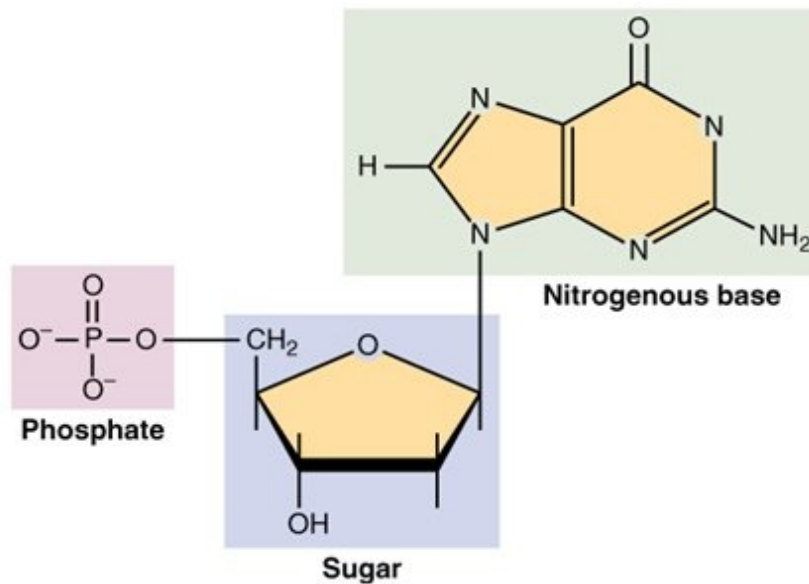
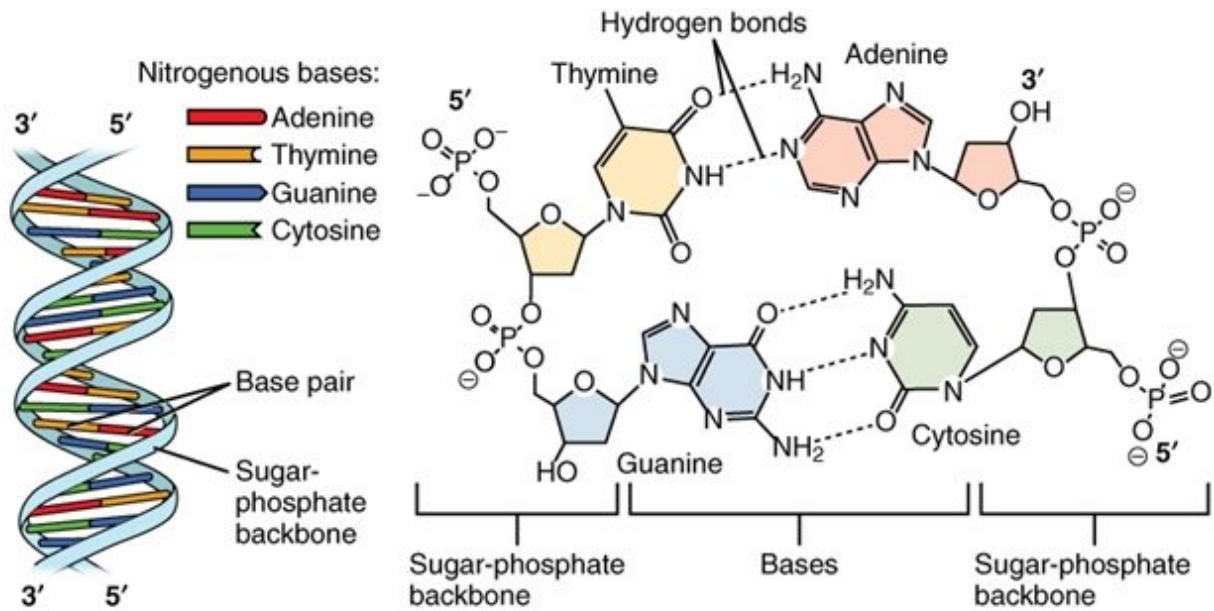


# Statistical test relates pathogen mutation to infectious disease progression

December 28 2017, by Lina Sorg



Ryosuke Omori and Jianhong Wu develop an inductive algorithm to study site-specific nucleotide frequencies using a multi-strain susceptible-infective-removed (SIR) model to better understand infectious disease epistemology, pathogen evolution, and population dynamics. Credit: Wikimedia Commons.

