# Herbarium-Genomik: Erforschung pflanzlicher und mikrobieller Evolution unter Verwendung von historischen und neuen Proben

# Herbarium genomics: investigating plant and microbial evolution using historic and modern samples

Burbano, Hernán A.

Max-Planck-Institut für Entwicklungsbiologie, Tübingen

Korrespondierender Autor/in

E-Mail: hernan.burbano@tuebingen.mpg.de

## Zusammenfassung

Heute lassen sich Genome aus Organismen rekonstruieren, die vor Tausenden von Jahren gelebt haben. In Museen lagern Millionen von getrockneten Pflanzenteilen, deren DNA sich für die Genomsequenzierung eignet. Die Pflanzenproben enthalten unentdeckte Nachweise für die globale Diversität der letzten rund 450 Jahre, und die kombinierte Verwendung von historischen und aktuellen Pflanzenproben führt so eine neue, zeitliche Dimension in die Evolutionsforschung ein. Untersucht wird die Evolution von Pflanzen, ihren Pathogenen und das dazugehörige Mikrobiom unter Verwendung von neuer und alter DNA.

## Summary

Today it is possible to retrieve genomes from organisms that perished thousands of years ago. Millions of desiccated plant specimens are stored in museums and contain DNA suitable for genome sequencing. These specimens harbor an untapped record of global biodiversity spanning the last 450 years. The combined use of historical and modern plant samples introduces a temporal scale to evolutionary studies, allowing the interrogation of allele frequencies through time. The research group studies the evolution of plants, their pathogens, and the associated microbiome using modern and ancient DNA.

## Einführung

Die Evolutionsforschung beschäftigt sich mit der Entstehung und Ausbreitung der Biodiversität auf einem Planeten, der reich ist an ökologischen Nischen und Ressourcen, aber auch voller Feinde, Widrigkeiten und Konkurrenten. Wichtige Aspekte der Evolution, wie etwa die Vielfalt und die Erhaltung der Arten, sowie Umweltveränderungen können in unterschiedlichen Zeitskalen untersucht werden. Die moderne Evolutionsforschung kann daher aktuell auch als eine Mischung aus Geschichte und Biologie betrachtet werden; sie hat sich zum Ziel gesetzt, die Beziehungen und die Dynamik von Populationen innerhalb und zwischen verschiedenen Arten im Zeitverlauf zu erforschen. Typischerweise erfolgt dies entweder durch die direkte Untersuchung von Fossilien, historischen Belegen oder durch Echtzeitbeobachtung von noch bestehenden Populationen. Alternativ lassen sich Spuren früherer Ereignisse durch den Vergleich von

© 2014 Max-Planck-Gesellschaft 

Genomen eng verwandter Organismen verfolgen.

Üblicherweise wurde die Vergangenheit stets auf der Grundlage heutiger Daten abgeleitet. Der Nachteil dabei ist aber, dass der Zeitpunkt wichtigen Evolutionsgeschehens äußerst ungenau bleibt. Will man wiederum vergangene Ereignisse genauer auswerten, so kann man indirekt eine Zeitreise zurück in die Vergangenheit antreten, indem man DNA aus Organismen gewinnt, die in der Vergangenheit gelebt haben. Das DNA-Material, das aus Museumsexemplaren, Fossilrückständen und archäologischen Funden gewonnen wird und als *ancient* DNA oder antike DNA (aDNA) bezeichnet wird, hat eine faszinierende, neue Dimension der Evolutionsforschung eröffnet [1]. Weltweit gibt es mehr als 3.000 Herbarien [2], jedoch haben Evolutionsbiologen, die sich mit der Genomik beschäftigen, diese wertvollen Sammlungen bisher ignoriert, obwohl in dieser globalen Datenbank rund 450 Jahre Pflanzenevolution gespeichert sind!

