic sign depends on its 'causal connection' to a specific succession of evolutionary events. This 'causal connection' can best be measured by its corresponding process probabilities. A concept of the ideal sign to phylogeny is presented and its specific properties are taken as criteria to test real signs in respect to their similarity to this ideal. The first test, the *identity test*, is located within the character analysis, testing the identity of traits which are hypothesized to represent *replica* of the same *type* of phylogenetic sign. The second test, the *congruence test*, is located within the cladistic analysis, testing the congruent relation between different putative phylogenetic signs. Furthermore, the relation and agreement of the semiotic approach and falsificationism is shown and the semiotic process of inferring phylogenies is analysed. Here, the three logical steps of inference – *abducion, inducion* and *deducion* - which C. S. Peirce developed, are recognized as steps in phylogenetic research within the character analysis as well as within the cladistic analysis. Since phylogenetic research, like phylogeny itself, both represent semiotic processes, the knowledge of their general mechanisms results in qualitatively better phylogenetic hypotheses. Thus, there exists a reciprocal illumination between the knowledge of general phylogenetic mechanisms and the inference of specific singular phylogenetic events. Therefore, it is advisable to consider all available and sufficiently corroborated relevant knowledge in each step of phylogenetic inference. On this ground, a justification of the superiority of parsimony methods over likelihood methods - as some advocates of falsificationism in phylogenetic research claim, because parsimony would not consider process probabilities and would minimize the assumed background knowledge, is not tenable.

Keywords: background knowledge, corroboration, evolutionary process probabilities, falsification, Peirce, phylogeny, Popper, semiotics, sign, weighting

Introduction

The classification of organisms belongs to the oldest sciences and, thus, the application of theories from epistemology and philosophy within systematic research and methodology has a long lasting tradition. In recent years, especially the interpretation of Popper's falsificationism gave rise to controversial discussions (e.g. Farris, 1970; Bock, 1973; Platnick and Gaffney, 1977; Brower 2000; de Queiroz and Poe, 2001; Faith and Trueman, 2001; Kluge, 2001). In particular the role which assumptions on evolutionary process probabilities play within the weighting of cladistic characters, as well as the amount of relevant background knowledge that has to be assumed when inferring phylogenies, represent open questions.

Since organismic traits are interpreted in respect to their own history to give evidence to the phylogeny of the corresponding species, those traits function as signs to phylogeny. The first to represent a modern triadic conception of sign was C. S. Peirce, therewith founding a new field, called semiotics. The application of semiotic theory and biology is known as *biosemiotics*. While behavioural biology, neurology, cognition, and artificial life/intelligence represent major fields of biosemiotical research, the application of semiotic theory in evolutionary biology and especially in systematics is still poor.

Within this paper, I want to give an introduction to Peirce's conception of semiotic epistemology. A general application of his conception on different steps of phylogenetic inference is presented. Additionally, a semiotic perspective is given on the role of assumptions of evolutionary process probabilities and the amount of relevant background knowledge within systematics.

Peirce's Semiotic Epistemology

When Robinson Crusoe encounters footprints of 'Friday' on 'his' island for the first time, his attention is drawn to an unusual structure in the sand at the beach. While examining this structure, he recognizes the imprint of toes, of a heel and of all the other parts usually belonging to a human foot. By interpreting these findings, he concludes that there must have been and probable still is another human being on 'his' island (Defoe, 1719). This footprint represents a sign which is interpreted by Robinson Crusoe (see also Sebeok, 1990).

Another example is found in the ancient Greek anecdote of the philosopher Aristippos, a pupil of Sokrates (Cicero, 1995). After shipwrecked, he strands at a region unknown to him. He is desperate until he discovers geometrical figures in the sand, which he interprets as representing signs to the presence of other 'educated' humans.
It seems as if understanding the world depends on the ability to correctly interpret signs. Thus, these two examples give rise to the more general question of which role play signs in science and especially in phylogenetic research.

Since we have no possibility to directly get to know and understand the being and since we do not directly experience matter in its totality, we have to rely on perceptions. To be able to rely on perceptions, we have to reflect and interpret them, thus we have to have a consciousness of past perceptions and of a self or we should at least have established a specific habit that lets us react 'correctly' upon specific perceptions. Already at this basal level of cognition – close from the starting point of any epistemological process - signs play a fundamental role. Without signs we would have no knowledge.
Some signs mediate between matter and consciousness. As a mediator, they have to have a connection to matter as well as to our consciousness. Since this connection is not necessarily linearly causal, the interpretation of material signs and with it of our perceptions is not unproblematic and principally fallible. We have the concept of truth and falsehood, which we, in case of truth, use to express our believe in interpretations being consistent not only to our experiences, but also to reality, and which we, in case of falsehood, use when experiencing insuperable resistance against believing a specific statement. Thereby, we take advantage of methods from an epistemological logic, which help us to distinguish steps of interpretation, which are thruth-keeping and therefore unproblematic from steps which are not.

Firstness, Secondness and Thirdness

Charles S. Peirce was the first to start to work out a theory of signs (*semiotic*) which broke with the classical dualistic view of a sign, the representation theoretical model of a sign, consisting only of *representamen* and *object*, (*symbol* and *nominator*, *designat* and *denotat*, *significant* and *significat* respectively). Instead, he based his conception on a triadic relation, corresponding to his modification of Aristoteles' and Kant's apriorisms.

By asking and analyzing which are the necessary conditions for the possibility of consciousness, experience and thus of empirical knowledge in general, he reduced Kant's 12 categories (Kant, 1998) to three. Those three categories, which he calls *firstness*, *secondness* and *thirdness*, play a central role in most of Peirce's work, including his semiotic as well as his logical theory (*three-valued logic*, *relative terms*, *quantification theory* and *existential graphs*), which should be treated as a cohesive epistemological system, constituting his semiotic epistemology. Peirce spent most of his work on those three apriorisms, rethinking and developing them continuously. They are indispensably necessary for the understanding of his triadic conception of a sign. Here, I can only give a short description of them.

In the beginning Peirce distinguishes three independent ontological classes, which are operationally linked to another (Pape, 2000a). During his research he develops them to universal categories.

Firstness represents the conception of *qualia* in reference to *ground*. No object could be recognized, were there no quality which could be ascribed to the object. Thus, with such a quality we simultaneously refer to a ground – the fact, that such a property exists. The *first* is the term, to exist independent of anything else (Pape, 1991).

However, *qualia* presuppose that we generalize and that we distinguish objects which typify qualities. Thus, there is the need of the conception of *relation*, in reference to a *correlative*. This conception is called *secondness* (Peirce, 1866). The *second* is the term, to be in relation to something else, to react on something else (Pape, 1991). The introduction of the term *correlative* intents that we can, under reference to a *mediating representation*, refer to a plurality of objects as correlatives. This mediating representation has the *interpretant* as its ontological correspondent (Pape, 2000). The conception of a mediating representation is called *thirdness*. Thus, the *third* is the term of mediation, by which a first and a second are correlated (Pape, 1991).

Peirce's Conception of Sign

The combination and interaction of these three categories represent the object of semiotic inferences. It is this triadic process, that is called *semiosis*, by which a first determines a third to refer to a second, to which it refers to itself. Peirce took the term semiosis from the epicurean philosopher Philodemus, who called every conclusion from a sign a semiosis (Deledalle, 2000).

In correspondence with Peirce's (Peirce, 1903) three universal categories, every sign consists of three components: *object*, *representamen* and *interpretant* (Fig. 1). This conception of a sign ties up to the three universal categories in that way, that the *object* refers to the *ground* (*firstness*), the *representamen* to the *correlative* (*secondness*) and the *interpretant* to the *interpretant* (*thirdness*), constituting semiotics as a general epistemological logic.

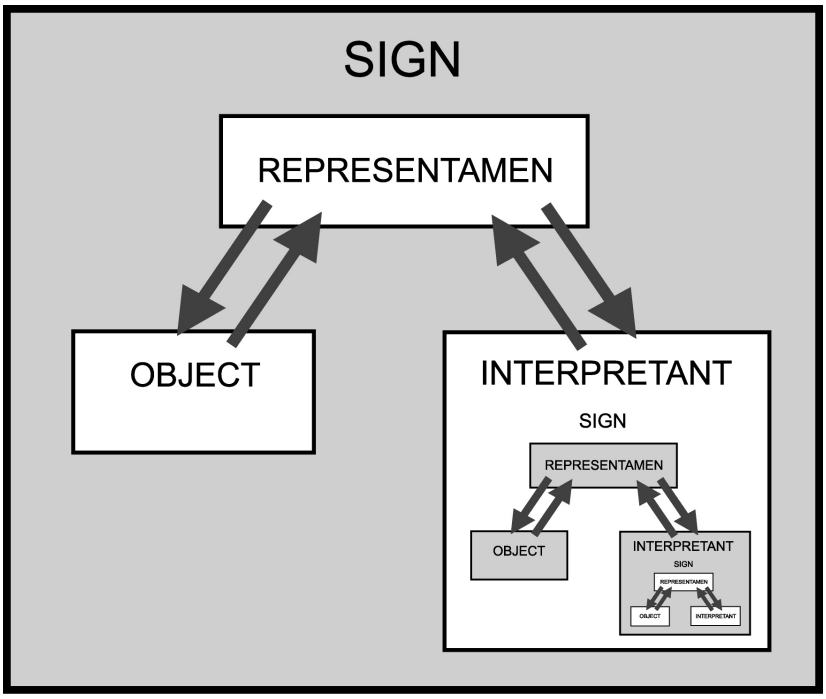

Fig. 1: A sign consists of three components: the representamen, the object and the interpretant. For more details see text.

