

PETER UETZ, JOHANNES GOLL & JAKOB HALLERMANN Die TIGR-Reptiliendatenbank

Aus der EMBL-Reptiliendatenbank wurde vor kurzem die TIGR-Reptiliendatenbank. Geändert hat sich nicht nur der Name, sondern auch die Oberfläche und einige Funktionen dieser umfassendsten Internetdatenbank zum Thema Reptiliensystematik. Geblieben ist die seit 2001 aktive Adresse: <http://www.reptile-database.org>

Einleitung

Als die EMBL-Reptiliendatenbank Ende 1995 gegründet wurde, war die Entwicklung des Internets noch am Anfang. Nach einer Studie von ARD und ZDF stieg die Zahl der deutschen erwachsenen Internetnutzer von 6,5% im Jahr 1997 auf rund 63% im Jahr 2007. Damit dürfte auch die Mehrzahl der DGHT-Mitglieder Zugang zu diesem und anderen Online-Diensten haben.

Einer von uns [PU] begann die Arbeit an der EMBL-Reptiliendatenbank als Doktorand am Europäischen Molekularbiologischen Labor (EMBL) Ende 1995. Mitte 1996 ging die Datenbank mit einer Suchmaschine online (UETZ & ETZOLD 1996a,b). Die Benutzeroberfläche lehnte sich eng an die EMBL-DNA-Sequenzdatenbank an, die von Thure Etzold am EMBL entwickelt worden war. Der letzte Mitarbeiter dieses Teams, Ramu Chenna, verließ aber Ende 2006 das EMBL, worauf das EMBL die Unterstützung der Datenbank aufkündigte. Wir waren daher gezwungen, auf einen anderen Anbieter umzuziehen. Aus privaten, beruflichen und politischen Gründen zogen wir [PU + JG] außerdem zurück in die USA, wo unser neuer Arbeitgeber, The Institute of Genomic Research (TIGR), anbot, die Datenbank bereit zu stellen. Aus diesem Grund haben wir uns

auch entschlossen, den Namen entsprechend zu ändern. Während die Startwebseiten immer noch auf Servern in Deutschland gespeichert sind, steht der Server mit der eigentlichen Artendatenbank jetzt in Rockville im US-Bundesstaat Maryland.

Beschreibung

Die Datenbank konzentriert sich nach wie vor primär auf taxonomische Daten auf Artniveau. Dabei werden Artname, Autor, Beschreibungsjahr sowie die bibliographischen Daten erfasst. Hinzu kommen Typus-Angaben, Lokalität und Sammlungsdaten (Tabelle 1). Diese Informationen sowie die Verbreitung auf Länderebene sind praktisch vollständig für alle 8589 Arten vorhanden. Damit lassen sich auch leicht Checklisten für einzelne Länder zusammenstellen. Hinzu kommen pro Jahr 50-100 neue Arten, erstaunlicherweise mit steigender Tendenz (Abb. 1).

Neuerdings versuchen wir, aus den zunehmend elektronisch vorliegenden Originalbeschreibungen die Diagnosen und weitere Informationen direkt in die Datenbank zu kopieren. Die aktuelle Versi-

Tabelle 1. Kennzahlen der Version Juni 2007

Arten	8589
Referenzen	23252
Typuslokalität	5943
Typusdaten	3840
Fotos (lokal und WWW)	3206
Fotos (lokal)	1334
Trivialnamen	6282
Trivialnamen	1240

Zahlen (außer Referenzen) geben die Artenzahl an, für die diese Information vorliegt. „Fotos (lokal)“ liegen auf unserem eigenen Server, die anderen nur im Internet.

