

CHÚNG TA LÀ AI

VÀ CHÚNG TA ĐÃ
ĐẾN ĐÂY THẾ NÀO?

WHO WE ARE AND HOW WE GOT HERE

Copyright © 2018 by David Reich and Eugenie Reich. All rights reserved.

CHÚNG TA LÀ AI VÀ CHÚNG TA ĐÃ ĐẾN ĐÂY THẾ NÀO?

DNA cổ đại và khoa học mới về quá khứ của loài người

Tác giả: David Reich

Bản quyền tiếng Việt © Công ty Cổ phần Sách Omega Việt Nam, 2026

Bản dịch tiếng Việt © Công ty Cổ phần Sách Omega Việt Nam

Công ty Cổ phần Sách Omega Việt Nam xuất bản, bìa mềm, tay gấp, 2026

Thiết kế và trình bày: Công ty Cổ phần Sách Omega Việt Nam

Công ty Cổ phần Sách Omega Việt Nam - thành viên của Alpha Publishing Group.

CÔNG TY CP SÁCH OMEGA VIỆT NAM (OMEGA PLUS) thành lập tháng 9/2016 với định hướng xuất bản tri thức nền tảng trong lĩnh vực Khoa học Nhân văn, Khoa học Tự nhiên, Chính trị - Xã hội, Triết học, Nghệ thuật... Đến nay, chúng tôi đã xuất bản gần 500 đầu sách thuộc 7 tủ sách, bao gồm: Kinh điển, Sử Việt, Khoa học, Nhân vật Vĩ đại, Nghệ thuật, Tủ sách Đời người và Tủ sách Y sinh. Thông qua việc xuất bản các ấn phẩm có giá trị, mang hàm lượng tri thức cao, qua các hoạt động nhằm đưa tri thức hữu ích đến cộng đồng, Omega Plus mong muốn đóng góp theo cách riêng vào sự phát triển đang ngày càng mạnh mẽ của xã hội.

Không phần nào thuộc sở hữu của Công ty Cổ phần Sách Omega Việt Nam và được chủ sở hữu nhượng quyền trong xuất bản phẩm này được phép sao chép, lưu trữ trong hệ thống truy xuất hoặc truyền đi dưới bất kỳ hình thức nào hoặc bằng bất kỳ phương tiện nào gồm điện tử, cơ khí, sao chụp, ghi âm hoặc mọi hình thức và phương tiện khác mà không có sự cho phép trước bằng văn bản của Công ty Cổ phần Sách Omega Việt Nam. Chúng tôi luôn mong muốn nhận được những ý kiến đóng góp của quý vị độc giả để sách ngày càng hoàn thiện hơn.

Biên mục trên xuất bản phẩm của Thư viện Quốc gia Việt Nam

Reich, David

Chúng ta là ai và chúng ta đã đến đây thế nào? : DNA cổ đại và khoa học mới về quá khứ của loài người / David Reich ; Dịch: Trương Cao Minh... - H. : Dân trí ; Công ty Sách Omega Việt Nam, 2026. - 420 tr. ; 24 cm

1. Di truyền học người 2. Gen học 3. ADN

572.86 - dc23

DTF0643p-CIP

Góp ý về sách, liên hệ về bản thảo và bản dịch: info@omegaplus.vn

Liên hệ ebooks, hợp tác xuất bản & truyền thông trên sách: truyenthong@omegaplus.vn

Liên hệ dịch vụ bản quyền, xuất bản, ký gửi và phát hành: dichvu@omegaplus.vn

David Reich

CHÚNG TA LÀ AI

VÀ CHÚNG TA ĐÃ
ĐẾN ĐÂY THẾ NÀO?