A representamen is everything that relates to a second, which is called object, so that it is able to constitute a third, the interpretant, to stand in the same triadic relation to the relation to the object, as the representamen does. This means, that the interpretant is a sign itself - and this goes on *at infinitum*, leading to an interpretant and a sign that is 'self similar' (Fig. 1).

As the interpretant of a sign is a sign itself, interpreting a sign causes the formation of another sign in the mind of the interpreter, which itself is open to interpretation again, thus, constituting a potentially infinite process of interpretation. In this process, every sign refers to a precursor-sign by interpreting the precursor and, therewith, trying to

generate its interpretant as the representamen of the interpreting sign (Deledalle, 2000). Thus, an essential property of any sign is that it generates a sequence of signs of interpretation.

It is reasonable to distinguish between a single occurrence of a sign and its underlying rule. The occurrence is not independent of its rule. Therefore, one can differentiate the *replica* of a sign (also called *token*) as represented by a single occurrence of a sign, and its *type*, which refers to the rule. Thus, a sign does only exist through its replica and is therefore not a real thing itself (Peirce, 1904). And each replica refers to its corresponding type, as the word 'and' is printed several times within the issue you hold in your hands right now, each 'and' standing for one *replicate* of the same *type* of sign. Thus, one can follow, that if an interpretation of a sign generates the interpretant of the sign as the representamen of the new sign without modification, both signs (initial sign and its interpretation) would be identical and would thus represent replica of the same type.

Icon, Index and Symbol

Fundamental for the understanding of a sign is the relation between representamen and object and, with it, between representamen and interpretant. Peirce (1893) distinguishes three types or aspects of signs in respect to the nature of the relation of representamen to its object: icon, index and symbol.

An *icon* relates to an object by its own properties only - its *qualia* (*firstness*); whether an object exists or not. As a *potentiality*, the representamen is this quality only. Therefore, this relation is based on pure similarity, as e.g. a portrait relates to the person being portrayed. An icon can be a picture, a diagram, an analogy or a metaphor (Peirce, 1903).

An *index* possesses a physical and thus causal relationship to its object. In this *relation* (*secondness*), the object acts on the representamen, like e.g. a sun-dial is physically linked to daytime (Peirce, 1893), a symptom to its disease (Oehler, 2000) or a photography to a specific momentary view. Thus, the representamen stands for its object through a real *existing connection* to it (Peirce, 1895; Oehler 2000). Herein, an index always includes some kind of icon. The footprint of 'Friday', which Robinson Crusoe encounters at the beach, is a typical example of an index including an icon. The

imprint represents, by its iconic relation to the human foot, an index to the presence of a human being on the island.

A *symbol* relates to its object by referring to a convention or a *law*, *representing* it (*thirdness*), like e.g. any word and sentence. A symbol always includes an index and an icon.

Indexicality

The kernel of any empirical sign and the key to its understanding lies within the indexical relation of the representamen and its object. The mediating property of a sign depends on this indexical relation. Therein, the relation resembles the reversal of the causal connection of the object to its representamen. One could say that indexicality twists causality upside down (Sebeok, 1990). Therefore represent causality and indexicality two sides and possible perspectives for physical (and mental) processes - the former represents the ontological and the latter the epistemological side. Thus, the indexical relation of representamen and object of a sign refers to its complementary

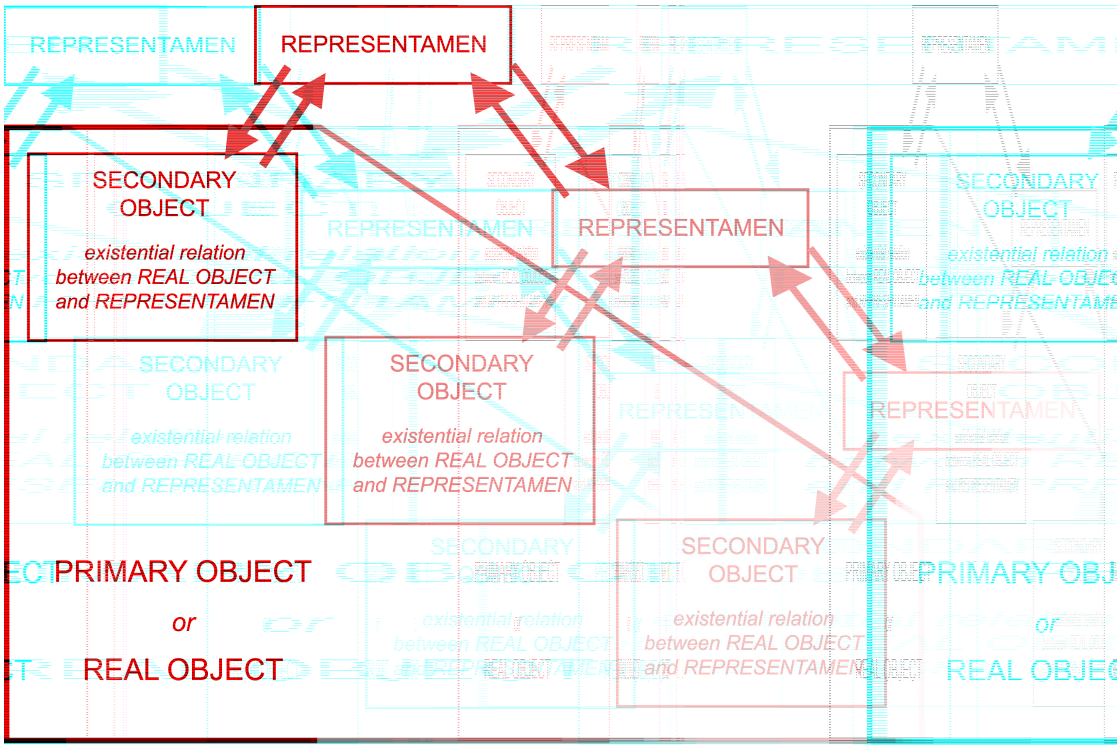

Fig. 2: The existential relation between the real (primary) object and the representamen is the object of the sign, as the interpretant represents it. Because of this, the real object, which is the correlative of this existential relation, is an object of the sign as well. Thus, the sign represents properties of the real object.

part, the causal relation of *before* and *after* of its corresponding physical or mental process (Peirce mentions (1903a), that the concept of the index and the concept of the past correspond with each other). This indexicality between object and representamen is projected onto the relation of representamen to its corresponding interpretant. That means, that the interpretant relates to the representamen by representing the real existential relation between the representamen and its real object. Therefore, since this represented existential relation is an object of the sign, the real object, which is the correlative of this relation, is an object of the sign as well. In this way does the representamen mediate between real object and interpretant (Fig. 2). Thus, it holds true that, what cannot be an object of the representamen, cannot be an object of the interpretant.

In this way do signs mediate between being and perception, between perceptions and perceptual judgments, between those judgments and thoughts, between thoughts and language, and thus between matter and consciousness. Therefore, one can conclude that causality and indexicality represent the two sides of this mediation.

The interpretation of a sign by a human being is restricted by what Eco (2000) calls limits of interpretation. It is the being that sets our liberty of the word its boundaries. There are *lines of resistance* (Eco, 2000) that are set by the being, and these lines restrict the semiosis of interpretation. Considering indexicality, the strength of the indexical relation of representamen and object determines those lines of resistance, thereby determining the limits of interpretation. By testing the indexical relation of representamen and object, the lines of resistance are evaluated and therewith the limits of interpretation. This is an ongoing process and represents necessary conditions for the possibility of cognition altogether.

In other words: Peirce calls the *real object*, or the *primary object* (Fig. 2), which can be understood as the real itself – Kant would probably call it the *noumenon*, the thing-as-such (Kant, 1998) -, also the *dynamical object*, because the epistemically decisive move of resistance – the lines of resistance - starts from it (Oehler, 2000). This *dynamical object* induces us, when confronted with it, to generate a representamen, which generates a *direct object*, which corresponds to the *secondary object* (Eco, 2000). The *dynamical object* persists as a thing-as-such, always present, but never penetrable, unless through semiosis (Eco, 2000). When considering natural signs (Emmeche, 1991), those limits of interpretation can be understood as being located within human cognitive dispositions as well as within the habits of those signs themselves show (induce). By

systematically covering the evaluation of the interpretational limits within a clearly phrased methodology, they are open to constructive scientific critique. One promising method for interpreting natural signs is represented by the recent improvements in the conceptualization of the principle of causality (Mackie, 1974, 1985; May, 1999; Spohn *et al*., 2001). This especially seems to be a fruitful approach due to the specific relation of indexicality and causality mentioned above.

Under the conception stated above, it is possible to infer specific causal relations by studying signs and referring to their indexical relations. And under this conception it is possible to reconstruct the past by looking for relevant signs, by investigating their corresponding icons and indices and by interpreting and evaluating them.

Semiotics and Falsificationism

The process of interpreting and evaluating an index belongs to the scientific field of testing and corroborating empirical hypotheses. To understand a material sign, as for instance an index of the type of the sun-dial or a symptom of a disease, and to interpret it correctly, the interpreter has to know the possible causing relations between object and representamen to understand their indexical relation. The interpreter, thus, relies on a law-like causal relation and, therewith, on some kind of background knowledge, when interpreting the material sign – she has to know how the shadow of the sun-dial is physically linked to daytime or the symptom to a specific disease to be able to interpret the corresponding material signs correctly. In this way, by relying on a law-like hypothesized relation between object and representamen, the interpretation of every material sign that is an index represents, at its turn, a *symbol* and, therefore, an *argument* of rational conclusion (either abduction, induction or deduction – three types of logical inference, which he formulated after reinterpreting Aristotelian logic; see Peirce 1866a, 1867, 1983, 1898, 1901).

