

PNAS

Proceedings of the National Academy of Sciences of the United States of America

www.pnas.org

The Modern View of Domestication Special Feature



This work is reprinted from the *Proceedings of the National Academy of Sciences of the United States of America*, vol. 111, no. 17, pp. 6139–6197, April 29, 2014, and includes articles from the Special Feature on The Modern View of Domestication. Any opinions, findings, conclusions, or recommendations expressed in this work are those of the authors and do not necessarily reflect the views of the National Academy of Sciences.

The National Academy of Sciences is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare. Upon the authority of the charter granted to it by the US Congress in 1863, the Academy has a mandate that requires it to advise the federal government on scientific and technical matters.

© Copyright by the National Academy of Sciences, USA
All rights reserved. Published 2014

The Modern View of Domestication

Special Feature

National Academy of Sciences
2014 Washington, DC

Don't have time to dive into research papers?

Explore the PNAS Front Matter Portal
frontmatter.pnas.org

Check out the expanded front section of PNAS written at a general level of discussion, tackling the stories of science in interesting ways.

PNAS

www.pnas.org



Cover image: Pictured are the Gamo-Gofa highlands of southern Ethiopia, a traditional agricultural landscape dotted with domesticated plants and animals such as hybrid cattle. Domesticated plants of diverse geographical origins include maize, sorghum, barley, Ethiopian banana, palm kale, and castor oil plant. The Modern View of Domestication Special Feature, appearing in this issue, presents recent genetic and archaeological evidence regarding the origin and spread of domesticated plants and animals, and addresses questions including those concerning the speed and intentionality of early domestication. See the Introduction to the Special Feature by Greger Larson et al. on pages 6139–6146. Image courtesy of Dorian Fuller.

The Modern View of Domestication Special Feature

PNAS is promoting research in anthropology. This special feature is part of that initiative. Other special features are planned, and the Editorial Board encourages submissions in these disciplines. Please see the Information for Authors at www.pnas.org.

Contents

INTRODUCTION

- 6139 **Current perspectives and the future of domestication studies**
 Greger Larson, Dolores R. Piperno, Robin G. Allaby, Michael D. Purugganan, Leif Andersson, Manuel Arroyo-Kalin, Loukas Barton, Cynthia Climer Vigueira, Tim Denham, Keith Dobney, Andrew N. Doust, Paul Gepts, M. Thomas P. Gilbert, Kristen J. Gremillion, Leilani Lucas, Lewis Lukens, Fiona B. Marshall, Kenneth M. Olsen, J. Chris Pires, Peter J. Richerson, Rafael Rubio de Casas, Oris I. Sanjur, Mark G. Thomas, and Dorian Q. Fuller

PERSPECTIVES

- 6153 **Evaluating the roles of directed breeding and gene flow in animal domestication**
 Fiona B. Marshall, Keith Dobney, Tim Denham, and José M. Capriles
- 6171 **Particularism and the retreat from theory in the archaeology of agricultural origins**
 Kristen J. Gremillion, Loukas Barton, and Dolores R. Piperno
- 6190 **Archaeological and genetic insights into the origins of domesticated rice**
 Briana L. Gross and Zhijun Zhao