**DNA cổ đại
và khoa học mới
về quá khứ
của loài người**

Trương Cao Minh, Nguyễn Hồng Hà,
Trần Tuấn Hiệp *dịch*

⊕ MEGA⁺  NHÀ XUẤT BẢN
DÂN TRÍ

ĐỘI NGŨ TRIỂN KHAI OMEGA+

Chịu trách nhiệm xuất bản: Trần Hoài Phương

Phụ trách xuất bản: Hà Kim Ngân

Điều phối xuất bản: Nguyễn Hương

Phụ trách bản quyền: Thiên Hương

Biên tập viên: Ngô Tuấn, Ái Nhi

Thiết kế bìa: Hoàng Duy Khánh

Trình bày: Vũ Lê Thư

Thư ký xuất bản: Bùi Huyền

ĐƠN VỊ HỢP TÁC TRUYỀN THÔNG



<https://www.facebook.com/tramdocvn>
website: <http://tramdoc.vn/>

Mục lục

Lời cảm ơn.....	9
Mở đầu.....	11
PHẦN I. LỊCH SỬ XA XƯA CỦA CHÚNG TA	
Chương 1 Cách hệ gen cho biết chúng ta là ai	31
Chương 2 Những lần gặp mặt với người Neanderthal	55
Chương 3 DNA cổ đại mở ra kho thông tin	87
PHẦN II. CÁCH CHÚNG TA HIỆN HỮU NHƯ NGÀY NAY	
Chương 4 “Những bóng ma” của loài người.....	115
Chương 5 Sự hình thành châu Âu hiện đại.....	139
Chương 6 Cuộc va chạm hình thành nên Ấn Độ	167
Chương 7 Đi tìm tổ tiên của người châu Mỹ bản địa.....	203
Chương 8 Nguồn gốc di truyền của người Đông Á.....	243
Chương 9 Châu Phi tái xuất trong câu chuyện về loài người.....	265

PHẦN III. HỆ GEN VÀ TÁC ĐỘNG XÃ HỘI

Chương 10	Nghiên cứu di truyền về bất bình đẳng	288
Chương 11	Nghiên cứu di truyền về chủng tộc và bản sắc	308
Chương 12	Tương lai của DNA cổ đại	341
Ghi chú về hình minh họa.....		357
Chú thích		362
Mục từ tra cứu.....		409

Dành tặng Seth và Leah

Lời cảm ơn

Đầu tiên, tôi xin dành lời cảm ơn tới vợ tôi, Eugenie Reich. Cuốn sách này ra đời sau một năm làm việc tất bật với cô ấy. Chúng tôi cùng soạn dàn ý, chuẩn bị những bản nháp đầu tiên của các chương, và không ngừng nói về cuốn sách khi nó được hoàn thiện. Cuốn sách này sẽ không thể thành hình được nếu không có cô ấy.

Tôi cũng xin gửi lời biết ơn tới Bridget Alex, Peter Bellwood, Samuel Fenton-Whittet, Henry Louis Gates Jr., Yonatan Grad, Iosif Lazaridis, Daniel Lieberman, Shop Mallick, Erroll McDonald, Latha Menon, Nick Patterson, Molly Przeworski, Juliet Samuel, Clifford Tabin, Daniel Reich, Tova Reich, Walter Reich, Robert Weinberg và Matthew Spriggs về những góp ý mang tính xây dựng cho toàn bộ cuốn sách.

Tôi cảm ơn David Anthony, Ofer Bar-Yosef, Caroline Bearsted, Deborah Bolnick, Dorcas Brown, Katherine Brunson, Phan Xảo Mai, David Goldstein, Alexander Kim, Carles Lalueza-Fox, Iain Mathieson, Eric Lander, Mark Lipson, Scott MacEachern, Richard Meadow, David Meltzer, Priya Moorjani, John Novembre, Svante Pääbo, Pier Palamara, Eleftheria Palkopoulou, Mary Prendergast, Rebecca Reich, Colin Renfrew, Nadin Rohland, Daniel Rozas, Pontus Skoglund, Vương Truyền Siêu, và Michael Witzel đã góp ý cho các chương riêng lẻ. Tôi cũng cảm ơn Stanley Ambrose, Graham Coop, Dorian Fuller, Éadaion Harney, Linda

Heywood, Yousuke Kaifu, Kristian Kristiansen, Michelle Lee, Daniel Lieberman, Michael McCormick, Michael Petraglia, Joseph Pickrell, Stephen Schiffels, Beth Shapiro và Bence Viola đã đánh giá các phần của cuốn sách cho chính xác nhất.

Tôi biết ơn Trường Y Harvard, Viện Y khoa Howard Hughes và Quỹ Khoa học Quốc gia, tất cả đều hào phóng hỗ trợ về các vấn đề khoa học khi tôi viết cuốn sách này, và xem nó như là sự bổ sung cho nghiên cứu ban đầu của tôi.

Cuối cùng tôi xin cảm ơn những người đã nhiều lần khuyến khích tôi viết cuốn sách này. Tôi đã gạt bỏ ý tưởng đó trong nhiều năm vì không muốn bị phân tâm khỏi việc nghiên cứu khoa học, và vì với các nhà di truyền học, công bố khoa học là tiền bạc, không phải là sách. Nhưng suy nghĩ của tôi đã thay đổi khi các đồng nghiệp bao gồm các nhà khảo cổ, nhân học, sử học, ngôn ngữ học và những người khác ngày càng muốn nắm bắt được cuộc cách mạng DNA cổ đại. Có nhiều bài nghiên cứu tôi đã không viết, và nhiều bài phân tích tôi đã không hoàn thành, vì thời gian đó tôi dành để viết cuốn sách này. Mong rằng quý độc giả sẽ có một cái nhìn mới về bản chất con người chúng ta.