Robinson Crusoe encounters some structures in the sand which call for his attention. By analyzing the structures he concludes, while considering all his experience (background knowledge), that they cannot represent the result of processes caused by weather and ocean alone. He recognizes the iconic relation (similarity) of the structures and his own foot, and concludes *abductively*, that the structure represents the imprint of a human foot. That abductive conclusion is a symbol – it refers to a law-like relation of object and representamen - and an interpretation of the material sign. In a next step of

interpretation, Robinson Crusoe gives an explanation for the occurrence of this imprint, by recalling that this imprint cannot stem from his own foot. He concludes again by abduction, that this imprint must have been caused by a human being who walked along this beach.

Interpretations, if possible, should be empirically tested to receive some kind of corroboration. This is done by deducing predictions – qualities and relations – that the representamen of the material sign has to show in case its interpretation is correct. While performing the test, one tests the interpretation against the lines of resistance to evaluate the borders of interpretation. This reveals the logical and methodological linkage of semiotics and a popperian falsificationist approach of testing empirical hypotheses by empirical observations (see also von Pückler, 2002).

Semiosis and Epistemology

The three categories mediate operationally between *being* (*'Sein'*, as pure conjunction of subject and predicate; Eco, 2000) and *substance* (as undetermined 'it'; Eco, 2000), which mark the margins in between which all epistemological processes take place (Pape, 2000). Peirce's three categories also built the basis for his *logic of relatives* (Peirce, 1897) with its three classes of relations: monadic, dyadic and triadic (one, two and three figure) relations (Oehler, 2000). As such they represent universal categories and at the same time three necessary elements of thinking and rationality (Peirce, 1903a), that constitute the basis for Peirce's conception of signs as well as for his conception of an epistemological logic. Peirce defines logic as the science of the conditions which capacitate signs to refer to objects (Peirce, 1865). Thus, why are signs essential for cognition?

A very basal sign in the process of perception (Roesler, 2000) could be understood as follows: The object is the causing stimulus in a sense organ, which is represented by its percept, and the interpretant is the judgment of the perception, which leads by its interpretation to the generation of another sign, the perceptual statement (Roesler, 2000). This process goes on and on, and one can conclude, that thinking is basically nothing else but a semiosis. Therefore, everything we know are signs. This holds true also for science. Scientific inference also represents 'nothing' but a specific kind of semiosis. And any scientific hypothesis can be interpreted as a sign which needs severe testing of its mediating properties.

The application of a semiotic conception, though, does not necessarily have to be that abstract. In many cases it provides a suitable basis for investigating and solving concrete empirical problems. I want to investigate, in which way the epistemological conception of semiotics, the science of semiosis, is applicable on phylogenetic research, and I want to examine its logical conclusions for the theory and methodology of systematics. Furthermore, I will demonstrate how a semiotic and a refutationist approach to science can be reasonably combined in cladistic research. Especially the question of what belongs to the relevant background knowledge of inference and of testing hypotheses and which part do considerations of process probabilities take within the whole procedure can be answered from this new perspective. Those two questions build the source for an ongoing and current discussion about the implementation of a popperian falsificationist approach to phylogenetic inference and the choice of either Maximum Parsimony or Maximum Likelihood as the best method of cladistic analysis, that is presently available (e. g., Farris, 1970, 2000; Platnick and Gaffney, 1977; Brower 2000; de Queiroz and Poe, 2001; Kluge, 2001).

Semiosis and Phylogeny

The application of semiotic theory in biology represents a lively field of research, called *biosemiotics*, which especially focuses on behavioral studies, cognition and artificial life/intelligence (e.g. Emmeche, 1991; Hoffmeyer, 1998; Brauckmann, 1999; Brier, 1999; Brogaard, 1999; Kawade, 1999; Laubichler, 1999; Stjernfelt, 1999; von Uexküll, 1999; Kull, 2000). An interpretation of semiotic theory within phylogenetic methodology is still needed.

Phylogeny is basically historical research with methods from natural science. Though the history of single individuals is in parts directly observable, reconstructing the evolutionary relationships of groups of organisms or species is not directly possible. Therefore, nothing but indices, traces of history, remain as empirical evidence for testing cladistic hypotheses. Thus, a student of phylogeny is like a detective at the site of crime trying to correctly recognize and interpret the traces, clues and hints which are left behind by the relevant evolutionary events. To be phylogenetically informative, those traces must be representations of material change in evolutionary lines of

organisms, which could by heredity pertain their structure until nowadays. How is this translated into terms of semiotics?

A semiotic process does not necessarily premise a living interpreter. A billiard ball rolling with a certain speed, transfers its energy to another billiard ball by hitting it. This process of transference can be understood as a semiosis, in which the second ball interprets the precursor-sign in respect to its movement-energy. If this process runs under ideal conditions, the second ball takes over the entire movement-energy from the first without any loss and, thus, interprets this material sign by completely 'generating' its interpretant as the interpreting sign in the moment of the collision.

Thus, it is possible to distinguish at least three classes of semiotic processes.

A process of semiosis, in which only material signs are involved – the sign that is interpreted as well as the interpreter itself. The collision of the two billiard balls represents an example of such a process.

A process of semiosis, in which material as well as immaterial signs are involved; e.g. a physicist consciously percepts the process of the collision of the two billiard balls.

A process of semiosis, in which only immaterial signs are involved; e.g. the physicist interprets his perceptions of the collision in the light of the mechanic theories known to him.

The processes of all three classes have in common, that the interpretation of a sign may generate the interpretant of the 'template' with or without change as the representamen of the interpretation-sign. Therefore, with the process of interpretation, the emerging signs may differ from their 'template'-signs, thus, constituting a process of evolution of signs. Phylogeny can be viewed as a special case of such an evolution of natural material signs.

Therefore, to be able to correctly reconstruct this semiotic process of phylogeny, one first has to understand it. One should investigate and recognize its possible types of processes and their corresponding products, to be able to trace back the to us observable products of phylogeny – the distribution of identical traits – to their putatively causing events; processes, that took place sometime back in the corresponding line of descent.

Thus, to be able to develop good hypotheses of the indexical properties of observed material signs and therefore to be able to interpret them, one has to know their causal relations. This elucidates the relation of reciprocal illumination in phylogenetic research. The better we understand the general processes involved in phylogeny, the

better we reconstruct the actual phylogeny and thereby the better we explain the distribution of identical traits among members of different organisms and species, and vice versa (Fig. 3). The growing knowledge of general processes of phylogeny and insights into their mechanisms provide a growing set of possible types of processes and their corresponding results and probabilities, which we can utilize when hypothesizing a specific event to explain our actual experiences. The larger this knowledge, the smaller the amount of misinterpretations and of wrong hypotheses.

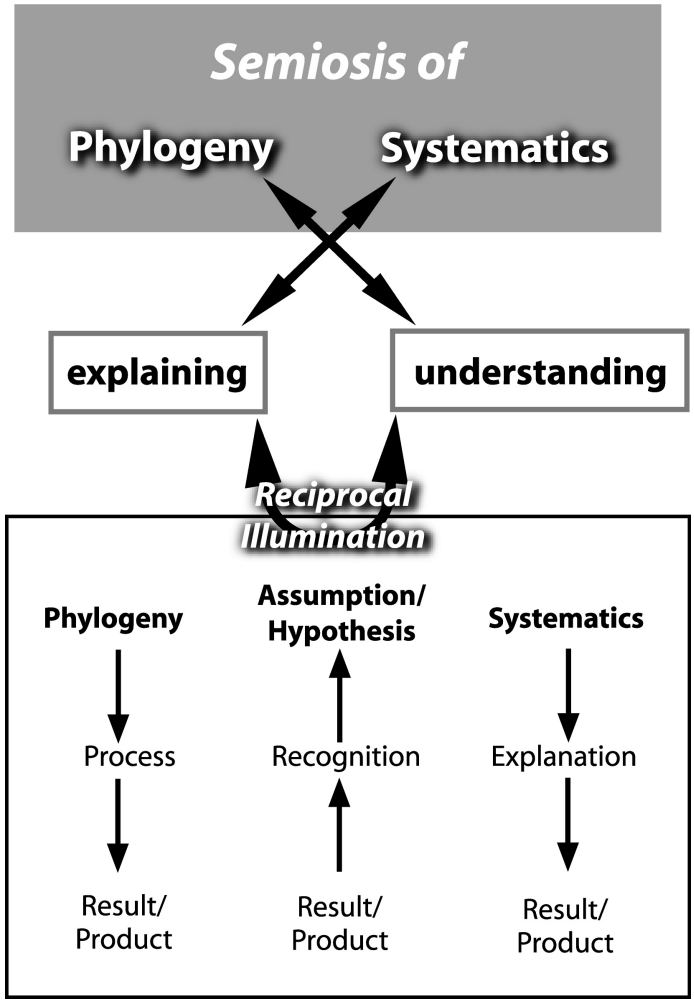

Fig. 3: The semiosis of systematics is the attempt to reconstruct the semiosis of phylogeny. The quality of the interpretation of the signs of phylogeny depends on the knowledge we have of the general processes of phylogeny and their corresponding products. The better our understanding of those processes, the better is our explanation of the results and thus the better is our attempt to reconstruct phylogeny (and vice versa).