RESEARCH ARTICLES

- 6147 **Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record**
 Dorian Q. Fuller, Tim Denham, Manuel Arroyo-Kalin, Leilani Lucas, Chris J. Stevens, Ling Qin, Robin G. Allaby, and Michael D. Purugganan
- 6159 **Storytelling and story testing in domestication**
 Pascale Gerbault, Robin G. Allaby, Nicole Boivin, Anna Rudzinski, Ilaria M. Grimaldi, J. Chris Pires, Cynthia Climer Vigueira, Keith Dobney, Kristen J. Gremillion, Loukas Barton, Manuel Arroyo-Kalin, Michael D. Purugganan, Rafael Rubio de Casas, Ruth Bollongino, Joachim Burger, Dorian Q. Fuller, Daniel G. Bradley, David J. Balding, Peter J. Richerson, M. Thomas P. Gilbert, Greger Larson, and Mark G. Thomas
- 6165 **Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico**
 Kraig H. Kraft, Cecil H. Brown, Gary P. Nabhan, Eike Luedeling, José de Jesús Luna Ruiz, Geo Coppens d'Ecckenbrugge, Robert J. Hijmans, and Paul Gepts
- 6178 **Beyond the single gene: How epistasis and gene-by-environment effects influence crop domestication**
 Andrew N. Doust, Lewis Lukens, Kenneth M. Olsen, Margarita Mauro-Herrera, Ann Meyer, and Kimberly Rogers
- 6184 **Establishing the validity of domestication genes using DNA from ancient chickens**
 Linus Girdland Flink, Richard Allen, Ross Barnett, Helena Malmström, Joris Peters, Jonas Eriksson, Leif Andersson, Keith Dobney, and Greger Larson

PNAS Simplifies Online Submission.

**Authors can now submit
to PNAS using the fast and
easy "Express Submission."**

When you choose Express Submission for both Direct Submissions and for PNAS PLUS, a single PDF file containing manuscript text and any supporting information may be uploaded, saving you time and effort. To submit to PNAS, please visit www.pnascentral.org.

PNAS

www.pnas.org

PNAS is committed to transparency in its editorial review process. FAQs for authors and other resources explaining the editorial process are available at www.pnas.org/site/authors/index.xhtml

Current perspectives and the future of domestication studies

Greger Larson^{a,1}, Dolores R. Piperno^{b,c}, Robin G. Allaby^d, Michael D. Purugganan^e, Leif Andersson^{f,g}, Manuel Arroyo-Kalin^h, Loukas Bartonⁱ, Cynthia Climer Vigueira^j, Tim Denham^k, Keith Dobney^l, Andrew N. Doust^m, Paul Geptsⁿ, M. Thomas P. Gilbert^o, Kristen J. Gremillion^p, Leilani Lucas^h, Lewis Lukens^q, Fiona B. Marshall^r, Kenneth M. Olsen^j, J. Chris Pires^s, Peter J. Richerson^t, Rafael Rubio de Casas^u, Oris I. Sanjurc^c, Mark G. Thomas^v, and Dorian Q. Fuller^h

^aDurham Evolution and Ancient DNA, Department of Archaeology, Durham University, Durham DH1 3LE, United Kingdom; ^bDepartment of Anthropology, Smithsonian National Museum of Natural History, Washington, DC 20560; ^cSmithsonian Tropical Research Institute, Balboa, Republic of Panama; ^dSchool of Life Sciences, University of Warwick, Coventry CV4 7AL, United Kingdom; ^eDepartment of Biology, New York University, New York, NY 10003-6688; ^fDepartment of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden; ^gDepartment of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden; ^hInstitute of Archaeology, University College London, London WC1H 0PY, United Kingdom; ⁱDepartment of Anthropology, Center for Comparative Archaeology, University of Pittsburgh, Pittsburgh, PA 15260; ^jDepartment of Biology and ^kAnthropology Department, Washington University in St. Louis, St. Louis, MO 63130; ^lSchool of Archaeology and Anthropology, College of Arts and Social Sciences, Australian National University, Canberra, ACT 0200, Australia; ^mDepartment of Archaeology, University of Aberdeen, Aberdeen AB24 3UF, Scotland; ⁿBotany Department, Oklahoma State University, Stillwater, OK 74078; ^oDepartment of Plant Sciences/MS1, Section of Crop and Ecosystem Sciences, and ^pDepartment of Environmental Science and Policy, University of California, Davis, CA 95616; ^qCentre for GeoGenetics and Section for Evolutionary Genomics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark; ^rDepartment of Anthropology, Ohio State University, Columbus, OH 43210; ^sDepartment of Plant Agriculture, University of Guelph, Guelph, ON, Canada N1G 2W1; ^tDivision of Biological Sciences, University of Missouri, Columbia, MO 65211; ^uDepartamento de Ecología, Facultad de Ciencias, Universidad de Granada, E-18071 Granada, Spain; and ^vResearch Department of Genetics, Evolution and Environment, University College London, London WC1E 6BT, United Kingdom