Überall in der Welt dokumentieren Sammlungen zahlreiche faunistische und botanische Expeditionen des 18. und 19. Jahrhunderts. Die berühmteste Expedition war vielleicht die Reise auf dem Schiff Beagle und die Finkensammlung, die Darwin von den Galapagos-Inseln mitbrachte. Während seiner fünfjährigen Expeditionszeit hat Darwin aber nicht nur Tiere gesammelt und untersucht, sondern auch eine riesige Zahl von botanischen Exemplaren erforscht. Die historischen Sammlungen von Darwin und anderen warten nun auf genetische Untersuchungen und Vergleiche mit der heute bestehenden genetischen Diversität.

In diesem Artikel werden unsere neuesten Ergebnisse zur Evolution von Pflanzen und Pflanzenpathogenen aus Studien vorgestellt, in denen sowohl aktuelle als auch historische Pflanzenproben untersucht wurden. Darüber hinaus werden die Herausforderungen und Chancen des sich entwickelnden Forschungsfeldes der Herbarium-Genomik erläutert.

## Antike DNA: Tierfossilien und getrocknetes Pflanzengewebe im Vergleich

Antike DNA ist typischerweise eine komplexe DNA-Mischung, die Moleküle aus den Artenproben (endogene DNA) und mikrobielle DNA enthält, die entweder bereits zum Zeitpunkt der Sammlung vorhanden war oder das Gewebe *post mortem* kolonisiert hat. Die biochemischen Signaturen der antiken DNA (aDNA) unterscheiden sich von derjenigen DNA, die aus frisch gesammelten Gewebeproben stammt. Insbesondere ist die aDNA stark fragmentiert und enthält biochemische Veränderungen, die während des Sequenzierens zum Fehleinbau von Nukleotiden führen. Die Definition von aDNA berücksichtigt diese Merkmale, beschreibt aber keine Zeitgrenze zwischen antiker und moderner DNA [1]. Deshalb bezeichnen wir DNA, die aus jahrhundertealten Pflanzenproben rekonstruiert wurde, trotz ihres evolutionär jungen Alters als aDNA.

Obwohl aDNA aus Herbarproben stark fragmentiert ist, entfällt im Allgemeinen ein Großteil auf endogene DNA, sodass sie sich für die direkte Sequenzierung eignet. Außerdem können im Verlauf der genomischen *Library Preparation* biochemische Schäden in der aDNA repariert werden. Zudem wurden inzwischen bioinformatische Algorithmen entwickelt, mit denen sich aDNA-Sequenzen mit sehr kurzen Fragmentlängen analysieren lassen. Die Haupteinschränkung kurzer Fragmentlängen besteht jedoch in der Schwierigkeit von *de novo* Assemblierungen antiker Genome. Dennoch ist es heute möglich, vollständige Genome aus historischen Proben in einer Qualität darzustellen, die mit den Genomdaten aus modernen Proben konkurrieren kann.

## Suche nach Pflanzenpathogenen in Herbarien

Interessanterweise haben die Botaniker nicht nur gesundes Pflanzenmaterial gesammelt, sondern sie haben oft auch Exemplare mit Läsionen, die charakteristisch für Erregerinfektionen sind, in ihre Sammlungen

mitaufgenommen. So liefert das Herbariummaterial ganz neue Perspektiven auf die Pflanzenepidemien der Vergangenheit. Es entstehen faszinierende neue Möglichkeiten, Pflanze-Pathogen-Interaktionen durch die gleichzeitige Untersuchung beider Genome zu erforschen [3]. Mit diesem Ansatz wurden mehrere Genome des Pathogens *Phytophthora infestans* aus Proben aus dem 19. Jahrhundert rekonstruiert, die ursprünglich an verschiedenen geografischen Standorten gesammelt wurden [4]. *P. infestans* ist der ursächliche Erreger der Kraut- und Knollenfäule der Kartoffel, die einer der Auslöser für die große irische Hungersnot Mitte des 19. Jahrhunderts war. Durch einen Vergleich von aktuellen und historischen Proben konnte die Beziehung zwischen einzelnen Erregerstämmen nachgewiesen werden.