And both processes, phylogeny and systematics, resemble semiotic processes. The semiosis of phylogenetic research is concerned with the correct interpretation of the material signs *of* phylogeny, to convert them – the phylogenetically informative signs – into signs *to* phylogeny. Those signs, which are, as interpretations of material signs, symbols, serve as arguments and, thus, as empirical evidence for the reconstruction of the corresponding relationships and the evolutionary past of species.

Semiosis at the Level of Individual Organisms

Phylogenetic processes take place within individual organisms and in between them and their environment. Whenever a cell cleavage takes place, the genome of the precursor-cell, and with it its sequences and genes, are copied and the resulting sequences are in most of the cases identical to their precursors. Thus, as long as no error occurs during the replication and cell cleavage, the 'new' sequence represents a true replicate of its corresponding precursor. If one understands the genome or a sequence of a genome as a sign, the precursor as well as the 'new' genome represent two *replica* of the same *type*. The precursor is interpreted during the process of DNA replication by the resulting 'new' sequence, and the interpretant of the former constitutes the representamen of the latter. By exhibiting a strong iconic relation to its precursor, the resulting sequence indicates a strong indexical relation to its object.

Therefore, whenever information is transferred and structures are copied, the corresponding process can be referred to as a semiotic process. Thus, whenever parents reproduce and whenever traits are inherited, it is a phylogenetic semiosis that takes place.

However, since reproduction and cell cleavage are processes of material entities, and since those processes are not free of 'errors', the results are influenced by effects of chance. Not always within the phylogenetic semiosis is the resulting sign identical to its precursor. In these cases the interpretation changes the sign. The resulting sign then represents the first *replicate* of a new *type*. This happens, whenever a mutation or recombination takes place and whenever a transformation of a trait occurs. Therefore, a phylogenetic semiotic *type* can be understood as a general type of trait – a class of necessarily identical traits. Talking about a type in systematics makes only sense as long as there exists a mechanism which somehow guarantees that a replicate of a type generates a replicate of the same type and DNA replication represents such a mechanism. By referring to the distribution pattern of identical traits among individual organisms, one has in principle an index to the classification of all organisms in respect to the history of the corresponding phylogenetic semiosis.

Semiosis at the Level of Populations and Species

In case of sexually reproducing organisms, besides cell cleavage, there is also the possibility of the fusion of two haploid cells bearing different genomes. These genomes are combined to a new resulting genome. Moreover, recombination, speciation and other processes may occur. All those processes can be interpreted in terms of semiotic.
By referring to the semiosis discussed above, one should also be able to classify populations and species in respect to the history of their phylogenetic semiosis and therefore to their phylogeny respectively.

As a consequence, identical traits represent indices to systematics – the distribution of them serves as an analytic sign for the reconstruction of the phylogeny of the corresponding organisms and species.
However, this is only possible as long as *replica* are to the scientist identifiable as belonging to a specific *type* and as long as *replica* of different *types* can be distinguished.
Therefore, a phylogenetic sign has to have three specific properties. On one side it has to operate as an index to phylogeny as described above. On the other side it has to affect the attention of an observer and has to enable the correct construction of its interpretant. From this perspective, reconstructing phylogenies can be referred to as a semiotic problem – the question of identifying and correctly interpreting the phylogenetically relevant signs. Knowing which types of properties of organisms are linked to their phylogeny and how they are linked would give the basis for developing a methodology for identifying and interpreting these indexical signs correctly; thus enabling one to perform empirical tests on cladistic hypotheses as severe as possible.

Phylogenetic Signs

A cladistic hypothesis is a hypothesis of a succession of speciation events which gave rise to the to us observable species. If one wants to infer this succession to receive the evolutionary relationships of the investigated organisms, one has to take the linear structure of time into account, thus its irreversibility. In doing so and considering four types of evolutionary events – heredity, reproduction, transformation and speciation -, one can deductively predict the existence of a specific type of distribution pattern of

identical and different organismic traits among individual organisms and species. While this pattern has universal properties, the actual structure of such a distribution depends directly on the real historical succession of those four types of events and their specificity.
This dependence is utilized in phylogenetic analyses, because it represents the indexical part of phylogenetic signs. The universal properties at their turn are employed to test the hypotheses of synapomorphy empirically.

The Ideal Phylogenetic Sign

An ideal phylogenetic sign would give definite, unambiguous evidence for a specific relationship, when interpreted correctly. Of course, such a sign probably does not exist, and even if such a sign would exists, we would not know and recognize it – it represents an ideal. However, we can construct such a sign in theory and investigate and deduce its properties for testing the existing signs and their similarity to the ideal sign.

Trait *X* has to meet the following conditions to be considered an ideal sign to the monophyly of the species whose members carry this trait:

1. *X* is a specific type of trait that is recognizable to an observer and distinguishable from any other trait '*non-X*'.
2. *X* can *only* be caused by one type of transformation event that happens only once in evolution and will never happen again.
3. A subsequent alteration of *X* – either modification or reduction - is not possible.

X would represent a synapomorphy for all species possessing *X* and would be an ideal index to the monophyly of these species, because its indexical function is strict and unambiguous, and so is its interpretation to the observer.

The Semiosis of Phylogenetic Research

Phylogeny is an ongoing semiosis and so is phylogenetic research.
Phylogenetic research is the attempt to reconstruct the semiosis of phylogeny by interpreting the material signs that resulted from this semiotic process. Those immaterial interpretations of material phylogenetic signs represent process types of

class 2 and 3 (see chapter *Semiosis and Phylogeny*) and are, thus, conclusions, which represent symbols of the type of an *argument* (Peirce, 1866, 1905). Argumentation, in general, is the formulation of a judgment (Peirce, 1895), and an argument is a sign by itself, that explicitly represents the interpretant (conclusion), which to determine it was intended (Peirce, 1901). A possible syntax of an argument consists of the premise, the fact and the conclusion. Those three elements are accomplished semantically by observation/evidence, hypothesis and theory. In dependence on how the interpretation interprets the relation of representamen and object – as *iconic*, *indexical* or *symbolic* -, there are three different types of arguments: *abduction*, *induction* and *deduction* (Peirce, 1901).

With *abduction* Peirce made a logical discovery (Zeidler, 2000) – deduction and induction alone cannot provide a self-contained logical basis for a theory of semiosis. The subsumption of a 'particular' under a 'universal' does not only premise the inductive inference of the universal, but also the representational determination of the particular. The abduction is a logic synthesis, which infers the instance to the rule and therefore the applicability of the rule – it formulates the act of representational determination and of the formation of terms (Zeidler, 2000). While deduction and induction split the logical synthesis in the two distinct aspects of a merely formal and a merely empirical statement and rational, abduction mediates both scopes by achieving the representational identification of the object, giving the deduction its logical objects and induction its empirical representations (Zeidler, 2000). Thus, with abduction, logic combines what is empirically appropriate (correspondence theory), formally correct (coherence theory) and inter-subjectively valid (consensus theory) (Zeidler, 2000).

Phylogenetic Arguments and Character Analysis

Any attempt to reconstruct the phylogeny of a group of species begins with the character analysis (Fig. 4). Within the character analysis, individual organisms are studied and perceptions of their specific organismic properties serve as basal empirical evidence. The first semiotic step in any phylogenetic analysis, and thus, the initial representamen of a beginning semiosis of phylogenetic inference is represented by the empirical perception of an organismic property. Thereby does the perception represent an interpretation of the reaction of our senses to the confrontation with a specific feature

of an organism. Often, this step - I call it *abductive construction* -, passes on unconsciously (ratiomorph *sensu* Riedl, 2000), resulting in a *perceptual judgment*. Nevertheless, it is possible to analyze this *abductive construction* a posteriori, and Peirce and his semiotics understand perception therefore as a form of conclusion (Eco, 2000). Each perceptual judgment represents a sign and stands as a single replicate of a specific type. The replica of the study of several individuals of different organisms are compared to one another and identical replica are subsumed to one class of the same type, which results in the formulation of a *character hypothesis*. Character hypotheses at their turn constitute a universe/domain of discourse and represent empirical experience. In character analysis, this argument of *abductive construction* has the form:

Observation:	Several replica (perceptual judgments) stand in an *iconic* (*firstness*) relation to one another, i.e. they cannot be distinguished from another.
Theory:	There exists a natural representational class, i.e. a *type*, whose *replica* are identical with the observed *replica*.

Conclusion

Hypothesis:	The observed replica are replica of the same type, i.e. they are elements of a common conceptual class of a predicate.

Thus, this step conceptualizes identical perceptions of parts of organisms to identical properties of organisms, thereby assigning those organisms a predicate. The basis for this step is the empirical evidence together with the comparative relation of similarity – analogy.

Within the next step (logical and not chronological step) of the character analysis, our *abductively constructed* empirical experience, which is a set of distribution patterns of identical traits among members of different species, needs scientific explanation: How does it come that we perceive traits of members of different species that are identical?
It is always experience which poses problems. And, actually, this is the main task science has to perform: It has to give convincing explanations for problems emerging from our experience. And those explanations should be causal to be scientific explanations. As, for instance, a statement like “When the barometer falls, the weather will get worse” does not represent a causal or *mechanismic explanation* (*sensu* Mahner

and Bunge, 1997), but a conditional phrase, which rather describes than explains an empirical fact. Therefore, an explanation is a statement, that, by referring to a known process, gives a mechanismic and thus causal explanation for our experience in form of a hypothesis (this is due to the specific relation of indexicality and causality – see chapter *Indexicality*).

