Stochastische Modelle der Evolution

Matthias Birkner

https://www.staff.uni-mainz.de/birkner/

Hauptseminar, WS 2025/26

Stochastische Modelle helfen, die genetische Variabilität von Populationen zu verstehen und die zugrundeliegenden evolutionären „Kräfte“, wie

- genetische Drift
- Selektion
- Mutation
- Rekombination
- räumliche Strukturierung

zu quantifizieren.

Ein zentrales Thema der Theorie ist dabei das Zusammenspiel zwischen Vorwärts-Zeitentwicklung der Population und der Rückwärts-Sicht auf die Genealogien von Stichproben.

Ein Beispiel

Robert L. Dorit, Hiroshi Akashi und Walter Gilbert, *Science* 268, 1183–1185 (1995):

Absence of Polymorphism at the *ZFY* Locus on the Human Y Chromosome

Robert L. Dorit,* Hiroshi Akashi, Walter Gilbert

DNA polymorphism in the Y chromosome, examined at a 729–base pair intron located immediately upstream of the *ZFY* zinc-finger exon, revealed no sequence variation in a worldwide sample of 38 human males. This finding cannot be explained by global constraint on the intron sequence, because interspecific comparisons with other nonhuman primates revealed phylogenetically informative sequence changes. The invariance likely results from either a recent selective sweep, a recent origin for modern *Homo sapiens*, recurrent male population bottlenecks, or historically small effective male population sizes. A coalescence model predicts an expected time to a most recent common ancestral

- Weltweite Stichprobe von 38 Männern (*homo sapiens*)
- jeweils ein (729 BP langes, nicht-kodierendes) Stück des Y-Chromosoms wurde sequenziert
- keinerlei Mutationen (alle 38 Stichproben identisch),

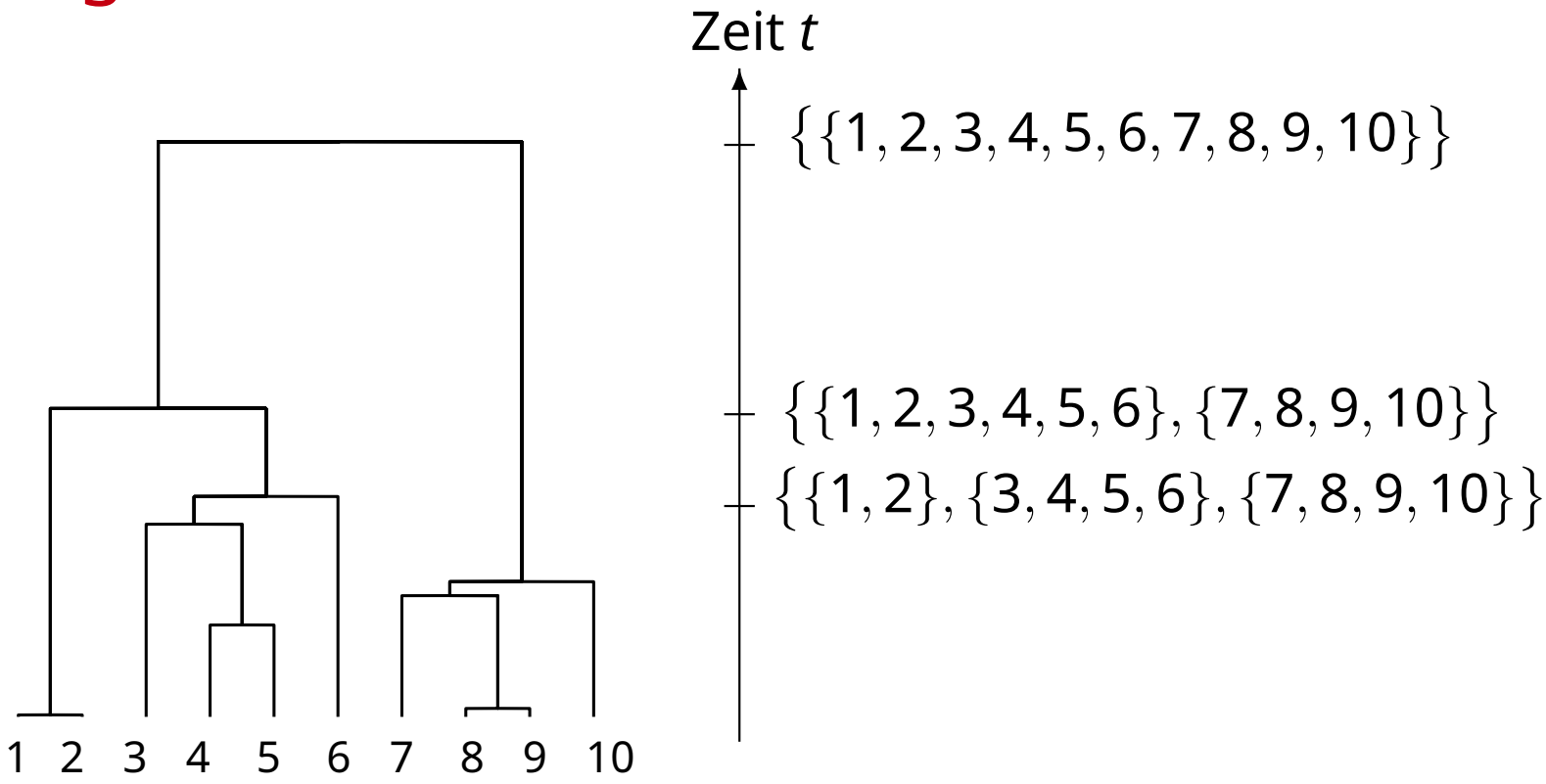
  aber prinzipiell möglich (Inter-Spezies-Vergleich mit Schimpanse, Gorilla, Orang-Utan)
- Mutationsrate $\approx 1{,}35 \times 10^{-3}$ (Mutationen pro Basenpaar pro Million Jahre)
  (Molekulare Uhr-Annahme, Fossilbefund)