It is difficult to overstate the cultural and biological impacts that the domestication of plants and animals has had on our species. Fundamental questions regarding where, when, and how many times domestication took place have been of primary interest within a wide range of academic disciplines. Within the last two decades, the advent of new archaeological and genetic techniques has revolutionized our understanding of the pattern and process of domestication and agricultural origins that led to our modern way of life. In the spring of 2011, 25 scholars with a central interest in domestication representing the fields of genetics, archaeobotany, zooarchaeology, geoarchaeology, and archaeology met at the National Evolutionary Synthesis Center to discuss recent domestication research progress and identify challenges for the future. In this introduction to the resulting Special Feature, we present the state of the art in the field by discussing what is known about the spatial and temporal patterns of domestication, and controversies surrounding the speed, intentionality, and evolutionary aspects of the domestication process. We then highlight three key challenges for future research. We conclude by arguing that although recent progress has been impressive, the next decade will yield even more substantial insights not only into how domestication took place, but also when and where it did, and where and why it did not.

evolution | selection | agriculture | human ecology | human history

The domestication of plants and animals was one of the most significant cultural and evolutionary transitions in the ~200,000-y history of our species. Investigating when, where, and how domestication took place is therefore crucial for understanding the roots of complex societies. Domestication research is equally important to scholars from a wide range of disciplines, from evolutionary biology to sustainability science (1, 2). Research into both the process and spatiotemporal origins of domestication has accelerated significantly over the past decade through archaeological research, advances in DNA/RNA sequencing technology, and methods used to recover and formally identify changes

in interactions among plants and animals leading to domestication (2–4). In the spring of 2011, 25 scholars with a central interest in domestication and representing the fields of genetics, archaeobotany, zooarchaeology, geoarchaeology, and archaeology met at the National Evolutionary Synthesis Center to discuss recent progress in domestication research and identify challenges for the future. Our goal was to begin reconsidering plant and animal domestication within an integrated evolutionary and cultural framework that takes into account not just new genetic and archaeological data, but also ideas related to epigenetics, plasticity, gene-by-environment interactions, gene-culture co-

evolution, and niche construction. Each of these concepts is relevant to understanding phenotypic change, heritability, and selection, and they are all fundamental components of the New Biology (5) and Expanded Modern Evolutionary Synthesis (6).

Author contributions: G.L., D.R.P., R.G.A., M.D.P., L.A., M.A.-K., L.B., C.C.V., T.D., K.D., A.N.D., P.G., M.T.P.G., K.J.G., L. Lucas, L. Lukens, F.B.M., K.M.O., J.C.P., P.J.R., R.R.d.C., O.I.S., M.G.T., and D.Q.F. wrote the paper.

The authors declare no conflict of interest.

¹To whom correspondence should be addressed. E-mail: greger.larson@durham.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1323964111/-/DCSupplemental.

This PNAS Special Feature presents a collection of papers emanating from that meeting. Some evaluate past evidence and views on fundamental aspects of plant and animal domestication and offer a consensus perspective through the lens of more recent empirical findings and ideas. Others explore how best to investigate challenging research questions. All of the papers provide examples of how domestication research has illuminated, and will continue to enrich, our understanding of evolutionary and cultural change. In this introduction to the Special Feature, we present an outline of what is currently known about the pattern and process of domestication and we discuss foundational issues in domestication research, both in general and in light of the collected contributions. We conclude with a summary of outstanding questions and challenges.