Unsere Studien ergaben, dass ein einziger Stamm, nämlich HERB-1, für die Epidemie im 19. Jahrhundert verantwortlich war. Die klonale Abstammungslinie HERB-1, eine Schwestergruppe des in den 1980er-Jahren am weitesten verbreiteten Stammes, konnte bisher unter den im 20. Jahrhundert gesammelten Erregern noch nicht identifiziert werden. Das lässt vermuten, dass sie heute ausgelöscht oder zumindest sehr selten geworden ist [4]. Aufstieg und Fall von HERB-1 stimmen mit der historischen Dynamik von *P. infestans* in der gesamten Welt überein [3]. Die Erforschung der historischen Dynamik von *P. infestans* und anderen Pflanzenpathogenen wird uns somit helfen, sich gegen künftige Epidemien zu rüsten.

Darüber hinaus konnten anhand von *P. infestans*-Proben, die vor über einem Jahrhundert gesammelt wurden, die Veränderungen in Genen mit einer Schlüsselfunktion für den Infektionsprozess identifiziert werden. Sehr wahrscheinlich sind diese Veränderungen und der dafür zugrundeliegende selektive Druck durch das Aufkommen der modernen Landwirtschaft und durch das Einkreuzen von Resistenzgenen in Wirtspflanzen mittels traditioneller Pflanzenzuchtmethoden entstanden [4]. In den letzten 200 Jahren gab es zahlreiche Epidemien durch Pflanzenpathogene. In künftigen Studien werden sich Herbarproben, die mit diesen Zeiträumen in Verbindung gebracht werden können, dazu nutzen lassen, die mit diesen katastrophalen Ereignissen assoziierten Genomveränderungen zu erforschen.

## Rückverfolgung von Pflanzenkolonisationen und -invasionen

Pflanzen sind zwar sessile Lebewesen. Ihre Saat kann aber, besonders durch anthropogene Verbreitung, weite, sogar interkontinentale Reisen zurücklegen. Besonders interessant sind Erkenntnisse darüber, wie sich eine Art an eine neue Umgebung anpasst und welche Faktoren über ihren Erfolg als Kolonisator entscheiden. *Arabidopsis thaliana*, ein bedeutender Modellorganismus der Pflanzenbiologie, hat sich gut an viele verschiedene ökologische Bedingungen angepasst, was in einem breiten Spektrum an genetischer und phänotypischer Variation resultiert. Dieser Modellorganismus kann zur Erforschung des Evolutionsprozesses der Kolonisation dienen. In diesem Zusammenhang wurden im Rahmen des „1001 *A. thaliana* Genomprojekts" über 1.000 Genome aus vorhandenen *A. thaliana-Accessions* (Stämmen) sequenziert [5].

© 2014 Max-Planck-Gesellschaft

**Abb. 1:** Herbariumexemplare von Wildpopulationen der Pflanze *Arabidopsis thaliana* aus Nordamerika. Sammler: Jane Devos und Gautam Shirsekar, Max-Planck-Institut für Entwicklungsbiologie, Tübingen.
© Max-Planck-Institut für Entwicklungsbiologie/Abendroth

Wie viele andere Pflanzen, siedelte sich *A. thaliana* in jüngster Geschichte in Nordamerika an. Heute kommt in nordamerikanischen Populationen ein dominanter Genotyp vor, der als HPG1 bezeichnet wird [6]. Die Verwendung von historischen Exemplaren aus dem kolonisierten Umfeld in Nordamerika in Verbindung mit jetzt gesammelten Stämmen macht es möglich, die Dynamik dieses Genotyps im Zeitverlauf zu untersuchen (**Abb. 1**). Mit den gesammelten Daten aus den historischen Proben, die als interne Kalibrierpunkte in einem phylogenetischen Rahmen dienen, ist es möglich, langfristige Mutationsraten zu berechnen, dann das Kolonisationsereignis zeitlich zu bestimmen und die Divergenz zwischen verschiedenen Abstammungslinien und Unterlinien von *A. thaliana* in Nordamerika auszuwerten. Zudem lassen sich mit diesen Informationen die Änderungen in der Populationsgröße im Zeitverlauf abschätzen und die phylogeographischen Muster der Arten seit ihrer Ankunft in Amerika rekonstruieren.