Mở đầu

Cuốn sách này được lấy cảm hứng từ một người có tầm nhìn xa, Luca Cavalli-Sforza, người sáng lập ra các nghiên cứu về gen trong quá khứ. Được đào tạo bởi một trong những học trò của ông, tôi là một phần của thế hệ kế nhiệm, được truyền cảm hứng từ tầm nhìn coi bộ gen như lăng kính soi chiếu lịch sử loài người.

Bước ngoặt trong sự nghiệp của Cavalli-Sforza đến vào năm 1994 khi ông xuất bản cuốn *The History and Geography of Human Genes* (Lịch sử và địa lý của gen người), tổng hợp những gì đã được biết đến tại thời điểm đó từ khảo cổ học, ngôn ngữ học, lịch sử học và di truyền học để kể một câu chuyện lớn về cách mà thế giới các dân tộc có được như ngày nay.¹ Cuốn sách là cái nhìn tổng quan về quá khứ xa xưa. Nhưng nó dựa trên những gì đã được biết vào thời điểm đó nên vẫn còn khiếm khuyết bởi dữ liệu di truyền sẵn có khi đó rất ít ỏi, hạn chế đến mức gần như vô dụng so với thông tin phong phú từ khảo cổ học và ngôn ngữ học. Dữ liệu di truyền thời đó đôi khi có thể tiết lộ các mô hình phù hợp với những gì đã được biết đến, nhưng thông tin mà nó cung cấp không đủ để chứng minh bất cứ điều gì thực sự mới. Thật vậy, một số tuyên bố mới to tát của Cavalli-Sforza về cơ bản đã bị chứng minh là sai. Hai thập niên trước, tất cả mọi người, từ Cavalli-Sforza đến những sinh viên mới bắt đầu tốt nghiệp như tôi, đang làm việc trong thời đại tăm tối của khoa học DNA.

Cavalli-Sforza đã có một tuyên bố lớn vào năm 1960 sẽ quyết định toàn bộ sự nghiệp. Ông đặt cược rằng có thể tái tạo những cuộc di cư vĩ đại trong quá khứ, hoàn toàn dựa trên sự khác biệt về gen giữa các dân tộc ngày nay.²

Từ nghiên cứu này đến nghiên cứu khác trong suốt năm thập niên sau đó, Cavalli-Sforza dường như đang tiến rất tốt trên con đường đã vạch ra. Khi ông bắt đầu công việc, công nghệ nghiên cứu sự đa dạng của con người còn quá lạc hậu nên cách duy nhất là đo các protein trong máu, sử dụng các biến thể như nhóm máu A, B và O được các bác sĩ kiểm tra để có sự tương đồng giữa người hiến máu và người nhận. Tính tới thập niên 1990, ông và đồng sự đã thu thập dữ liệu từ hơn một trăm biến thể như vậy trong vô vàn quần thể đa dạng. Họ dùng dữ liệu này để phân nhóm cá thể theo châu lục, một cách đáng tin cậy, dựa trên tần suất đối sánh với nhau ở những biến thể này: ví dụ, người châu Âu có tỉ lệ trùng khớp cao với người châu Âu vùng khác, người Đông Á với người Đông Á cũng vậy, người châu Phi với người châu Phi cũng thế. Trong những năm 1990 và 2000, họ đã đưa công việc này lên tầm cao mới bằng cách vượt ra ngoài sự biến đổi của protein và kiểm tra trực tiếp DNA, mã di truyền của chúng ta. Họ đã phân tích tổng cộng khoảng một nghìn cá thể từ khoảng năm mươi quần thể trải dài trên khắp hành tinh, kiểm tra sự biến đổi ở hơn ba trăm vị trí trong bộ gen.³ Khi họ sử dụng máy tính – vốn không có kiến thức về các nhân quần thể – để phân cụm những cá thể thành năm nhóm, các kết quả có độ trùng khớp cao với trực giác mà họ thường có về sự phân tách sâu sắc ở loài người (người Tây Á-Âu, người Đông Á, người châu Mỹ bản địa, người New Guinea và người Phi).

Cavalli-Sforza đặc biệt lưu tâm đến việc giải mã những cụm di truyền trong lịch sử những quần thể. Ông và đồng sự đã phân tích dữ liệu nhóm máu bằng cách sử dụng công nghệ giúp xác định sự kết hợp của các biến thể sinh học hiệu quả nhất trong việc tóm tắt sự khác biệt giữa những cá thể với nhau. Vẽ biểu đồ kết hợp các