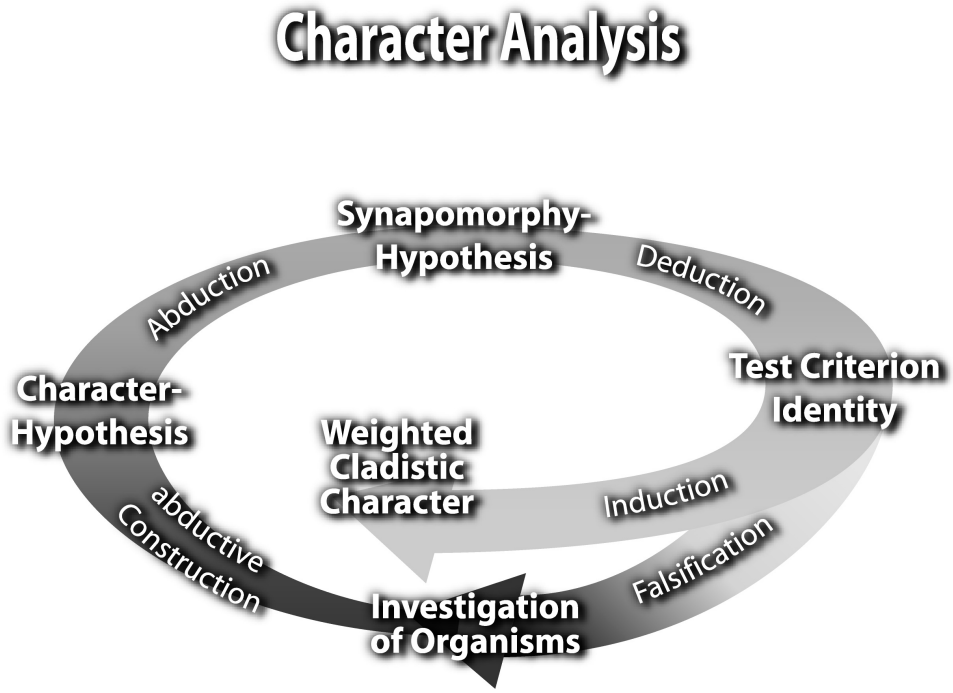

Fig. 4: Cycle of semiotic steps of argumentation within the process of character analysis. Starting point are the perceptions of organismic properties. By an abductive step a term is referred to each perception, constituting a character hypothesis. This character hypothesis is explained by another abductive step, which puts a known process in relation to the character, resulting in a hypothesis of synapomorphy. From the hypothesis of synapomorphy properties are deduced, which a synapomorphy bears in general, constituting a theory consisting of the hypothesis of synapomorphy as well as of the predicted perceptions, which necessarily belong to the character, if it really resembles a synapomorphy. These prediction are tested against new perceptions of the relevant organisms and against all we know of them by now. If this test is failed, the complete hypothesis is falsified and the inference has to start from new. If this test is passed successfully, the whole cycle, and especially the two steps of abduction, have to be evaluated and their explanatory power has to be estimated within a step of induction, to justify and substantiate the hypothesis.

Therefore, in our case of an identical trait among different species, one possible mechanismic explanation would be, that there happened a transformation process in a species some time back in the evolutionary past, resulting in that trait (property); and subsequently a specific number of speciation events took place, leading to identical traits among members of different species. Moreover, if one can determine the type of transformation process which caused this property, it is even possible to evaluate the explanatory power of this explanation in comparison to other alternative hypotheses

(see inductive step, below). The result of this step is the formulation of a *hypothesis of synapomorphy*. And, again, the conclusion is gained *abductively*, by referring to known processes and *iconicaly* (*firstness*) relating their resulting product with the observed product (on the basis of analogy):

Observation:	Several organisms exhibit an indistinguishable property *X*.
Theory:	There exists an evolutionary type of process *Y*, that causes such a heritable property *X*. Speciation events can result in the distribution of heritable properties of a common ancestor among the members of the resulting species.

Conclusion

Hypothesis:	The observed identical property *X* represents a property, which was caused by *Y* in the common ancestor of the species exhibiting *X*. And since the transformation process took place, speciation events must have been occurred.

The result of the abductively inferred mechanismic explanation of the perceptual judgment (character hypothesis) is represented by a hypothesis of synapomorphy.

In terms of semiotics, this abductive step is therewith founded, that an observed property of an organism, in its function as representamen, relates to a specific class of objects and not necessarily only to a single individual object. This corresponds to the fact that a representamen relates only to a secondary object (direct object), which is an existential relation between the real object (dynamical object) and the representamen (see Fig. 2). Though it relates in this way to the real object as well, it does so only in terms of its existential relation to the representamen. This means, that the representamen can only represent those properties of the real object, that affect the representamen. Therefore does a specific trait represent a class of one or more corresponding types of causing processes. This is because the indexical relationship between object and representamen is not that strict that the interpretation of the representamen is absolutely unambiguous (see *Indexicality* and the limits of interpretation). Following the methodological criterion of parsimony, identical traits are interpreted as evidence for the common descent of all organisms that bear those identical properties. Thus, identical

representamen of different organisms belong to the same single symbolic sign and their iconic relation is explained by abductively hypothesizing an identical indexical relation of the specific representamen and their corresponding objects. From this perspective, it is obvious that the iconic relations among traits serve as the empirical basis to estimate their indexical properties to their evolutionary past and the past of the corresponding lines of descent.

From the resulting hypothesis of synapomorphy, in reference to universal properties of synapomorphies, we can deduce empirical predictions, which should be fulfilled, if the hypothesis were true. The predictions represent an instance of empirically testing the hypothesis. Since this step of inference refers to a *law-like relation* (*thirdness*), it resembles a classical logical step, a *deduction*. It has the form:

Theory:	If after a transformation event only speciation events take place and the product of this initial transformation event is not subsequently reduced or modified, then all descendants *Y* will bear this property and this property will be identical among them. Therefore, it would resemble an epistemologically ideal synapomorphy, an ideal phylogenetic sign.
Hypothesis:	Property *X* represents such an ideal synapomorphy.

Conclusion

Observation/Prediction:	Property *X* is identical among all descendants of *Y*.

At this point, it is possible to test our past and future relevant experience against this *test criterion of identity* (Fig. 7) (Vogt, 2002). The test, and with it the falsification or corroboration of the hypothesis, represents the next step of inference. Here, the other steps are evaluated and additional empirical evidence is considered. It also resembles a step of justification of the hypothesis. Semiotically speaking, it is the step of the evaluation of the *indexical strength* of a phylogenetic sign, and it has the form:

1. In the case of corroboration

Observation:	Those replica, which represent property *X*, are iconically related to one another, i.e. they cannot be distinguished.

Theory: A synapomorphic property *X* is identical among all organisms that bear *X*.

Conclusion

Hypothesis: The identical replica '*X*' represent a putative synapomorphy.

2. In the case of falsification

Observation: Those replica, which represent property *X*, are *not* iconically related to one another, i.e. they *can* be distinguished.

Theory: A property *X* cannot be a recognizable synapomorphy, if it is not identical among all organisms bearing *X*.

Conclusion

Hypothesis: The non-identical replica do *not* represent a putative synapomorphy.

This result corresponds with a popperian approach to phylogenetic inference (Vogt, 2002).

In case of a falsification, the cycle has to begin from the starting point again. In case it passes the test successfully, the hypothesis gained corroboration and the result is a weighted cladistic character. And since a transformation process involves the substrate as well as the product, the hypothesis of synapomorphy should include both organismic conditions, that represent instances of the transformation, thus the condition *before* and the condition *after* the transformation event. Therefore, a 'complete' explanation of the observed distribution pattern, and a 'complete' hypothesis of synapomorphy, consists of the *apomorphic* condition as well as the *plesiomorphic* condition, if it should represent the conclusion of an epistemologically well substantiated argumentation (Vogt, 2002), and a complete historical explanation of the empirical evidence.

As the result of the character analyses, one receives weighted cladistic characters, since the explanatory power of the different hypotheses of synapomorphy vary. And those weights correspond with the indexical strengths of the corresponding phylogenetic signs. As the whole process resembles a semiotic process, one receives a *cladistic sign* (Fig. 5), which consists of an *object*, that is the transformation event and its specific number of succeeding speciation events within the corresponding line of descent; a

representamen, that is the cladistic character, consisting of two character states *X* and *Y*, which represent the two conditions before and after the transformation, and their distribution pattern; and an *interpretant*, which is the weighted hypothesis of synapomorphy for the character state *Y* and, therefore, the hypothesis of monophyly of all species, whose members bear that character state *Y*.

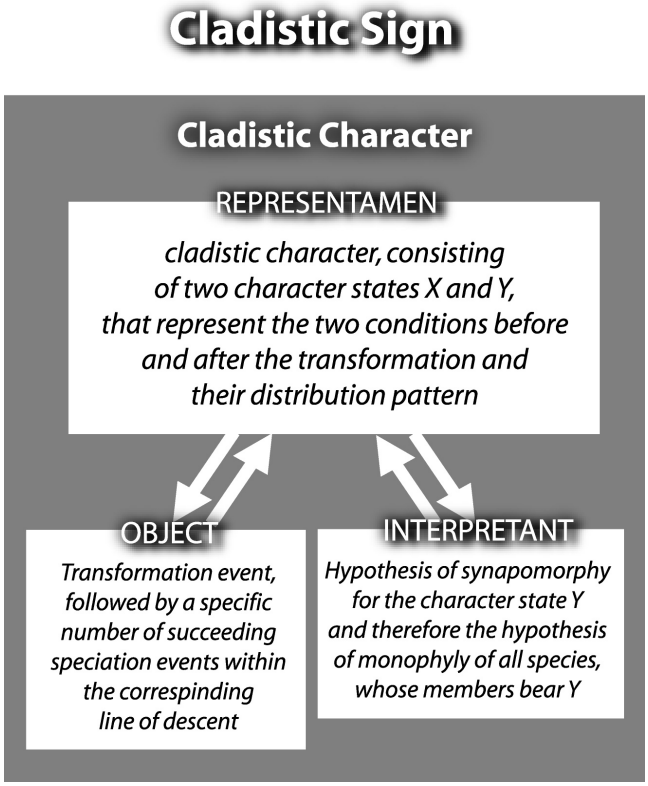

Fig. 5: The result of the character analysis is the weighted cladistic character – a cladistic sign, which consists of an object, a representamen and an interpretant. For more details see text.