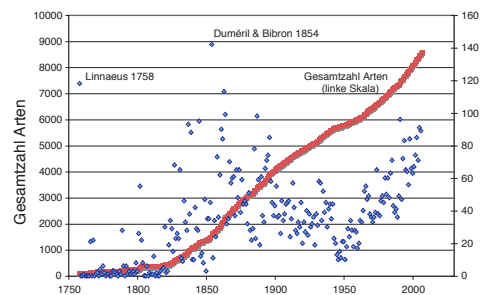


Abb. 1. Zahl bekannter Arten. (Rote) Linie: Gesamtzahl der bekannten Arten (linke Skala). (Blaue) Punkte/Rauten: Pro Jahr neu beschriebene Arten (rechte Skala). Hervorgehoben sind ausserdem die Jahre 1758 und 1854, in denen zwei Klassiker mit vielen neuen Arten veröffentlicht wurden.

on enthält 436 Diagnosen. Daneben enthält die Datenbank Angaben zur Etymologie von 1534 Arten,

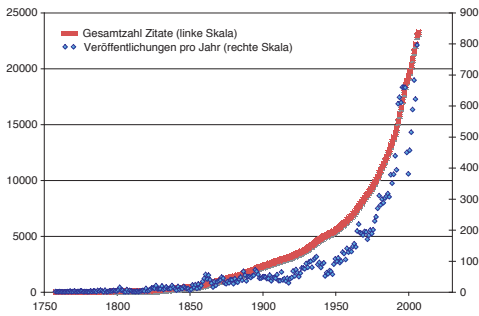


Abb. 2. Literaturzitate in der Datenbank. Blaue Punkte: Zahl der im jeweiligen Jahr publizierten Arbeiten, die in der Datenbank zitiert werden. Rote Linie: Kumulierte Gesamtzahl der Publikationen, d.h. die Summe der bis zu diesem Jahr publizierten (und in der Datenbank verzeichneten!) Arbeiten.

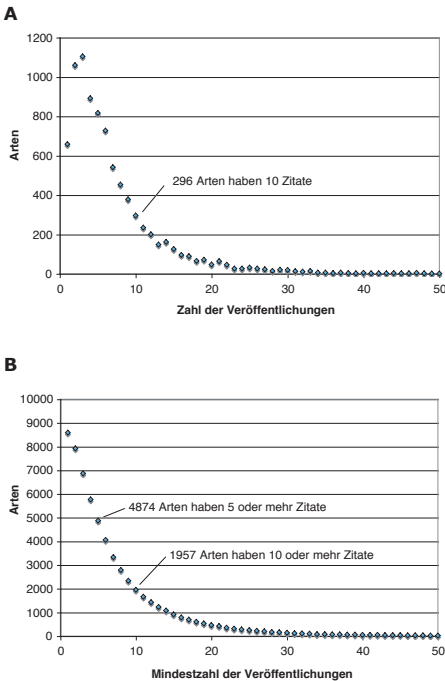


Abb. 3A. Referenzen pro Art. Die Mehrzahl der Arten hat 2-6 Literaturzitate. 3B. Kumulierte Anzahl von Referenzen für Arten.

zumeist mit Bezug auf den Artnamen. Solche, weniger systematik-relevante Daten werden derzeit im "Kommentar"-Feld gespeichert, aus dem sie in eigene Datenfelder ausgelagert werden, sobald eine grössere Anzahl vorhanden ist. Zum Beispiel haben wir derzeit Höhenangaben für 782 Arten; in absehbarer Zeit wird diese Information in ein eigenes Datenfeld abgetrennt werden. Weitere Informationen lassen sich in der Regel über die Literaturhinweise finden, die für etliche Arten mittlerweile Dutzende von Zitaten umfassen (Tabelle 2, Abb. 2 & 3).

Familien

„Familien“ dienen als praktische Orientierungshilfe, um sich in der fast unüberschaubaren Vielfalt der Reptilien und anderer Lebewesen zurechtzufinden. Familien sind aber noch schwieriger und damit noch willkürlicher zu definieren als Arten (siehe unten). Die Datenbank verwendet zwar ein eher konservatives Familienkonzept, zum Beispiel bei der Behandlung der Großleguane (Iguaninae) als Unterfamilie, folgt aber auch neuen Erkenntnissen, wenn beispielsweise DNA-Sequenzen die Aufspaltung einer Familie erzwingen um paraphyletische Beziehungen zu vermeiden.