## Ausblick: Metagenomik historischer Pflanzenarten

Da ein erheblicher Teil der aus den Herbarproben abgeleiteten Sequenzstücke nicht von den gesammelten Pflanzen, sondern von Mikroorganismen stammt, eröffnet die Klassifizierung dieser Sequenzen die Möglichkeit, das Metagenom historischer Arten zu erforschen. Dieser Ansatz kann der Identifizierung von Pathogenen in Herbarproben ohne offensichtliche Läsionen, die auf Infektionen hinweisen, dienen. Ebenso kann man das Vorkommen verschiedener Mikroorganismen im Kontext gesunder und infizierter Pflanzen oder in Proben derselben Arten, die in verschiedenen Umgebungen gesammelt wurden, untersuchen. Die Erforschung des Metagenoms ist mit enormen Herausforderungen verbunden. Die größten Hürden ergeben sich aus dem Bedarf an einer rechnerisch effizienten und exakten Annotation der Metagenomik-*Reads* sowie der Entwicklung von Methoden zur Unterscheidung zwischen dem *bona fide*-Metagenom, das während der Sammelzeit vorhanden war, und Mikroorganismen, die *post mortem* im Gewebe kolonisieren. Die Verwendung neuer Algorithmen für die schnelle Annotation von Metagenom-*Reads* mit Hilfe leistungsstarker Computer sowie die präzise Identifizierung von DNA-Fragmenten werden die weitere Ausschöpfung einer bisher unzureichend genutzten historischen Ressource weiter voranbringen.

**Literaturhinweise**

[1] **Shapiro, B.; Hofreiter, M.**

**A paleogenomic perspective on evolution and gene function: new insights from ancient DNA**
Science 343: 1236573 (2014)
DOI: 10.1126/science.1236573

[2] **Thiers, B.**

**Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium**
http://sweetgum.nybg.org/science

[3] **Yoshida, K.; Burbano, H. A.; Krause, J.; Thines, M.; Weigel, D.; Kamoun, S.**

**Mining herbaria for plant pathogen genomes: back to the future**
PLoS Pathogens 10: e1004028 (2014)
DOI: 10.1371/journal.ppat.1004028

[4] **Yoshida, K.; Schuenemann, V. J.; Cano, L. M.; Pais, M.; Mishra, B.; Sharma, R.; Lanz, C.; Martin, F. N.; Kamoun, S.; Krause, J.; Thines; M.; Weigel, D.; Burbano, H. A.**

**The rise and fall of the *Phytophthora infestans* lineage that triggered the Irish potato famine**
eLife 2: e00731 (2013)
DOI: 10.7554/eLife.00731

[5] **Weigel, D.; Mott, R.**

**The 1001 genomes project for *Arabidopsis thaliana***
Genome Biology 10: 107 (2009)
DOI: 10.1186/gb-2009-10-5-107

[6] **Hagmann, J.; Becker, C.; Müller, J.; Stegle, O.; Meyer, R. C.; Wang, G.; Schneeberger, K.; Fitz, J.; Altmann, T.; Bergelson, J.; Borgwardt, K.; Weigel, D.**

**Century-scale methylome stability in a recently diverged *Arabidopsis thaliana* lineage**
PLoS Genetics 11: e1004920 (2015)
DOI: 10.1371/journal.pgen.1004920

© 2014 Max-Planck-Gesellschaft 