Spatial and Temporal Patterns of Domestication

The beginnings of plant and animal domestication related to food production began globally 12,000–11,000 y ago at the end of the most recent ice age and during the transition to the Present Interglacial Period (7) (Figs. 1 and 2). Although often characterized as rapid and the result of explicit human intention (8, 9), domestication is a complex process along a continuum of human, plant, and animal relationships that often took place over a long time period and was driven by a mix of ecological, biological, and human cultural factors (2, 3). The process

encompassed a wide range of relationships, from commensalism/mutualism to low-level management, and directed control over reproduction (10, 11), although these stages did not necessarily progress in a ratchet-like fashion from wild to domestic.

The addition of a human selective component on top of a natural selection regime has enhanced the power of domestication to reveal insights into long-standing evolutionary issues, including those highlighted below. Although we eschew one-size-fits-all definitions for either plants or animals, domestication can be generally considered a selection process for adaptation to human agro-ecological niches and, at some point in the process, human preferences. Importantly, the wild progenitor species of domesticated taxa must have possessed the potential to live in the context of human ecologies, and to express traits that were favorable for human use, harvesting, and edibility. Finally, the presence of gene flow between populations of domestic and wild plants and animals [and members of the same or closely related but geographically and genetically differentiated domesticated species (12)] often results in modern populations that appear as if they arose outside the regions where the initial domestication process took place (13). As a result, it is crucial that researchers carefully evaluate whether multiple domestications of a single species occurred (13, 14), making sure to reserve the term “domestication” solely for the initial independent process, and to avoid

using the term to refer to subsequent admixture that often incorporated genetic and morphological characteristics of wild populations that were never independently domesticated (12, 13).

An increasingly rich and diverse corpus of data from molecular and archaeological research generated over the past 15 y now makes it clear that agriculture began independently over a much larger area of the globe than was once thought, and included a diverse range of plant and animal taxa (Figs. 1 and 2). At least 11 regions of the Old and New World were involved as independent centers of origin, encompassing geographically isolated regions on most continents, but several more have been suggested (Fig. 1) (3, 7, 11, 15, 16). Some of these regions were the sources of major domesticates that spread to adjacent regions, whereas others involved more regionally important species often regarded as “minor” crops today (7, 17). The combined data also clearly show that two major chronological periods are of greatest interest: the transition to the Holocene from about 12,000–9,000 B.P. (all ages are calendar years before CE 1950), and the middle Holocene between 7,000 and 4,000 B.P. (Fig. 1). Dogs were a significant exception and were certainly domesticated in the late Pleistocene before the establishment of agriculture, although both the geographical origins of dog domestication and claims for domestic dogs in ~30,000-y-old contexts remain contentious (18). In the New World, crop domestication occurred thousands of years before animal domestication, whereas the opposite was true in areas such as Africa, Arabia, and India. Some of the asynchronous patterns in individual plant and animal species from different regions may be the result of patchy evidence, as well as the lack of a clear distinction between “primary” (truly independent) vs. “secondary” (e.g., inspired by diffusional processes) domestication.

Hunting and gathering was the primary subsistence strategy for more than 95% of the time since the origin of *Homo sapiens* 200,000 y ago (19). Theories and explanations for why human cultures abandoned this long-term and apparently successful subsistence strategy and turned to food production continue to attract discussion and intense debate. Traditionally, the transition to agriculture was viewed as the result of a few single agents or “prime movers” that operated at the onset of the Holocene. Climate change, human population pressure, and culturally driven alternatives, such as “competitive feasting,” are among numerous additional agents proposed by generations of archaeologists (20–24). Simple unidirectional