Obwohl sich die Arten am leichtesten über die Suchmaschine finden lassen, erlaubt die Webseite auch, die Arten über höhere Kategorien wie Ordnungen und Familien zu erschließen. Dazu linkt die

Tab. 2. Top-20-Arten (mit 49 oder mehr Literaturzitaten)

Art	Zitate
<i>Hemidactylus turcicus</i>	91
<i>Thamnophis sirtalis</i>	86
<i>Testudo graeca</i>	80
<i>Hemidactylus frenatus</i>	79
<i>Lampropeltis triangulum</i>	75
<i>Hemidactylus mabouia</i>	69
<i>Trachemys scripta</i>	68
<i>Varanus salvator</i>	67
<i>Lepidodactylus lugubris</i>	66
<i>Natrix natrix</i>	65
<i>Emys orbicularis</i>	63
<i>Sceloporus undulatus</i>	61
<i>Podarcis muralis</i>	60
<i>Laudakia stellio</i>	59
<i>Boa constrictor</i>	58
<i>Cnemidophorus tigris</i>	55
<i>Cyrtopodion kotschy</i>	53
<i>Chelonia mydas</i>	51
<i>Eutropis multifasciata</i>	50
<i>Hemidactylus brookii</i>	49

Hauptseite von den Ordnungen bzw. Unterordnungen zu Familien, die jeweils kurz auf einer eigenen Seite charakterisiert werden und Listen aller 1052 Reptiliengattungen enthalten. Ein Klick auf die Gattungsnamen zeigt alle Arten der Gattung an. Die Familienseiten enthalten in den meisten Fällen Stammbäume bis Gattungs- oder Artniveau, die die Beziehung der Gattungen untereinander darstellt. Derzeit sind ungefähr die Hälfte der rund 65 Familien in Stammbäumen abgebildet. Allerdings werden die grossen Familien der Geckos, Skinke, und Colubriden in Stammbäumen nur teilweise abgebildet. Interessanterweise enthalten die 7 artenreichsten Gattungen schon 1030 oder 12% aller Reptilienarten (vgl. Tab. 3).

Artkonzepte

Ein Grund für die rasch zunehmende Artenzahl ist nicht zuletzt die Anwendung alternativer Artkonzepte. Während historisch das biologische Artkonzept dominierte (das reproduktive Isolation der Artdefinition zugrunde legt), werden zunehmend evolutionäre oder phylogenetische Artkonzepte angewandt (DE QUEIROZ 2005). Letztere definieren „diagnostizierbare“ oder phylogenetisch definierte Populationen als Art. In der Datenbank wird tendenziell das biologische Artkonzept bevorzugt, obwohl sich in den meisten Fällen nicht wirklich ent-

scheiden lässt, ob reproduktive Isolation vorliegt. Allerdings sind auch phylogenetische Definitionen nicht wirklich befriedigend, da sich eine phylogenetische Gruppe („Clade“) theoretisch über einzelne DNA-Basenpaare definieren lassen, was das Konzept mitunter *ad absurdum* führt. Da wir dazu tendieren, das biologische Artkonzept beizubehalten, ist auch die aktuelle Artenzahl eher konservativ, d.h. die Datenbank würde mit konsequenter Anwendung des evolutionären Artkonzeptes deutlich mehr Arten enthalten. Allerdings lässt sich über die Synonymie leicht feststellen, von welchen Autoren zum Beispiel eine Unterart als volle Art betrachtet wird. Phylogenetisch betrachtet könnte man einen Großteil der Unterarten als „volle“ Arten betrachten. Derzeit sind für 1307 Arten 4496 Unterarten (einschließlich Nominatformen) verzeichnet.