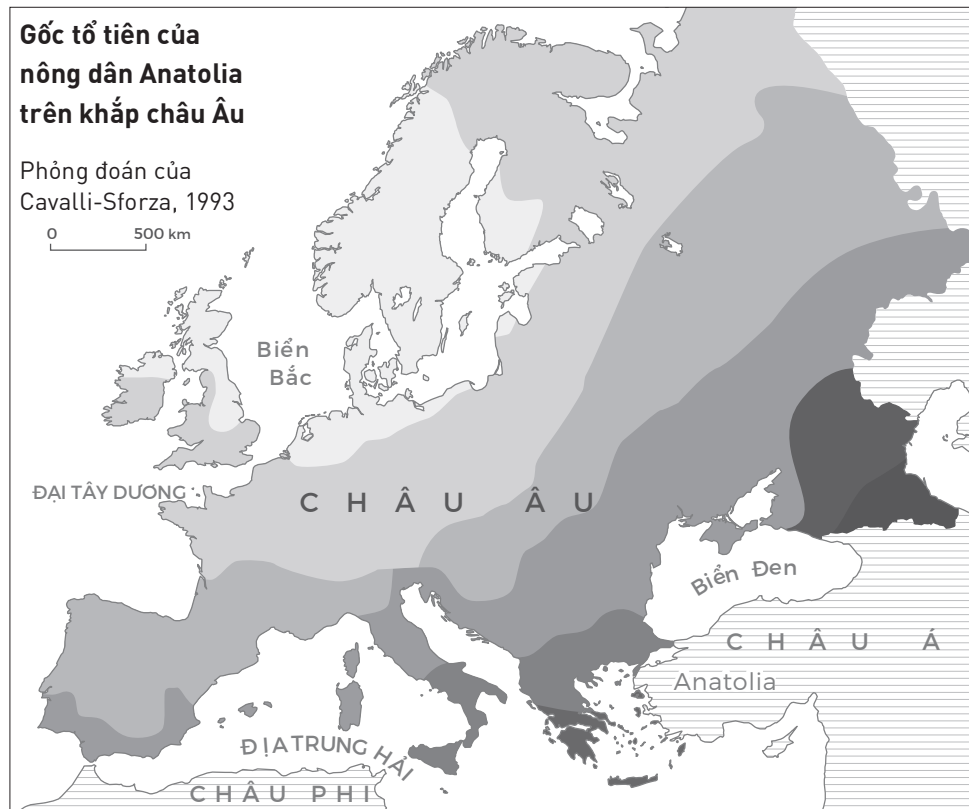
nhóm máu này lên bản đồ Tây Á-Âu, họ thấy khác biệt nhiều nhất giữa các quốc gia đạt giá trị cực đại ở vùng Cận Đông và giảm dần từ đông nam sang tây bắc khi hướng vào châu Âu.⁴ Họ giải thích đây là dấu vết di truyền của cuộc di cư của những nông dân từ vùng Cận Đông vào châu Âu, theo khảo cổ học là đã xảy ra từ 9.000 năm trước. Họ được gợi ý rằng sau khi đến châu Âu, những nông dân đầu tiên đã “hòa trộn” với những người săn bắt hái lượm địa phương, tích lũy thêm tổ tiên săn bắt hái lượm trong quá trình mở rộng lãnh thổ, một quá trình mà họ gọi là “sự phân tán diệu kỳ.”⁵ Cho đến gần đây, nhiều nhà khảo cổ thấy mô hình phân tán dân cư này là ví dụ cho việc kết hợp những hiểu biết sâu sắc từ khảo cổ học và di truyền học.

Mô hình mà Cavalli-Sforza và đồng sự đề xuất để phân tích dữ liệu rất hấp dẫn về mặt tư duy, nhưng tiếc là nó đã sai. Những sai sót bắt đầu trở nên rõ ràng vào năm 2008, khi John Novembre và đồng sự chứng minh các đường dốc quan sát được ở châu Âu có thể xuất hiện mà không cần sự di cư.⁶ Sau đó, theo một cách phản trực giác, họ cho thấy việc sản xuất nông nghiệp lan tỏa từ vùng Cận Đông sang châu Âu có thể làm phát sinh hiện tượng toán học mà Cavalli-Sforza sử dụng để tạo ra một đường dốc trực giao với hướng di cư chứ không song song với nó như đã thấy trong dữ liệu thực.⁷

Cuộc cách mạng được thực hiện nhờ khả năng phân lập DNA từ xương cổ đại – hay được biết đến với cái tên “cuộc cách mạng DNA cổ đại” – đã khép lại tấm màn về mô hình phân tán diệu kỳ. Cuộc cách mạng DNA cổ đại đã ghi lại rằng những nông dân đầu tiên ngay cả ở những vùng xa xôi nhất của châu Âu, như Anh, Scandinavia và Iberia, có rất ít gốc tổ tiên liên quan đến những người săn bắt hái lượm. Trên thực tế, họ có ít gốc tổ tiên săn bắt hái lượm hơn so với những quần thể đa dạng khác ở châu Âu ngày nay. Tỷ lệ tổ tiên nông dân ban đầu cao nhất ở châu Âu ngày nay không phải ở Đông Nam Âu, nơi mà Cavalli-Sforza cho rằng nó phổ biến