Character Analysis – Testing against the Ideal Phylogenetic Sign

The testing and evaluating of the indexical strength of the interpretations of phylogenetic signs can also be viewed as testing the signs in respect to *the ideal phylogenetic sign* (see above). The better the sign matches the ideal sign, the stronger its indexical property for reconstructing phylogeny.

Looking at the properties of an ideal phylogenetic sign, one notes how they are related to the different steps of inference of the character analysis:

1. Property: *X* is a specific type of trait that is recognizable to an observer and distinguishable from any other trait '*non-X*'.

This refers to the step of *abductive construction*, which results in the perceptual judgment (*empirical experience*) of a character hypothesis. In case of the ideal phylogenetic sign, a predicate can unambiguously be applied on a specific property. Since this step deals with our sensational abilities to distinguish objects belonging to different natural classes, all the knowledge about possible artifacts in the production of the relevant objects (microscopy, PCR-methods, DNA-extraction, histology etc), knowledge of weaknesses of our sensational apparatus and of problems of comparison and description should be taken into account, when evaluating the reliability of this step.

2. Property: *X* can *only* be caused by one type of transformation event (…).

This refers to the step of *abduction*, which results in a *mechanismic explanation* in the form of a hypothesis of synapomorphy. In case of the ideal phylogenetic sign, a property can unambiguously be referred to a specific type of process of transformation. If we know (from e.g. experiments in the lab), that a specific type of trait can result from different types of transformational processes, the actual sign is, in respect to this specific property, not much alike an ideal phylogenetic sign.

3. Property: (…) that happens only once in evolution and will never happen again.

This refers to the step of *induction*, which results in an *evaluation* and *justification* of the hypothesis. In case of the ideal sign, the process probability of the transformation is extremely low and converges towards zero. Therefore, the higher the process probability of the actual transformation, the lesser resembles the observed sign an ideal phylogenetic sign. One can conclude the following relation:

The more probable the specific transformation process, the lesser ideal is the putative phylogenetic sign and the smaller is the cladistic weight, that should be given to this specific trait (Vogt, 2001a; contradicting Kluge, 1997, 1997a).

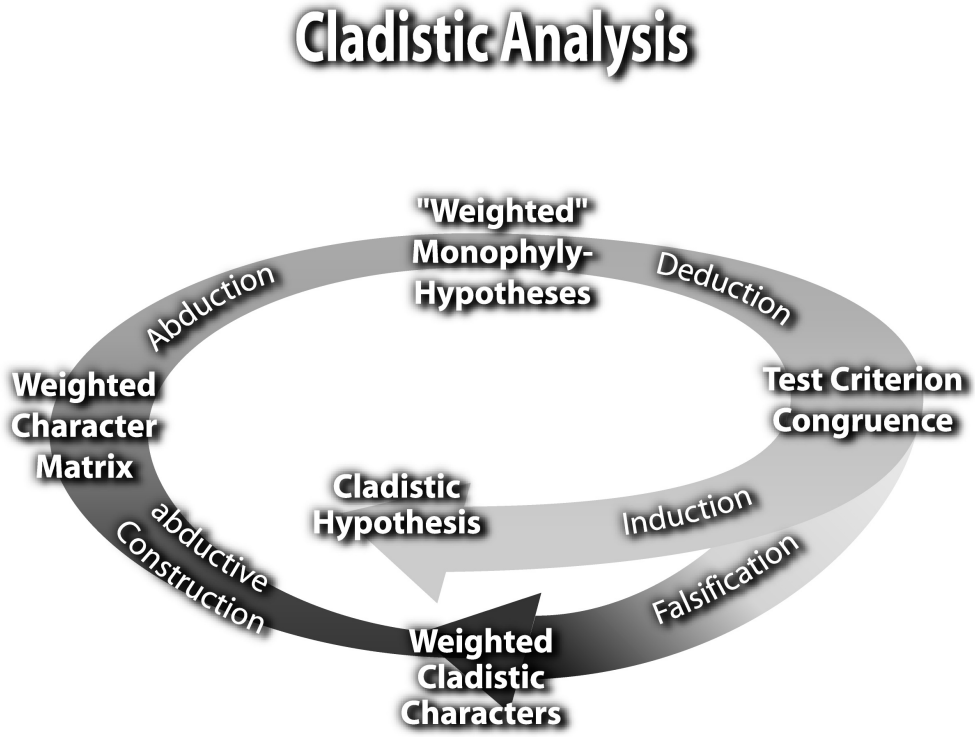

Fig. 6: Cycle of semiotic steps of argumentation within the process of cladistic analysis. Starting point is a set of weighted cladistic characters. By an abductive step an abstract term is referred to each character state, constituting a weighted character matrix. This character matrix is explained by another abductive step, resulting in a set of weighted hypotheses of monophyly. From the hypotheses of monophyly properties are deduced, which a monophylum bears in general, constituting a theory consisting of the hypothesis of monophyly as well as of the predicted perceptions, which necessarily belong to them, if it really resembles a monophylum. These predictions – the congruent distribution of synapomorphic characters within monophyletic groups - are tested against the empirical evidence. If this test is failed, the complete set of hypotheses(the set of weighted cladistic characters) is falsified and the inference has to start from new. If this test is passed successfully, the whole cycle, and especially the two steps of abduction, have to be evaluated and their explanatory power has to by estimated within a step of induction, to justify and substantiate the resulting cladistic hypothesis.

Correspondingly the same applies to the cladistic analysis (Fig. 6). Again, four steps of argumentation can be differentiated within this cycle. Starting point is a set of weighted cladistic characters. A step of abductive construction transforms the set of weighted cladistic characters into a weighted character matrix. This step resembles an abstraction, by which the results of the character analysis, which are primarily results in form of descriptions, are translated into an operationally accessible ‘digitalized’ form. The character matrix resembles a representation of our experience by summarizing the results of the character analysis and, thus, needs explanation. Especially the distribution pattern of traits, rather than the fact that they are identical, has to be explained. This

explanation is given in form of a set of (weighted) hypotheses of monophyly. As we know that synapomorphies of (monophyletic) groups exhibit a distribution pattern that is congruent to those groups and, thus, that this pattern features a hierarchical structure coding for nested groups of species, we can deduce from the set of hypotheses of monophyly the *test criterion of congruence* (Fig. 7) (Patterson, 1988; de Pinna, 1991; Kluge, 1997; Vogt, 2002; *sensu* criterion of coincidence of Wagner, 1986). This deductive argument has the form:

Theory:	Synapomorphies code for consistently nested groups of taxa (monophyla).
Hypothesis:	The weighted cladistic characters represent synapomorphies and, thus, code for monophyletic taxa.

Conclusion

Prediction:	The synapomorphies code for congruent groups of taxa.

Fig. 7: An index to phylogeny has to have two universal properties against which hypotheses of synapomorphy can be tested. These are: *synapomorphic character states have to be identical* and *the distribution pattern of synapomorphic character states has to constitute congruent, thus nested groups of species*.

If the set of hypotheses fails this test, because the distribution patterns contradict the requirement of a nested hierarchy, we can conclude, that not all groups represent monophyla. At least one of them is not a monophylum and thus, the set of hypotheses of

monophyly is falsified and the whole cycle of cladistic analysis has to start again, until a set of hypotheses of monophyly is found, that passes the congruence test successfully (Vogt, 2002, 2002a). If this test is passed successfully, the whole cycle, and especially the two steps of abduction, have to be evaluated and their explanatory power has to by estimated within a step of induction, to justify and substantiate the resulting cladistic hypothesis against other alternative hypotheses (see Vogt, 2002a).

Process Probabilities and Background Knowledge

A key problem to cladistic research is the question of how to weight cladistic characters (e.g. Farris, 1969; Neff, 1986; Bryant, 1989; Goloboff, 1993; Chippindale and Wiens, 1994; Allard and Carpenter, 1996; Milinkovitch *et al.*, 1996; Kluge, 1997a; Haszprunar, 1998; Trueman, 1998; Wiens, 2001). Within a semiotic approach to phylogenetic research, this can be interpreted as the question of evaluating the indexical strength of phylogenetic signs. Since interpretations of material signs always represent hypotheses, it is also the question of how severe those interpretations have been empirically tested against their *lines of resistance*.

Since the succession of speciation events is not directly observable, one has to refer to the distribution of the results of different transformation events to conclude from this distribution the evolutionary relationships. Therefore, one should consider the different types of transformation events which caused the identical and different traits when evaluating the severity of each identity test. Thus, testing against the lines of resistance resembles a test against the universal properties of ideal phylogenetic signs, which are empirically appraisable (Fig. 7).