CD-ROM

Die Datenbank hat mittlerweile einen Umfang angenommen, der es unmöglich macht, die Daten in Papierform zu publizieren. Ein Ausdruck würde mindestens 3000 Seiten in Schriftgröße 10 benötigen. Um die Datenbank regelmäßig zu archivieren, geben wir deshalb ein bis zwei mal jährlich eine CD-ROM heraus, die nicht nur die Datenbank sondern auch digitale Versionen historischer Bücher enthält (siehe unten). Die Datenbank auf CD-ROM ist als Filemaker-Programm implementiert, das sowohl unter Windows und unter MacOS X läuft und einige zusätzliche Fähigkeiten besitzt, die in der Webversion nicht angeboten werden (können), zum Beispiel Statistikfunktionen. Beispielsweise kann man sich alle afrikanischen Schildkrötengattungen mit jeweiliger Artenzahl anzeigen lassen, die vor 1900 beschrieben wurden.

Eine verkürzte Version der Reptiliendatenbank ist Bestandteil des „Catalogue of Life“ der mittlerweile über eine Million wissenschaftlich beschriebene Organismen auflistet und von der Species2000-Foundation herausgegeben wird (UETZ 2007).

Tab. 3. Die 7 artenreichsten Reptiliengattungen (mehr als 100 Arten)

Gattung	Artenzahl
<i>Anolis</i>	208
<i>Liolaemus</i>	183
<i>Norops</i>	152
<i>Typhlops</i>	149
<i>Sphenomorphus</i>	127
<i>Atractus</i>	108
<i>Leptotyphlops</i>	103
Summe:	1030

Tab. 4. Derzeit verfügbare digitalisierte historische Literatur und die Zahl der darin neubeschriebenen (und noch gültigen!) Arten

Werk	Arten
Linnaeus (1758) Systema Naturae, Herpetologischer Teil	117
Duméril & Bibron (1838-1854) Érpétologie Générale, 10 Bände	209
Annals and Magazine of Natural History (1838-1899)	381
Proceedings of the Academy of Natural Sciences of Philadelphia (1841-1899)	301
Zoologischer Anzeiger (1880-1905)	81
Boulenger (1885-1896) Catalogue of lizards/snakes [...], 6 Bände*	122
Summe	1211

*in Vorbereitung

Historische Literatur

Die TIGR-Reptiliendatenbank will nicht nur die Originalliteratur zu bestimmten Taxa erschliessen, sondern langfristig auch diese Originalliteratur zugänglich machen. Obwohl dies aus urheberrechtlichen Gründen nur bedingt für das neuere Schrifttum möglich ist, bieten wir historische Schriften sowohl im WWW als auch auf CD-ROM an, wobei der Schwerpunkt wiederum auf der Taxonomie liegt. Tabelle 4 listet die Werke auf, die bereits als digitale Versionen vorliegen (zumeist hochauflösende Scans mit 400-600 dpi). Die 1211 in diesen Werken beschriebenen Arten entsprechen rund einem Drittel der bis 1900 beschriebenen Arten und sind in der Datenbank indiziert sodass zu den Originalnamen auch die aktuell gültigen Namen leicht auffindbar sind (und umgekehrt). Damit kann man mit der CD-Version auf Knopfdruck alle Arten eines Werkes anzeigen. Für weitere Informationen zum Projekt „historische Literatur“ siehe UETZ (2002). Die Datenbank-Webseite führt außerdem zu weiteren digitalen Versionen historischer Literatur im Internet.