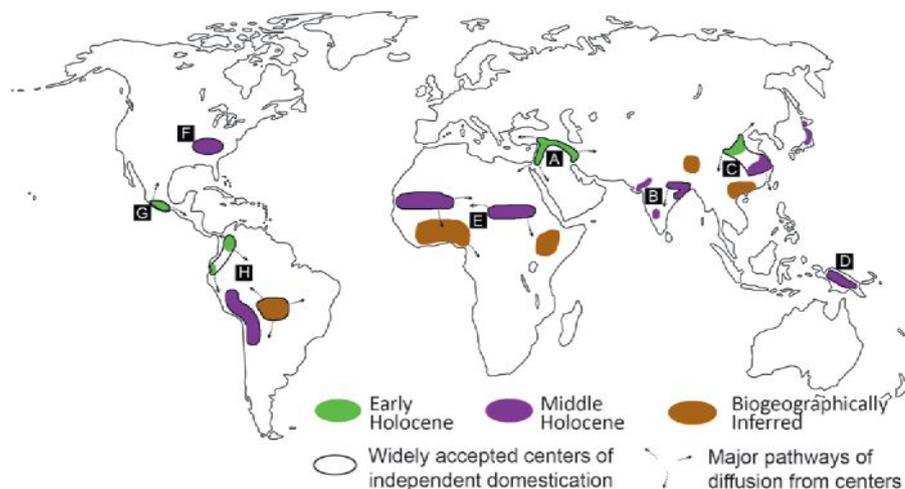


Fig. 1. A map depicting likely centers where the domestication of at least one plant or animal took place. Black outlines surround the most widely accepted independent centers of domestication, and sources of major diffusions of domesticates are indicated by arrows. Green and purple regions, respectively, are those where the domestication process took place during the late Pleistocene to early Holocene transition (12,000–8,200 B.P.), and in the middle Holocene (8,200–4,200 B.P.). Brown regions represent areas where, at present, the evidence for domestication is interpreted based upon the presence of domestic forms indigenous to these regions found outside of their native distributions. Letters A–H correspond to those listed in Fig. 2. Additional detail and references associated with each region are found in the *SI Text*.