Therefore, considering the ideal sign to phylogeny together with what we know from mutational mechanisms of evolution in general, one can conclude the following:

1. The result of a transformation can be traced back to a single specific type of transformational mechanism or to several alternative mechanisms. The more ambiguous the reference from the result to a specific causing mechanism is, the weaker is the indexical strength of the representamen to its object.
2. The probability of the transformation process stands in direct connection to the indexical strength of its corresponding representamen. The more likely a specific transformation event, the

weaker is the indexical strength of its corresponding representamen.

Considering these conclusions one expects that the quality and thus the phylogenetic information content of phylogenetic signs strongly depends on the specific probabilities of their originating processes and the ability of the observer to ascribe an observed trait to its corresponding type of transformational process. This result is, in principle, consistent to a falsificationist approach to phylogenetic inference (Vogt, 2002, 2002a), though it contradicts the ordinary cladistic interpretation (Kluge, 1997, 1997a), by recognizing process probabilities as the major key to weighting characters (see Vogt, 2002a).

The question of what belongs to the relevant background knowledge in phylogenetic research is part of an ongoing discussion, which especially concerns authors that focus their attention towards a falsificationist approach to phylogenetic inference (e.g. Kluge, 1997, 1997a; Brower, 2000). Within this paper I want to consider this question from a semiotic perspective.

To be able to use phylogenetic signs in cladistic research, one has to refer to some background knowledge which one necessarily has to assume to be able to interpret the signs as good as possible. This background knowledge must include everything that is necessary for the conception of a phylogenetic sign, including its iconic and indexical part as well as its interpretation. To begin with, and somehow self-evident, there is the requirement of a concept that explains what phylogeny is, before one can start to investigate what could serve as an index to a concrete phylogeny of specific species.

When considering the character analysis and the cladistic analysis it is evident, that semiotic inferences (and therefore scientific inferences in general) pass in cycles (see also Kluge's research cycles; Kluge, 1997a). Such a cycle (Fig. 8) includes aspects of description (*empirical*) as well as aspects of explanation (*rational*). At least four different steps of inference can be distinguished, which correlate (modified from Zeidler, 2000):

1. Perceptions/predictions with experience: step of *abductive construction*.

2. Experience with hypothesis: step of *abduction*.
3. Hypothesis with theory: step of *deduction*.
4. Theory with perceptions/predictions: step of *induction*.

All those steps rest, in respect to semiotic methodology, on Peirce's *apriorisms* of *firstness*, *secondness* and *thirdness*.

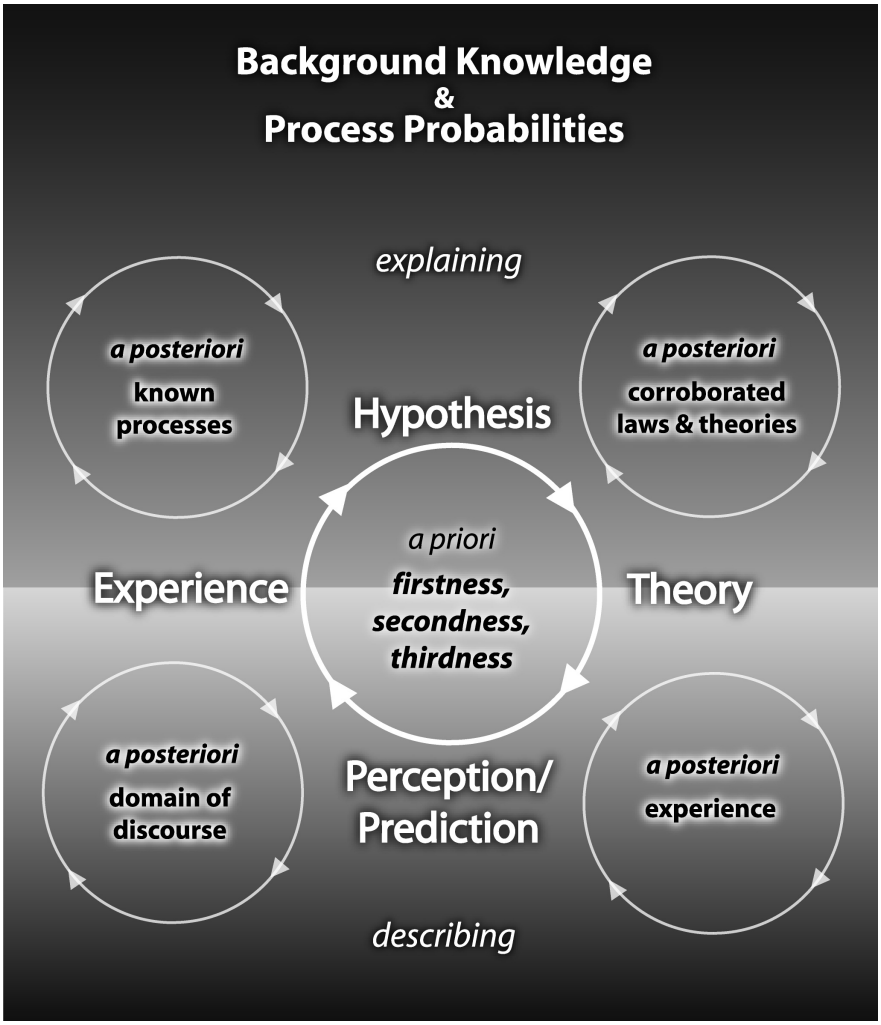

Fig. 8: The structure of cycles of inference, with their parts of description and of explanation. In every step of semiotic inference, the *apriorisms* of *firstness*, *secondness* and *thirdness* are involved. Furthermore does relevant a posteriori-knowledge influence each step of inference specifically. For more details see text.

And, by considering the steps, it is obvious, that each step is necessary - by removing one of those steps, the whole cycle would be interrupted. This holds true also for phylogenetic research. Therefore, all background knowledge that influences the steps of inference that are necessarily applied in phylogenetic research, starting from the empirical perceptions and ending up with the corroborated cladistic hypothesis, belongs to the relevant background knowledge.

The step of *abductive construction* considers all *a posteriori* knowledge (i.e. empirical knowledge) that is concerned with our general domain of discourse – knowledge of our perceptual apparatus, of the machines and methods that were applied to produce the objects that are studied, limits of our language etc. The step of *abduction* considers all *a posteriori* knowledge of the causal processes we know (from experiments etc.), including, if possible, an approximation of their specific process probabilities. The next step, the *deduction*, considers all *a posteriori* knowledge we have about relevant laws and theories, as for instance the species concept, the bifurcating mode of speciation, and more; and methodological rules as for instance the method of parsimony. The last step, the step of *induction*, considers *a posteriori* knowledge of our empirical experience.
All this speaks in favor of the consideration of all relevant empirical background knowledge we have, to be able to perform the inference as good as possible, and therewith contradicting some cladists, who argue for minimizing the amount of considered background knowledge (Kluge, 1997, 1997a; Brower, 2000).

Conclusion

Considering the perspective of a semiotic epistemology one has to rethink the conclusions for the methodology and theory of phylogenetic research taken from a specific interpretation of the popperian falsificationism. In this paper, it is shown at first sight, that the proposed semiotic epistemology is consistent with the general ideas of popperian falsificationism (see also von Pückler, 2002). But some of the popular conclusions taken from falsificationism for phylogenetic research are not tenable in the light of semiotics. For instance is an unweighted parsimony approach in cladistic analyses not justifiable and thus has to be refuted (contradicting Kluge, 1997). From a purely theoretical perspective, it is obvious that process probabilities of specific types of transformation processes represent key parameters for the choice and the quality of the chosen cladistic hypothesis and thus for its justification against alternative hypotheses (see Vogt, 2002a). The question of how to estimate those parameters empirically is a methodological problem.
Another popular ‘falsificationist’ conclusion is the maxim to minimize the amount of the necessarily assumed relevant background knowledge (e.g. Kluge, 1997, 1997a; Brower, 2000). Though, to my opinion, not resembling a necessary consequence of popperian logic, this point of view can go that far that even the assumption of evolution

as relevant background knowledge for cladistic analyses is neglected (Brower, 2000), thereby losing any basis for a general conception of phylogenetic relationship – I ask myself, how can one infer an actual phylogenetic relationship without referring to some kind of general conception of what *phylogeny* is. When considering the cycle of (semiotic) inference (Fig. 8) and the reciprocal illumination (Fig. 3) of (a) the recognition of general types of processes and their corresponding results together with understanding and evaluating their causal mechanisms, and (b) the explanatory power of the interpretations of the actual products of the semiosis of phylogeny in the light of this general knowledge (and vice versa), one has to conclude, that it is reasonable and advisable to take as much relevant and substantiated background knowledge into account as possible when analyzing the data.

Therefore, as some advocates of popperian falsificationism in phylogenetic research argue, that likelihood methods are to be rejected because they take assumptions of evolutionary process probabilities into consideration and because they do not minimize the assumed background knowledge (e.g. Kluge, 1997), is from a semiotic perspective not tenable. On this ground, a justification of the superiority of parsimony methods over likelihood methods seems to be not possible - and what is claimed to be a weakness of likelihood, may turn out to be its strength.

References

Allard, M.W. and Carpenter, J.M. (1996): On Weighting and Congruence. Cladistics 12: 183-198.

Bock, W. J. (1973): Philosophical foundations of classical evolutionary classification. Syst. Zool. 22: 375-392.

Brauckmann, S. (1999): On genes, cells, and memory. Semiotica 127: 151-167.

Brier, S. (1999): Biosemiotics and the foundation of cybersemiotics: Reconceptualizing the insights of ethology, second-order cybernetics, and Peirce's semiotics in biosemiotics to create a non-Cartesian information science. Semiotica 127: 169-198.