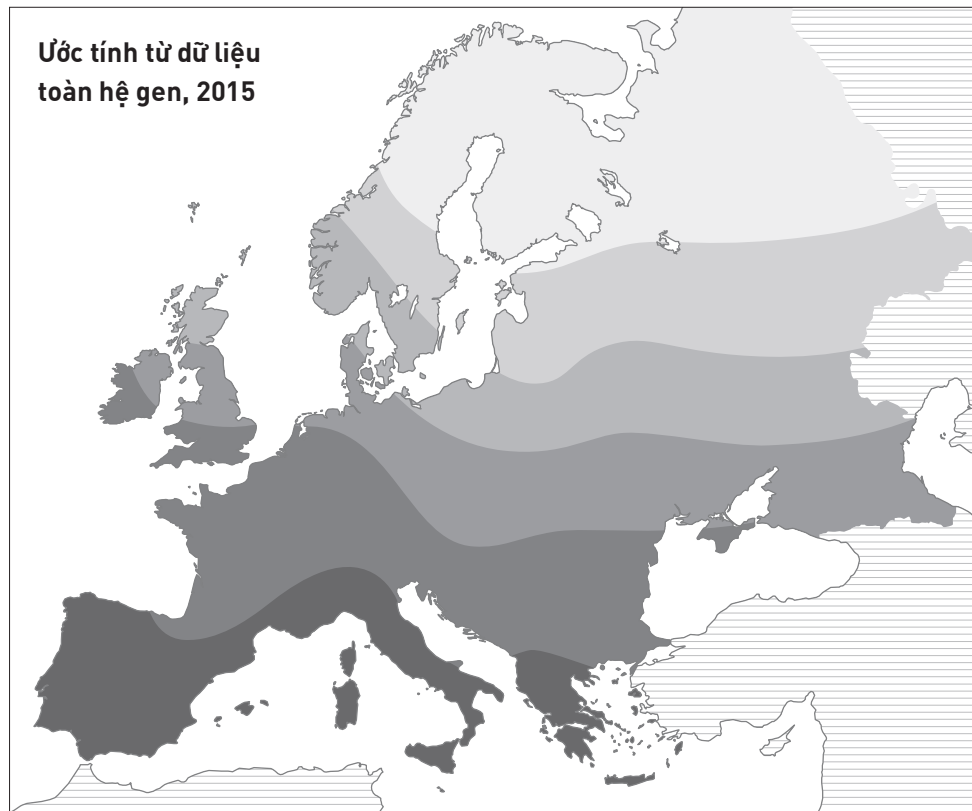
nhất dựa trên dữ liệu nhóm máu, thay vào đó là ở đảo Sardinia thuộc Địa Trung Hải ở phía tây Italy.⁸

Ví dụ về tấm bản đồ của Cavalli-Sforza cho thấy lý do mà vụ cá cược lớn của ông trở nên chua chát. Ông đã đúng khi cho rằng cấu trúc di truyền ngày nay của những quần thể lặp lại một số sự kiện trọng đại trong quá khứ của loài người. Chẳng hạn, sự đa dạng di truyền thấp hơn của người không phải châu Phi so với người châu Phi phản ánh sự đa dạng của quần thể người hiện đại lan tỏa từ châu Phi và vùng Cận Đông đã giảm vào khoảng 50.000 năm trước. Nhưng cấu trúc quần thể ngày nay của loài người không thể khôi phục những chi tiết tinh tế của các sự kiện cổ đại. Vấn đề không chỉ là mọi người đã pha trộn với những người hàng xóm của họ, làm mờ các dấu hiệu di truyền của các sự kiện trong quá khứ. Thực tế phức tạp hơn rất nhiều, vì từ DNA cổ đại, ta biết rằng những người sống ở một nơi cụ thể ngày nay hầu như không bao giờ chỉ đơn thuần là con cháu của những người sống ở cùng nơi đó vào thời xa xưa.⁹ Do đó mọi cố gắng tái tạo các đợt di chuyển quần thể trong quá khứ từ những quần thể ngày nay của bất kỳ nghiên cứu nào đều bị hạn chế. Trong *The History and Geography of Human Genes*, Cavalli-Sforza viết rằng ông đã loại trừ những quần thể được biết đến là sản phẩm của những cuộc di cư lớn, như người châu Mỹ có gốc tổ tiên châu Âu và châu Phi có nguồn gốc từ những cuộc di cư xuyên Đại Tây Dương trong 500 năm qua, hoặc các nhóm thiểu số châu Âu như người Digan và Do Thái. Ông đặt cược rằng quá khứ là một nơi đơn giản hơn nhiều so với hiện tại, và khi tập trung vào lịch sử được ghi lại về các quần thể ngày nay không chịu ảnh hưởng từ những cuộc di cư lớn, ông có thể nghiên cứu hậu duệ trực tiếp của những người sống đã có một thời gian dài sống ở cùng một nơi. Nhưng những nghiên cứu về DNA cổ đại ngày nay đã cho thấy quá khứ cũng phức tạp không kém so với hiện tại. Quần thể loài người đã nhiều lần bị xáo trộn.