Nucleic acid sequencing methods, which determine the order of nucleotides in DNA fragments, are rapidly progressing. These processes yield large quantities of sequence data—some of which is dynamic—that helps researchers understand how and why organisms function like they do. Sequencing also benefits epidemiological studies, such as the identification, diagnosis, and treatment of genetic and/or contagious diseases. Advanced sequencing technologies reveal valuable information about the time evolution of pathogen sequences. Because researchers can estimate how a mutation behaves under the pressure of natural selection, they are thus able to predict the impact of each mutation—in terms of survival and propagation—on the fitness of the pathogen in question. These predictions lend insight to infectious disease epistemology, pathogen evolution, and population dynamics.

In a paper published earlier this month in the *SIAM Journal on Applied Mathematics*, Ryosuke Omori and Jianhong Wu develop an inductive algorithm to study site-specific nucleotide frequencies using a multi-strain susceptible-infective-removed (SIR) model. A SIR model is a simple compartmental model that places each individual in a [population](#) at a given time into one of the three aforementioned categories to compute the theoretical number of people affected by an infectious disease. The authors use their algorithm to calculate Tajima's D, a popular statistical test that measures natural selection at a specific site by analyzing differences in a sample of sequences from a population. In a non-endemic situation, Tajima's D can change over time. Investigating the time evolution of Tajima's D during an outbreak allows researchers

to estimate mutations relevant to pathogen fitness. Omori and Wu aim to understand the impact of disease dynamics on Tajima's D, thus leading to a better understanding of a mutation's pathogenicity, severity, and host specificity.

The sign of Tajima's D is determined by both natural selection and [population dynamics](#). "Tajima's D equals 0 if the evolution is neutral—no natural selection and a constant [population size](#)," Omori said. "A nonzero value of Tajima's D suggests natural selection and/or change in population size. If no natural selection can be assumed, Tajima's D is a function of the population size. Hence, it can be used to estimate time-series changes in population size, i.e., how the epidemic proceeds."

Differential equations, which model the rates of change of the numbers of individuals in each model compartment, can describe population dynamics. In this case, the population dynamics of hosts infected with the strain carrying a given sequence are modeled by a set of differential equations for that sequence, which include terms describing the [mutation rate](#) from one sequence to another. When setting up their multi-strain SIR model, Omori and Wu assume that the population dynamics of the pathogen is proportional to the disease dynamics. i.e., the number of pathogens are proportional to the number of infected hosts. This assumption allows the value of Tajima's D to change.

In population genetics, researchers believe that the sign of Tajima's D is affected by population dynamics. However, the authors show that in the case of a SIR deterministic model, Tajima's D is independent of the disease dynamics (specifically, independent of the parameters for disease transmission rate and disease recovery rate). They also observe that while Tajima's D is often negative during an outbreak's onset, it frequently becomes positive with the passage of time. "The negative sign does not imply an expansion of the infected population in a deterministic

model," Omori said. "We also found the dependence of Tajima's D on the disease transmission dynamics can be attributed to the stochasticity of the transmission dynamics at the population level. This dependence is different from the aforementioned existing assumption about the relation between population dynamics and the sign of Tajima's D."

Ultimately, Omori and Wu prove that Tajima's D in a deterministic SIR model is completely determined by mutation rate and sample size, and that the [time evolution](#) of an infectious disease pathogen's genetic diversity is fully determined by the mutation rate. "This work revealed some dependence of Tajima's D on the (disease transmission dynamics) basic reproduction number ( $R_0$ ) and mutation rate," Omori said. "With the assumption of neutral evolution, we can then estimate mutation rate or  $R_0$  from sequence data."

Given the demand for tools that analyze evolutionary and disease dynamics, the observation that Tajima's D depends on the stochasticity of the dynamics is useful when estimating epidemiological parameters. For example, if sequences of pathogens are sampled from a small outbreak in a limited host population, then Tajima's D depends on both the mutation rate and  $R_0$ ; therefore, a joint estimate of these parameters from Tajima's D is possible. "We are applying this theoretic result to analyze real-world epidemiological data," Omori said. "We should also see if our approach can be used to investigate non-equilibrium [disease](#) dynamics with [natural selection](#)."

**More information:** Omori, R., & Wu, J. (2017). Tajima's D and Site-specific Nucleotide Frequency in a Population during an Infectious Disease Outbreak. *SIAM Journal on Applied Mathematics*, 77(6), 2156-2171.

Provided by Society for Industrial and Applied Mathematics

Citation: Statistical test relates pathogen mutation to infectious disease progression (2017, December 28) retrieved 27 April 2024 from <https://phys.org/news/2017-12-statistical-pathogen-mutation-infectious-disease.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.