Brogaard, B. O. (1999): An Aristotelian approach to animal behavior. Semiotica 127: 199-213.

Brower, A. V. Z. (2000): Evolution is not a necessary assumption of cladistics. Cladistics 16: 143-154.

Bryant, H.N. (1989): An evaluation of cladistic and character analyses as hypothetico-deductive procedures, and the consequences for character weighting. Systematic Zoology 38: 214-227.

Chippindale, P.T. and Wiens, J.J. (1994): Weighting, Partitioning, and Combining Characters in Phylogenetic Analysis. Systematic Biology 43: 278-287.

Cicero, M. T. (1995): De re publica. Cambridge University Press, Cambridge.

Defoe, D. (1719): The Life and strange and surprising Adventures of Robinson Crusoe. Edit. West, R., Carroll & Graf.

Deledalle, G. (2000): Semiotik als Philosophie. Pp. 31-43 in: U. Wirth (eds.), Die Welt als Zeichen und Hypothese. Suhrkamp taschenbuch wissenschaft (stw 1479), Frankfurt a.M.

de Pinna, M. C. C. (1991) : Concepts and tests of homology in the cladistic paradigm. Cladistics 7: 367-394.

de Queiroz, K. and Poe, S. (2001) : Philosophy and phylogenetic inference : A comparison of likelihood and parsimony methods in the context of Karl Popper's writings on corroboration. Syst. Biol. 50: 305-321.

Eco, U. (2000): Kant und das Schnabeltier. Carl Hanser Verlag, München.

Emmeche, C. (1991): A semiotic reflection on biology, living signs and artificial life. Biology and Philosophy 6: 325-340.

Faith, D. P. and Trueman, J. W. H. (2001): Towards an inclusive philosophy for phylogenetic inference. Syst. Biol. 50(3): 331-350.

Farris, J.S. (1969): A successive approximations approach to character weighting. Systematic Zoology 18: 374-385.

Farris, J. S. (1970): Methods for computing Wagner trees. Syst. Zool. 19: 83-92.

Farris, J. S. (2000): Corroboration versus "strongest evidence". Cladistics 16: 385-393.

Goloboff, P.A. (1993): Estimating character weights during tree search. Cladistics 9: 83-91.

Haszprunar, G. (1998): Parsimony Analysis as a Specific Kind of Homology Estimation and the Implications for Character Weighting. Molecular Phylogenetics and Evolution 9: 333-339.

Hoffmeyer, J. (1998): Biosemiotics - The study of living systems from a semiotic perspective. *In*: Encyclopedia of Semiotics (P. Bouissac, ed.), pp. 82-85. Oxford University Press, New York.

Kant, I. (1998): Kritik der reinen Vernunft. A & B edition. J. Timmermann (eds.), Felix Meiner Verlag GmbH, Hamburg.

Kawade, Y. (1999): The two foci of biology: Matter and sign. Semiotica 127: 369-384.

Kluge, A. G. (1997): Testability and the Refutation and Corroboration of Cladistic Hypotheses. Cladistics 13: 81-96.

Kluge, A. G. (1997a): Sophisticated falsification and research cycles: Consequences for differential character weighting in phylogenetic systematics. Zool. Scr. 26: 349-360.

Kluge, A. G. (2001): Philosophical conjectures and their refutation. Syst. Biol. 50: 322-330.

Kull, K. (2000): Organisms can be proud to have been their own designers. Cybernetics and Human Knowing 7: 45-55.

Laubichler, M. D. (1999): A semiotic perspective on biological objects and biological functions. Semiotica 127(1), 415-431.

Mackie, J. L. (1974): The cement of the universe. A study of causation. Oxford University Press, Oxford.

Mackie, J. L. (1985): Logic and knowledge. Clarendon Press, Oxford.

May, M. (1999): Kausales Schliessen. Eine Untersuchung über kausale Erklärungen und Theorienbildung. Graduiertenkolleg Kognitionswissenschaft Universität Hamburg, Bericht Nr. 64.

Mahner, M. and Bunge, M. (1997): Foundations of biophilosophy. Springer, Berlin.

Milinkovitch, M.C., LeDuc, R.G., Adachi, J., Farnir, F., Georges, M. and Hasegawa, M. (1996): Effects of Character Weighting and Species Sampling on Phylogeny Reconstruction: A Case Study Based on DNA Sequence Data in Cetaceans. Genetics 144: 1817-1833.

Neff, N.A. (1986): A rational basis for a priori character weighting. Systematic Zoology 35: 110-123.

Oehler, K. (2000): Einführung in den semiotischen Pragmatismus. Pp. 13-30 in: U. Wirth (eds.), Die Welt als Zeichen und Hypothese. Suhrkamp taschenbuch wissenschaft (stw 1479), Frankfurt a.M.

Pape, H. (1991): Einleitung. Pp. 11-109 in: H. Pape (eds.), Naturordnung und Zeichenprozeß. Suhrkamp taschenbuch wissenschaft (stw 912), Frankfurt a.M.

Pape, H. (2000): Einleitung. Pp. 7-83 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 2. Suhrkamp taschenbuch wissenschaft (stw 1481), Frankfurt a.M.

Pape, H. (2000a): Einleitung. Pp. 7-79 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Patterson, C. (1988): Homology in Classical and Molecular Biology. Mol. Biol. Evol. 5: 603-625.

Peirce, C. S. (1865) : Erste Harvard-Vorlesung (MS 340). Pp. 87-104 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1866) : The logic of science ; or, induction and hypothesis (Lowell Lecture X). In: E. C. Moore, M. H. Fisch, C. J. W. Kloesel, D. D. Roberts, L. A. Ziegler and N. P. Atkinson (eds.) Writings of Charles S. Peirce, Vol.1. Indiana University Press, Indiana.

Peirce, C. S. (1866a) : Neunte Lowell-Vorlesung (MS 357). Pp. 107-127 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1867) : Eine neue Liste der Kategorien. Pp. 147-159 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1893): Die Kunst des Räsonierens (MS 404). Pp. 191-201 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1895): Kurze Logik (MS 595). Pp. 202-229 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1897) : Die Logik der Relative (Translation form The Monist, Vol. 7, Logic of relatives, pp. 161-217). Pp. 269-335 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1898) : Grundbegriffe der Semiotik und formalen Logik. Pp. 336-375 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1901) : Minutiöse Logik (MS 1579 & MS 425 A). Pp. 377-408 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 1. Suhrkamp taschenbuch wissenschaft (stw 1480), Frankfurt a.M.

Peirce, C. S. (1903) : Syllabus of certain topics of logic. Pp. 37-183 in: H. Pape (eds.), Charles S. Peirce. Phänomen und Logik der Zeichen. Suhrkamp taschenbuch wissenschaft (stw 425), Frankfurt a.M.

Peirce, C. S. (1903a): Lowell-Lecture. Teil 1 des 3. Entwurfs der Dritten Vorlesung (MS 464). Pp. 127-145 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 2. Suhrkamp taschenbuch wissenschaft (stw 1481), Frankfurt a.M.

Peirce, C. S. (1904) : Neue Elemente (MS 517). Pp. 339-377 in: H. Pape (eds.), Naturordnung und Zeichenprozeß. Suhrkamp taschenbuch wissenschaft (stw 912), Frankfurt a.M.

Peirce, C. S. (1905): 2. Entwurf zum Aufsatz: Die Grundlagen des Pragmatizismus (MS 283). Pp. 308-353 in: H. Pape and J.W. Kloesel (eds.), Charles S. Peirce. Semiotische Schriften Band 2. Suhrkamp taschenbuch wissenschaft (stw 1481), Frankfurt a.M.

Platnick, N. I., and Gaffney, E. S. (1977): Systematics: A Popperian Perspective. Syst. Zool. 26: 360-365.

Riedl, R. (2000): Strukturen der Komplexität. Springer, Berlin.

Roesler, A. (2000): Vermittelte Unmittelbarkeit. Aspekte einer Semiotik der Wahrnehmung bei Charles S. Peirce. Pp. 112-129 in: U. Wirth (eds.), Die Welt als Zeichen und Hypothese. Suhrkamp taschenbuch wissenschaft (stw 1479), Frankfurt a.M.

Sebeok, T. A. (1990): Indexicality. The American Journal for Semiotics 7(4): 7-28.

Spohn, W., Ledwig, M. and Esfeld, M. (2001): Current issues in causation. Mentis, Paderborn.

Stjernfelt, F. (1999): Biosemiotics and formal ontology. Semiotica 127: 537-565.

Trueman, J.W.H. (1998): Reverse Successive Weighting. Systematic Biology 47: 733-737.

Vogt, L. (2002): Testing and weighting characters. O. D. E. 2:

Vogt, L. (2002a): Process probabilities and the weighting of characters in systematics.

von Pückler, T. (2002): Peirce und Popper über Hypothesen und ihre Bildung. http://www.rz.uni-frankfurt.de/~wirth/texte/Pückler.html.

von Uexküll, T. (1999): The relationship between semiotics and mechanical models of explanation in the life sciences. Semiotica 127: 647-655.

Wagner, G. P. (1986): The systems approach: an interface between development and population genetic aspects of evolution. *In* "Patterns and processes in the history of life" (D. M. Raup and D. Jablonski, Eds.), pp. 149-165. Springer, Berlin.

Wiens, J.J. (2001): Character analysis in morphological phylogenetics: Problems and solutions. Systematic Biology 50: 689-699.

Zeidler, K. W. (2000): Prolegomena zur Wissenschaftstheorie. Verlag Königshausen & Neumann GmbH, Würzburg.