Hình 1a. Một biểu đồ đường bao do Luca Cavalli-Sforza vẽ vào năm 1993 (như trên) gợi ý rằng sự di chuyển của nông dân từ phía đông có thể được tái hiện từ các mô hình biến đổi nhóm máu giữa những người sống ngày nay, với tỉ lệ tổ tiên trùng nhau cao nhất ở phía đông nam gần Anatolia.

Đóng góp mang tính đột phá của Cavalli-Sforza trong lĩnh vực nghiên cứu di truyền về thời tiền sử của loài người gợi lại câu chuyện về Moses, một nhà lãnh đạo có tầm nhìn xa, có thành tựu to lớn hơn bất kỳ ai về sau và là người tạo ra một khuôn mẫu mới để nhìn nhận thế giới. Kinh Thánh nói, “Chẳng có nhà tiên tri nào ở Israel được như Moses nữa,” nhưng cũng cho biết Moses không được phép đặt chân đến miền đất hứa như thế nào. Sau khi dẫn dắt dân tộc trong bốn mươi năm băng qua những miền hoang vu, Moses leo lên núi Nebo và nhìn về phía tây qua sông Jordan để ngắm miền đất hứa với người của ông. Nhưng ông không được phép vào vùng đất đó. Đặc ân đó dành cho những người kế vị ông.

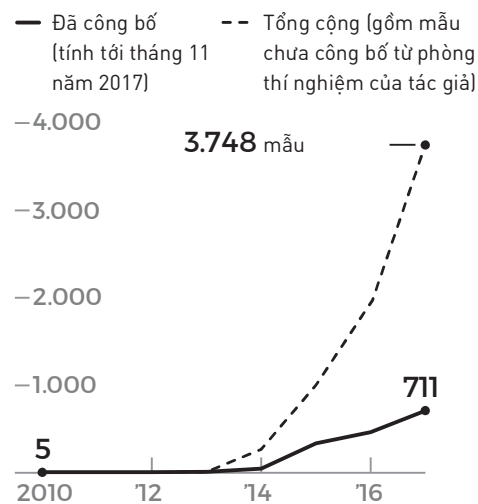


Hình 1b. Dữ liệu toàn hệ gen hiện đại cho thấy độ biến thiên tỉ lệ của gốc tổ tiên nông dân ở châu Âu không đi theo hướng đông nam–tây bắc mà theo hướng gần như vuông góc, đó là kết quả từ cuộc di cư lớn của những người chăn gia súc từ phía đông đã thay thế phần lớn gốc tổ tiên của những nông dân đầu tiên.

Đó là với nghiên cứu di truyền trong quá khứ. Cavalli-Sforza là người đầu tiên nhận thấy tiềm năng thực sự của di truyền học trong việc tiết lộ những điều chưa biết về quá khứ của loài người, tầm nhìn của ông đã có trước công nghệ cần thiết để hoàn thiện nó. Song, ngày nay mọi thứ đã rất khác. Chúng ta có dữ liệu nhiều hơn gấp vài trăm nghìn lần, ngoài ra chúng ta có quyền truy cập vào kho thông tin phong phú có trong DNA cổ đại, nguồn thông tin này đã trở thành một nguồn thông tin xác thực hơn về các đợt di cư trong quá khứ so với những công cụ khảo cổ học và ngôn ngữ học truyền thống.

Năm bộ gen người cổ đại đầu tiên được công bố vào năm 2010: một vài bộ gen cổ xưa của người Neanderthal,¹⁰ bộ gen người Denisova cổ đại,¹¹ và một cá thể khoảng 4.000 năm tuổi từ Greenland.¹² Vài năm tiếp theo chứng kiến việc công bố dữ liệu toàn hệ gen từ năm quần thể khác, tiếp theo là một loạt dữ liệu từ ba mươi tám cá thể vào năm 2014. Nhưng vào năm 2015, việc phân tích toàn hệ gen của DNA cổ đại đã được siêu máy tính giải quyết. Ba bài nghiên cứu đã bổ sung các bộ dữ liệu toàn hệ gen với 66 mẫu¹³ rồi 100 mẫu¹⁴ và sau đó là 83 mẫu.¹⁵ Đến tháng 8 năm 2017, chỉ riêng phòng thí nghiệm của tôi đã tạo ra dữ liệu toàn hệ gen cho hơn ba nghìn mẫu cổ. Chúng tôi hiện đang tạo ra dữ liệu nhanh đến mức độ trễ giữa quá trình tạo ra và công bố dữ liệu dài hơn thời gian cần để nhân đôi dữ liệu trên thực địa.

