

Science: Big Data erklärt Evolution der Vögel

Stammbaum der Vögel mittels Gen-Analysen und Supercomputern nachvollzogen / Neue Erkenntnisse über Grundlagen von Gesang, Federn, Biodiversität und Entwicklung der Vögel



Für die Familie der Bienenfresser (im Bild *Merops bullocki*) zeigte die Studie eine enge Verwandtschaft mit Singvögeln, Papageien und Raubvögeln. (Bild: Peter Houde)

Rund 95 Prozent der über 10 000 bekannten Vogelarten entwickelten sich erst nach dem Aussterben der Dinosaurier vor rund 66 Millionen Jahren. Computeranalysen der genetischen Daten zeigen auch, dass sich die heutige Vielfalt explosionsartig aus wenigen Spezies bereits nach 15 Millionen Jahren entwickelt hatte. KIT-Forscher entwarfen die Algorithmen für die umfassende Analyse der Vogel-Evolution. Für die jetzt im Fachjournal Science vorgestellten Ergebnisse wurde eine Rechenleistung von 300 Prozessorjahren benötigt. (DOI 10.1126/science.1253451)

„Die Berechnung dieser Stammbäume für die Evolutionsforschung ist ohne entsprechende Algorithmen und Supercomputer nicht möglich“, erklärt Alexandros Stamatakis, Professor für High Performance Computing in den Lebenswissenschaften am Karlsruher Institut für Technologie und Leiter der Forschungsgruppe „Scientific Computing“ am Heidelberger Institut für Theoretische Studien (HITS). „Moderne Sequenzanalysen liefern heutzutage umfangreiche Gendaten über zahlreiche Spezies. Aber mit der Aufgabe, auch evolutionäres Wissen aus diesen riesigen und komplexen Datenmengen zu erzeugen, waren Computerprogramme bislang selbst auf Supercomputern überfordert.“

Monika Landgraf
Pressesprecherin

Kaiserstraße 12
76131 Karlsruhe
Tel.: +49 721 608-47414
Fax: +49 721 608-43658
E-Mail: presse@kit.edu

Weiterer Kontakt:

Kosta Schinarakis
PKM – Themenscout
Tel.: +49 721 608 41956
Fax: +49 721 608 43658
E-Mail: schinarakis@kit.edu

Obwohl Supercomputer mittlerweile Tausende Prozessoren besitzen, war die Software für Stammbaumanalysen auf etwa 500 Prozessoren beschränkt. „Wir mussten daher das Kommunikationsschema zwischen den Programmteilen auf verschiedenen Prozessoren überdenken und neu entwickeln,“ so Stamatakis. Der neue Ansatz beschleunigte die Software um den Faktor 3 und ermöglicht es gleichzeitig die Rechnungen auf 4000 Prozessoren effizient zu verteilen. Informatiker sprechen hier von der Parallelisierung von Algorithmen. „Statt 24 Monate warten wir nun also nur 1 Monat auf die Ergebnisse.“

Die Berechnung von Stammbäumen gehört zu einer extrem rechenintensiven Klasse von mathematischen Problemen (NP-Hard genannt). „Schon bei 50 Spezies gibt es mehr als 10 hoch 76 mögliche Stammbäume, aus denen der Richtige gefunden werden muss“, erklärt Andre J. Aberer, Doktorand am KIT und Mitarbeiter des HITS, der die Computeranalysen durchgeführt hat. „Zum Vergleich: Im Universum gibt es etwa 10 hoch 78 Atome.“ Die Algorithmen filtern zunächst grob die unwahrscheinlichen Evolutionsszenarien heraus. Unter den verbliebenen wird dann anhand der Daten von jeweils 14 000 Genen von 48 repräsentativen Vogelspezies der evolutionäre Stammbaum berechnet, der die Daten plausibel erklärt. Die neue parallele Software wurde am Höchstleistungsrechner "SuperMUC" des Leibniz-Rechenzentrums der Bayerischen Akademie der Wissenschaften und zwei US-Rechenzentren ausgeführt. Die aufgewandte Rechenleistung entspricht einer Rechenzeit von 300 Jahren auf einem einzelnen Prozessor.

„Die von uns entwickelten Methoden zur Stammbaumberechnung sind auf alle Arten von Lebewesen anwendbar“, so Stamatakis. Sie ermöglichten bereits eine umfassende Studie zum Stammbaum der Insekten mit 144 Arten, die vor Kurzem im Fachmagazin Science veröffentlicht wurde. Aber auch der Ursprung und die Verbreitung von Viren und Bakterien lassen sich damit nachvollziehen, um etwa Krankheitserreger besser bekämpfen zu können. Die Analyse der genetischen Verwandtschaft von australischen Giftschlangen half entscheidend, die noch fehlende Gegengifte für einige Schlangenarten zu identifizieren.

Zu den neuen Erkenntnissen über Vögel, die nun zeitgleich in insgesamt 23 wissenschaftlichen Veröffentlichungen bei Science und weiteren Fachjournalen publiziert werden gehören, neben dem Stammbaum und der Evolution der Vögel:

- Genetische Grundlagen der Biodiversität.
- Die genetischen Grundlagen der Hirnregionen, die die Entstehung des Vogelgesangs steuern.
- Der Verlust von Zähnen bei Vögeln vor rund 100 Millionen Jahren.
- Die Verwandtschaft von Dinosauriern und Vögeln.
- Wie farbige Federn entstanden sind.

Die aktuelle Studie über die Entwicklung der Vögel wurde vom „Avian Phylogenomics Consortium“ durchgeführt, welches 200 Wissenschaftler an 80 Instituten in 20 Ländern umfasst. Koordinatoren waren Guojie Zhang vom BGI in China, Erich D. Jarvis von der Duke University in den USA und M. Thomas P. Gilbert vom Natural History Museum in Dänemark. Tandy Warnow von der University of Illinois und Alexandros Stamatakis waren die Koordinatoren der Computeranalysen. Alexandros Stamatakis und Andre J. Aberer sind die einzigen deutschen Ko-Autoren. Die Studie stellt die größte Genomanalyse einer Wirbeltierklasse dar und umfasst unter anderem Enten, Falken, Spechte und weitere Repräsentanten aller Zweige der modernen Vögel. Alle Daten und Methoden werden nun Forschern weltweit für weitere Studien kostenlos zugänglich gemacht.

Web-Site der Forschungsgruppe von Alexandros Stamatakis:

www.exelixis-lab.org

http://www.informatik.kit.edu/309_6333.php

Interview am KIT mit Alexandros Stamatakis:

<https://www.youtube.com/watch?v=6Laa58M-vk>

Das Karlsruher Institut für Technologie (KIT) ist eine Körperschaft des öffentlichen Rechts nach den Gesetzen des Landes Baden-Württemberg. Es nimmt sowohl die Mission einer Universität als auch die Mission eines nationalen Forschungszentrums in der Helmholtz-Gemeinschaft wahr. Thematische Schwerpunkte der Forschung sind Energie, natürliche und gebaute Umwelt sowie Gesellschaft und Technik, von fundamentalen Fragen bis zur Anwendung. Mit rund 9 400 Mitarbeiterinnen und Mitarbeitern, darunter mehr als 6 000 in Wissenschaft und Lehre, sowie 24 500 Studierenden ist das KIT eine der größten Forschungs- und Lehreinrichtungen Europas. Das KIT

**verfolgt seine Aufgaben im Wissensdreieck Forschung – Lehre
– Innovation.**

Diese Presseinformation ist im Internet abrufbar unter: www.kit.edu