Folgerung:

Der jüngsten gemeinsamen Vorfahren der gezogenen 38 Y-Chromosomen lebte vor ca. 195.000 Jahren

(95%-Kofidenzintervall: [74.000a, 404.000a])

Kingmans Koaleszent



Markovkette auf
$\mathcal{E}_n := \{\text{Partitionen von } \{1,2,\ldots,n\}\}$
Sprungraten

$$q_{\xi\eta} = \begin{cases} 1 & \text{falls Paarverschmelzung} \\ -\binom{|\xi|}{2} & \text{falls } \eta = \xi \\ 0 & \text{sonst} \end{cases}$$

$$\mathbb{P}\left(\tau_{\mathrm{jgV}}^{(n)} > t\right) = \sum_{i=2}^{n} \exp\left(-\binom{i}{2}t\right) \prod_{j=2, j\neq i}^{n} \frac{\binom{j}{2}}{\binom{j}{2} - \binom{i}{2}}$$

Voraussetzungen

- Fortgeschrittene Kenntnisse in Stochastik:
  Stochastik I,
  idealerweise Vertiefung in Stochastik
- Interesse an Anwendungen in Evolutionsbiologie

Termin

Blockseminar, voraussichtlich Januar 2026

(genaue Termine verhandelbar)

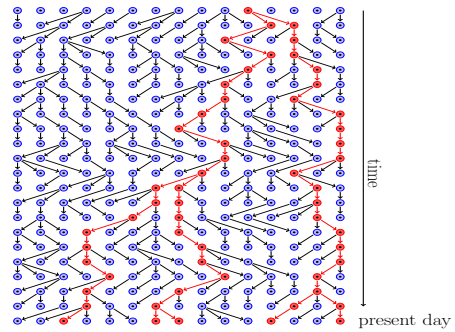
Stochastic population models[1]

Dozent/in: Prof. Dr. Matthias Birkner

Termine[2]: Thu, Fri 12-14

There is a lot of variability – both genetic and phenotypic – in real populations. Stochastic models can help to understand this and to quantify underlying evolutionary mechanisms like selection, mutation, genetic drift, recombination or spatial structure. They also form the basis for inference of biological mechanisms and their parameters based on observed genetic variability in sampled individuals. A central thread of the course will be the interplay between the forwards in time dynamics of the population and the backwards in time view on the genealogy of samples.

Some exemplary (technical[3], but see notes) keywords: Wright-Fisher model and diffusion, coalescents, Ewens' sampling formula, ancestral selection and recombination graphs, branching processes, stepping stone model, interacting particle systems



Schematic representation of a randomly reproducing population and a present-day sample together with its genealogy (in red)

Literature:

R. Durrett, *Probability Models for DNA Sequence Evolution*, Springer (2008).
W. Ewens, *Mathematical population genetics*, Springer (2004).
J. Wakeley, *Coalescent Theory: An Introduction*, Roberts & Company (2008).
M. Birkner, *Stochastische Modelle der Populationsbiologie*, Vorlesungsskript, JGU Mainz (2016) `https://www.staff.uni-mainz.de/birkner/SMPB1516/smpb1516.pdf`
S. Ethier, T. Kurtz, *Markov processes: characterization and convergence*, Wiley (1996).

Notes 1. The language can to some extent be negotiated at the beginning. A mixture of German and English is also an option.
2. Brown bag: Given the times, it is perfectly fine to bring a (small) lunch to lectures.
3. The principal ideas and phenomena can be understood and appreciated using notions from "Grundlagen der Stochastik". Advanced tools, e.g. diffusion processes or stochastic differential equations, may feature occasionally but will be motivated and illustrated by discrete approximations (and this course is of course also an invitation to learn about such objects). From the introduction of W. Feller's famous book on probability: "The traveler often has the choice between climbing a peak or using a cable car."
4. M.Ed. students (with our without biology as a second subject) are very welcome.

Quelle(n)

- Artikel aus der Fachliteratur
- Alison Etheridge, *Some mathematical models from population genetics*, École d'Été de Probabilités de Saint-Flour XXXIX, Springer, 2011
- Rick Durrett, *Probability Models for DNA Sequence Evolution*, 2nd ed., Springer, 2008
- M.B., Vorlesungsnotizen zu *Stochastische Modelle der Populationsbiologie*, WS 2015/16
  `https://www.staff.uni-mainz.de/birkner/SMPB1516/smpb1516.pdf`

Vorbesprechung und Themenvergabe

Mo., 25.8.2025, 14 c.t.

(ggf. hybrid möglich)