Số lượng mẫu tích lũy có dữ liệu toàn hệ gen kể từ năm 2010



Hình 2. Các phòng thí nghiệm DNA cổ đại hiện đang tạo ra dữ liệu nhanh đến mức độ trễ giữa quá trình tạo ra và công bố dữ liệu dài hơn thời gian cần để nhân đôi dữ liệu trên thực địa.

Phần lớn công nghệ phục hồi DNA cổ đại trên toàn hệ gen được phát minh bởi Svante Pääbo và đồng sự tại Viện Nhân học Tiến hóa Max Planck ở Leipzig, Đức. Họ đã phát triển nó để nghiên cứu các mẫu cực kỳ cổ xưa như người Neanderthal và người

Denisova. Tôi đã góp phần vào việc mở rộng quy mô nghiên cứu số lượng mẫu lớn và tương đối mới, gọi là mới song vẫn có hàng nghìn năm tuổi. Thời gian học nghề thuần túy là bảy năm, và tôi được đào tạo từ năm 2007 khi bắt đầu làm việc với Pääbo trong các dự án về bộ gen người Neanderthal và Denisova. Năm 2013, Pääbo giúp tôi thành lập phòng thí nghiệm DNA cổ đại riêng – phòng thí nghiệm đầu tiên ở Mỹ tập trung vào nghiên cứu DNA người cổ đại trên toàn hệ gen. Đối tác của tôi trong nỗ lực này là Nadin Rohland, người đã học việc bảy năm trong phòng thí nghiệm của Pääbo trước khi đến làm với tôi. Ý tưởng của chúng tôi là công nghiệp hóa DNA cổ đại bằng việc xây dựng một nhà máy sản xuất bộ gen kiểu Mỹ dựa trên các kỹ thuật được phát triển ở châu Âu để nghiên cứu các mẫu riêng lẻ.

Rohland và tôi nhận thấy một kỹ thuật do Matthias Meyer và Phan Xảo Mai phát triển trong phòng thí nghiệm của Pääbo có thể là mấu chốt cho việc nghiên cứu về DNA cổ đại theo hướng quy mô công nghiệp. Phát minh của Meyer và Phan Xảo Mai ra đời do sự cần thiết của nhu cầu trích xuất DNA từ người hiện đại khoảng 40.000 năm tuổi từ hang Tianyuan (Điền Viên) ở Trung Quốc.¹⁶ Khi Meyer và Phan Xảo Mai trích xuất DNA từ mẫu xương chân ở hang Tianyuan, họ nhận thấy chỉ có khoảng 0,02 phần trăm DNA trong số đó là từ chính bản thân người đó. Phần còn lại đến từ vi khuẩn đã thâm nhập vào xương sau khi người đó chết. Điều này làm cho việc giải trình tự trực tiếp trở nên quá đắt đỏ, thậm chí cả khi sử dụng công nghệ rẻ hơn hàng trăm nghìn lần đã phổ biến từ khoảng năm 2006. Để giải quyết thách thức, Meyer và Phan Xảo Mai mượn ý tưởng từ một trang trong cuốn sách hướng dẫn các phương pháp do các nhà di truyền y học phát triển. Cũng giống như các nhà di truyền y học đã phát minh ra phương pháp để phân lập DNA khỏi 2 phần trăm bộ gen cần quan tâm và loại bỏ 98 phần trăm còn lại, Meyer và Phan Xảo Mai đã phân lập được một phần nhỏ gen người từ mẫu xương ở hang Tianyuan và loại bỏ phần còn lại.