Weitere Entwicklung

Die Entwicklung jeder Datenbank ist ein Fass ohne Boden. Aus diesem Grund werden wir uns auch weiterhin auf taxonomische Informationen konzentrieren, auch wenn es verführerisch ist, Angaben zur Biologie aufzunehmen. Die Sammlung historischer Originalliteratur, von Fotos und von faunistischen Arbeiten wird aber auch in Zukunft ein Schwerpunkt der Datenbank bleiben. Ein wichtiges Ziel wird es aber auch sein, die Datenbank zu einem Bestimmungswerkzeug zu entwickeln, sodass beliebige Reptilien bestimmt werden können. Dies ist schon jetzt bedingt möglich, sofern die Herkunft eines Exemplars bekannt ist. In diesem Fall kann man nach allen Arten aus dieser Gegend suchen und dann gegebenenfalls mithilfe von Fotos oder weiteren Angaben in der Datenbank die Art bestimmen. Dies wird dadurch erleichtert, dass Fotos der häufigsten Arten entweder in der Datenbank oder auf verlinkten Webseiten zu finden sind. Allerdings wird die Bestimmung von Arten durch das phylogenetische Artkonzept wieder erschwert, zumal in vielen Fällen DNA-Sequenzen als diagnostische Merkmale herangezogen werden müssen.

Mitarbeit

Die Datenbank ist auf Ihre Mitarbeit angewiesen und Beiträge sind jederzeit willkommen. Vor allem sind wir ständig auf der Suche nach Originalbeschreibungen, Fotos, und Korrekturen existierender Daten. Die Datenbankeinträge der aktuellen Version weisen dafür einen link am unteren Seitenende

auf, über den man Korrekturen oder Kommentare zu einzelnen Arten einschicken kann. Machen Sie von diesem Link regen Gebrauch!

Es wäre freilich auch wünschenswert, die englischen Webseiten ins Deutsche zu übersetzen. Wenn Sie an einer Mitarbeit diesbezüglich interessiert sind, melden Sie sich bitte. Wie bisher erhalten Mitarbeiter kostenlose Kopien unserer CD-ROM. Weitere Angaben hierzu finden Sie auf der Webseite.

Dank

Am TIGR werden die Datenbank-Infrastruktur von HANK WU und die Webseiten von CONNIE SHIAU betreut. Die TIGR-Reptiliendatenbank selbst ist nur mit großzügiger Unterstützung zahlreicher freiwilliger Mitarbeiter möglich, die Daten und Fotos zur Verfügung gestellt haben. Eine Nennung ist aus Platzgründen leider nicht möglich, aber auf unserer Webseite findet sich eine weitgehend vollständige Liste. Wir bedanken uns außerdem bei unseren Sponsoren, die nicht nur Literatur zur Verfügung stellen, sondern auch die laufenden Kosten mittragen. Von 2003-2005 wurde die Datenbank von der EU im Rahmen des Species2000-Projektes gefördert.

Schriften

- DE QUEIROZ, K. (2005) Different species problems and their resolution. – *BioEssays* 27:1263–1269.
- UETZ, P. & T. ETZOLD (1996a) Die EMBL-Reptiliendatenbank. – *elaphe* 4/96: 49-53
- (1996b) The EMBL/EBI Reptile Database. – *Herpetological Review* 27(4): 174-175
- UETZ, P. (2002) Alte Bücher – neue Technik: Das Projekt „Historische Literatur“ bei der EMBL-Reptiliendatenbank. – *Beiträge zur Literatur und Geschichte der Herpetologie und Terrarienkunde*, Bd. II: 41-46
- (2007) Reptiles. – In: 2007 Catalogue of Life, CD-ROM, Species 2000, University of Reading. Online <http://www.species2000.org>

Autoren

Dr. PETER UETZ &
Dipl. Bioinformatiker JOHANNES GOLL
The Institute of Genomic Research
9712 Medical Center Drive, Rockville,
MD 20850, USA
E-Mail: uetz@jcvi.org,
E-Mail: johannes.goll@gmail.com

Dr. JAKOB HALLERMANN
Zoologisches Museum, Universität Hamburg,
Martin-Luther-King Platz 3, D-20146 Hamburg,
E-Mail: hallermann@uni-hamburg.de