



## Jeffrey L. Thorne, Ph.D.

Professor für Genetik und Statistik

North Carolina State University

Born in 1963 in Evanston, Ill., USA

Studied Mathematics and Molecular Biology at the University of Wisconsin and Genetics at the University of Washington and at Cornell University

SCHWERPUNKT

### ARBEITSVORHABEN

## Statistische Techniken zur Untersuchung der Evolution von DNA- und Protein-Sequenz

My research focus is on developing statistical techniques for extracting evolutionary information from DNA sequence data. During the last several years, I have mostly concentrated on connecting genotypic change (i. e., DNA sequence evolution) to phenotype (i. e., the characteristics of an organism). The relationship between genotype and phenotype is central to biology, but statistical techniques for studying how this relationship affects evolution are primitive.

A wide variety of computational techniques aim to predict aspects of phenotype (e. g., protein structure, RNA structure, antigenicity) from DNA sequences. We are working to improve probabilistic models of molecular evolution by incorporating these computational biology prediction systems. Because phenotype can induce dependence among changes that occur at different positions within a gene sequence, conventional procedures for making evolutionary inferences are not computationally feasible and we have developed new procedures.

On a separate topic, I am also interested in combining fossil and DNA evidence to estimate divergence times (i. e., the times since most recent common ancestry of related evolutionary lineages). Within this topic, another emphasis is to characterize the pattern by which chronological rates of sequence evolution change over time. This is important because the pattern of evolutionary rate

change has great potential to illuminate the process of biological evolution.

### Recommended Reading

Thorne, J. L., H. Kishino, and J. Felsenstein. "An evolutionary model for the maximum likelihood alignment of DNA sequences." *J Mol Evol* 33: (1991): 114-124.

Thorne, J. L., H. Kishino, and I. S. Painter "Estimating the rate of evolution of the rate of molecular evolution." *Mol Bio Evol* 15: (1998): 1647-1657.

Robinson, D. M., D. T. Jones, H. Kishino, N. Goldman, and J. L. Thorne. "Protein evolution with dependence among codons due to tertiary structure." *Mol Biol Evol* 20, 10 (2003): 1692-1704.

## Die DNA als historisches Instrument

Die wichtigsten Ereignisse in der Geschichte des Lebens ereigneten sich Millionen und Milliarden von Jahren vor dem Beginn der Menschheit und damit vor der Möglichkeit, diese Ereignisse aufzuzeichnen. Wie andere Historikern möchten auch die Evolutionsbiologen wissen, welche Ereignisse stattfanden, und die Kräfte bestimmen, die zu ihnen führten. Dokumentierte fossile Funde liefern dafür einige unentbehrliche Informationen. Jedoch sind diese Funde meist spärlich und schwer zu interpretieren.

Die andere Hauptinformationsquelle für die biologische Geschichte stellt die genetische Information noch vorhandener Organismen dar. Neueste technologische Entwicklungen haben wesentlich dazu beigetragen, genetisches Material wie beispielsweise die DNA-Sequenz besser zu bestimmen. Daher bilden die Daten der DNA-Sequenz heute eine wichtige Grundlage in der Erforschung der biologischen Geschichte.

Der Phänotyp eines Organismus ist seine äußere Erscheinung und sein Verhalten, der Genotyp wiederum seine DNA. Eine zentrale Frage in der Evolutionsbiologie ist der Einfluss des Phänotyps auf den Fortbestand des Genotyps. Die Herausforderung bei der Berechnung der natürlichen Selektion besteht darin, dass über die Beziehung zwischen Genotyp und Phänotyp wenig bekannt ist. Derzeit werden Rechenmethoden entwickelt, mit denen man gewisse Aspekte des Phänotyps der DNA-Sequenz vorhersagen kann. Unsere Strategie ist, diese phänotypischen Vorhersage-Schemata in Wahrscheinlichkeitsmodelle von DNA-Sequenz-Veränderungen zu integrieren. Die sich daraus ergebenden Evolutionsmodelle dienen als Basis für statistische Techniken. Von den Auswirkungen phänotypischer Merkmale können wir mit Hilfe dieser Techniken wiederum auf die Raten der DNA-Sequenz-Veränderung schließen.

Mit den meisten statistischen Verfahren zur Berechnung der natürlichen Selektion kann man die Datensätze der DNA verschiedener Individuen derselben Art analysieren. Im Gegensatz dazu analysieren wir mit unseren Verfahren Datensätze der DNA-Sequenz verschiedener Arten. Dieser Unterschied ist entscheidend, da die genetische Variation innerhalb einer Art für einige Aspekte des Phänotyps zu vernachlässigen ist oder uninteressant sein kann, wohingegen es sehr viel wahrscheinlicher ist, dass wichtige genetische Veränderungen gerade aus der Divergenz zwischen den verschiedenen Arten resultieren.

In meinem Vortrag gebe ich Ihnen einen Überblick über den statistischen Ansatz und zeige dessen Stärken und Schwächen auf. Ich werde mich dabei auf die Frage konzentrieren, wie wir den Ansatz angewendet haben, um die Einflüsse von drei-dimensionalen Proteinstrukturen und RNA-Sekundärstrukturen auf die Veränderung der DNA-Sequenz zu quantifizieren. Ich schließe meinen Vortrag mit einer Diskussion möglicher zukünftiger Richtungen für diesen Forschungszeitweig.

---

### PUBLIKATIONEN AUS DER FELLOWBIBLIOTHEK

Thorne, Jeffrey L. (2007)

Population genetics without intraspecific data

<https://kxp.k10plus.de/DB=9.663/PPNSET?PPN=1047193876>

Thorne, Jeffrey L. (2007)

Quantifying the impact of protein tertiary structure on molecular evolution

<https://kxp.k10plus.de/DB=9.663/PPNSET?PPN=1047193353>