Phương pháp phân lập DNA mà Meyer và Phan Xảo Mai phát triển đã mở màn cho thành công của cuộc cách mạng DNA cổ đại. Những năm 1990, các nhà sinh học phân tử đã học cách điều chỉnh kỹ thuật khắc laser, vốn được phát minh để in các mạch điện tử, nhằm gắn hàng triệu chuỗi DNA mà họ lựa chọn vào tấm wafer silic hoặc thủy tinh. Các chuỗi này sau đó có thể được cắt ra khỏi các tấm bằng chiếc kéo quy mô phân tử (enzyme) và giải phóng thành một hỗn hợp dung dịch. Meyer và Phan Xảo Mai đã tận dụng phương pháp này để tổng hợp chuỗi DNA dài năm mươi hai ký tự, chồng lên nhau như tấm lợp mái nhà, che phủ phần lớn nhiễm sắc thể số 21 của con người. Khi khai thác xu hướng liên kết của DNA với các chuỗi rất giống nhau, họ đã “bắt” và tìm ra các chuỗi DNA từ mẫu Tianyuan mà họ quan tâm bằng cách sử dụng “mồi” cho các chuỗi mà họ đã tổng hợp nhân tạo. Họ thấy phần lớn DNA mà họ thu được là từ bộ gen của mẫu Tianyuan. Không chỉ vậy, chúng còn tới từ các phần trong bộ gen của mẫu Tianyuan mà họ muốn nghiên cứu. Họ đã phân tích dữ liệu để cho thấy người ở hang Tianyuan là người hiện đại sơ khai, thuộc dòng dõi người Đông Á ngày nay. Anh ta không có nhiều tổ tiên cụ thể từ các dòng dõi con người cổ xưa, từng bị phân tách bởi quá trình pha trộn qua hàng nghìn năm với các dòng dõi người hiện đại, trái ngược với những tuyên bố trước đó dựa trên hình dạng xương cốt của anh ta.¹⁷

Rohland và tôi đã điều chỉnh kỹ thuật này để nghiên cứu toàn hệ gen. Chúng tôi làm việc với các đồng nghiệp từ Đức để tổng hợp chuỗi DNA dài năm mươi hai ký tự, bao gồm hơn một triệu vị trí khác nhau. Chúng tôi đã sử dụng những chuỗi mồi này để làm nổi bật DNA con người so với DNA vi sinh vật. Trong một số trường hợp, nó đã tăng phần DNA mà chúng tôi quan tâm lên hơn một trăm lần. Chúng tôi đạt được một bước nhảy vọt khác hiệu quả hơn khoảng mười lần khi chỉ nhắm tới mục tiêu là các vị trí có nghĩa trong bộ gen. Chúng tôi đã tự động hóa toàn bộ phương

pháp tiếp cận, xử lý DNA bằng cách sử dụng robot cho phép một người nghiên cứu hơn 90 mẫu cùng một lúc trong khoảng thời gian vài ngày. Chúng tôi đã thuê một nhóm kỹ thuật viên để nghiền bột từ các di vật cổ đại, để tách DNA từ bột, rồi biến DNA chiết xuất thành một dạng mà chúng tôi có thể giải trình tự. Công việc trong phòng thí nghiệm chỉ là bước khởi đầu. Một nhiệm vụ phức tạp không kém là phân loại hàng tỉ chuỗi DNA thành từng cá thể mà chúng thuộc về, phân tích dữ liệu và loại bỏ các mẫu bị tạp nhiễm và tạo ra một hệ thống dữ liệu dễ truy cập. Shop Mallick, một nhà vật lý gia nhập phòng thí nghiệm của tôi sáu năm trước, đã thiết lập để máy tính làm toàn bộ công đoạn này và liên tục cập nhật chiến lược xử lý dữ liệu khi bản chất của dữ liệu phát triển và quy mô dữ liệu tăng lên.

Kết quả thậm chí còn tốt hơn chúng tôi đã hi vọng. Chi phí cung cấp dữ liệu toàn hệ gen giảm xuống dưới năm trăm đô-la mỗi mẫu. Giá này rẻ hơn hàng chục lần so với việc giải trình tự toàn hệ gen theo kiểu truyền thống. Hơn thế, phương pháp của chúng tôi có thể lấy dữ liệu toàn hệ gen từ khoảng một nửa số mẫu xương mà chúng tôi sàng lọc, dù tỉ lệ thành công thay đổi tùy theo việc bảo quản chỗ di cốt mà chúng tôi nghiên cứu. Chẳng hạn, chúng tôi có tỉ lệ thành công khoảng 75 phần trăm với các mẫu cổ được bảo quản dưới khí hậu lạnh giá của Nga, nhưng chỉ khoảng 30 phần trăm với các mẫu từ vùng Cận Đông oi nóng.

Những tiến bộ này đồng nghĩa với việc nghiên cứu toàn hệ gen của DNA cổ đại không còn cần sàng lọc số lượng lớn hài cốt trước khi tìm thấy một vài cá thể có DNA cần được phân tích. Thay vào đó, một phần đáng kể trong các mẫu được sàng lọc có niên đại vào khoảng 10.000 năm qua hiện có thể được chuyển đổi thành dữ liệu toàn hệ gen đang hoạt động. Các phương pháp mới đã có thể phân tích hàng trăm mẫu trong một nghiên cứu. Với dữ liệu như vậy, việc tái tạo các chi tiết tinh tế về sự thay đổi quần thể là hoàn toàn khả thi, làm thay đổi hiểu biết của chúng ta về quá khứ.