

## Zusammenfassung der Dissertation "Relative Timing of Intron Gain and a New Marker for Phylogenetic Analyses" (Relative Datierung von Intron-Insertionsereignissen und ein neuer Marker für phylogenetische Analysen) von Jörg Lehmann:

Trotz jahrzehntelanger Bemühungen von Molekularsystematikern sind die Stammbäume eukaryotischer Organismen noch immer z.T. ungeklärt oder im Widerspruch miteinander. Eine fortwährend zunehmende Zahl an vollständig-sequenzierten Genomen verschiedener Eukaryoten erlaubt es, Gen- und Spezies-Phylogenien auf Genom-Ebene zu untersuchen. Jedoch zeigten solch Phylogenomics-basierten Ansätze auch, dass mehr und mehr Taxa und Gensequenzen nicht die ultimative Lösung darstellen, um diese Stammbaumkonflikte aufzuheben, und dass es einen Bedarf für Sequenz-unabhängige phylogenetische Meta-Merkmale gibt, die von Genomsequenzen abgeleitet sind.

Spliceosomale Intronen sind ein charakteristisches Merkmal von eukaryotischen Kerngenomen. Die relativ selten auftretenden Veränderungen in spliceosomalen Intron-Positionen wurden bereits als Meta-Merkmale sowohl zur Vorhersage von Intron-Evolution als auch Verwandtschaftsverhältnissen verwendet, jedoch mit unterschiedlichem Erfolg. In dieser Arbeit wird eine spezifische Teilmenge dieser Veränderungen betrachtet, und als neuartiger phylogenetischer Marker etabliert, genannt "Near Intron Pair" (NIP). Diese Marker werden aus homologen Genen abgeleitet, welche wechselseitig-exklusiv auftretende Intronen an paarweisen, nah-benachbarten kodierenden Sequenzpositionen aufweisen. Die Idee, dass NIPs mächtige Merkmale darstellen können, basiert auf der Annahme, dass sowohl die Existenz sehr kurzer Exonen als auch der multiple Intron-Gewinn selten ist.

Um ausreichende Mengen an NIP-Merkmalen aus genomischen bzw. Alignment-Datensätzen auf konsistente und flexible Weise zu erhalten, war die Umsetzung einer computergestützten Pipeline ein Hauptziel dieser Arbeit. Ausgehend von homologen Gen-Datensätzen, welche die genomischen Sequenzen und die zugehörigen Transkript-Annotationen umfassen, ist die Erstellung eines multiplen Alignments ein wesentlicher Teil dieser Pipeline. Dieses kann auf Aminosäure-Ebene ermittelt werden, und resultiert durch Rückübersetzung in einem Kodon-Alignment, in dem homologe Intron-Positionen zwischen den einzelnen Transkript-Sequenzen sichtbar werden. Die Pipeline liefert zunächst hieraus Alignment-Abschnitte, welche mindestens ein potentiell NIP-Merkmal enthalten. Im nächsten Schritt werden aus den gesammelten Abschnitten schliesslich 0/1-Merkmale extrahiert, welche die gültigen NIPs in Abhängigkeit von Qualitätsfilterkriterien wie dem Grad der Alignmentkonservierung um die Intronpositionen repräsentieren. Die Pipeline-Tools ermöglichen es zudem, NIP-Merkmalenmatrizen auszuwerten, welche für die Stammbaumrekonstruktion, z.B. mittels des Maximum Parsimony Ansatzes, verwendet werden können.

In einer ersten NIP-basierten Anwendung wurden die phylogenetischen Positionen von Hauptordnungen der holometabolen Insekten im kladistischen Sinne ausgewertet. Wie bereits während einer Studie zum eIF2gamma-Gen aufgrund von zwei NIPs vorgeschlagen (Krauss et al. 2005), unterstützte die genomweite Auswertung die Ordnung Hymenoptera als Schwestergruppe zu einem Verbund von Coleoptera und Mecoptera, in Übereinstimmung mit anderen Studien, aber entgegen der vorher etablierten Hypothese. Die NIP-Methode wurde im Rahmen des Genomprojektes zur Beschreibung einer neuen Art

von parasitären Fächerflüglern (Strepsiptera) angewandt, um deren Stellung innerhalb der Insekten mit aufzuklären. In Übereinstimmung mit Ergebnissen aus Analysen des Sequenzmusters und eines weiteren Meta-Charakters ergab sie, dass die Fächerflügler die unmittelbar nächsten Verwandten der Käfer darstellen. NIP-basierte Rekonstruktionen des Tier-Stammbaums, welche eine breite Auswahl an repräsentativen Arten abdeckten, identifizierten auch einige Schwachpunkte des NIP-Ansatzes, der z.B. unter ungenauen Alignment- oder Orthologen-Vorhersagen leidet, und gegenüber systematischen Verzerrungen empfindlich sein kann (long-branch-attraction Artefakt). In einer weiteren Studie konnte die Bestimmung von NIP Merkmalen innerhalb der evolutionär jungen Gattung *Drosophila* verwendet werden, um junge Intron-Entstehungsereignisse zu charakterisieren. Diese umfassen mehrere Fälle von Intron-Migration und Tandem-Exon-Duplikation, wenn auch die Entstehungsmechanismen für die Mehrheit der Intronen nicht aufgefunden werden konnten.