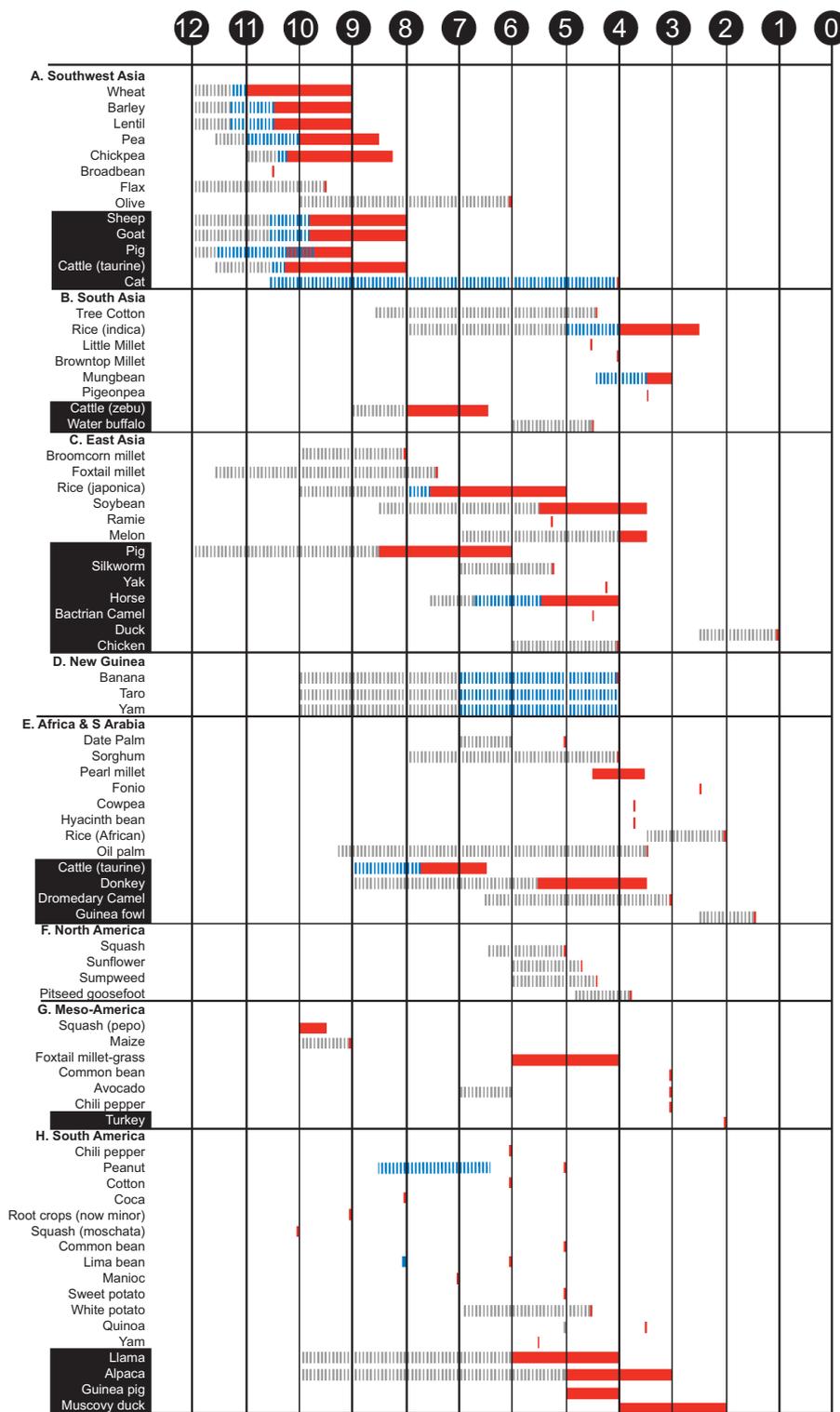


Fig. 2. A chronological chart listing the regions where, and the time frames over which, key plants and animals were domesticated. The numbers in the black circles represent thousands of years before present. Gray dashed lines represent documented exploitation before domestication or posited as necessary lead-time to domestication. Blue dashed lines represent either the management of plants or animals (including translocation) or predomestication cultivation of plants, neither of which were associated with morphological indications of domestication. Red bars frame the period over which morphological changes associated with domestication are first documented and a short, solid red bar represents the latest time by which domestication occurred. Although early Holocene plant domestication took place independently in both the Old and New Worlds, early Holocene animal domestication was restricted to the Near East. In addition, the majority of plants and animals on this list were domesticated in the middle Holocene. Additional details and references associated with each taxon are found in Table S1. Letters A–H correspond to those found in Fig. 1.

explanations, however, have proved unsatisfying for a number of researchers, and significant tensions remain between camps advocating different explanatory blueprints. The issue poses an important remaining challenge in domestication research (25) (see below).

Early Domestication Stages

The initial stages of the multispecies networks involved in domestication were critical because humans acted as: (i) dispersal agents (managing the reproduction of cultivated plants and controlling the mobility, range and density of domestic livestock); (ii) agents of (conscious or unconscious) selection, favoring the reproductive success of particular behavioral and phenotypic variants); and (iii) ecosystem modifiers, who (along with natural environmental changes) alter the developmental conditions and hence the characteristics of associated organisms.

So what is a domesticated plant or animal and how does it differ from its wild ancestor? From a present-day perspective, it is possible to recognize suites of common traits that make up the so-called “domestication syndrome” (26–28), and presumably many of these were key to early selection along the wild-to-domesticated trajectory. In plants, the syndrome is defined by a wide variety of traits that, depending on the species, may include: a reduced ability to disperse seeds without human intervention, reduction in physical and chemical defenses, reduction in unproductive side-shoots, reduction in seed dormancy, bigger seeds, more predictable and synchronous germination, and in some seed-propagated species, bigger and more inflorescences. In animals, these traits include: endocrine changes, increased docility, altered reproduction pattern and output, altered coat color, floppy ears, facial neotony, usually a reduction in size, and other changes in body proportions (26). Recent genetic and archaeological research, however, has demonstrated that not all of these traits arose at the same time in either plants or animals. In addition, it has been helpful to separate genes that controlled the traits that were under early selection (domestication genes) from those that were selected later to produce diversified and improved crops and animals (improvement genes) (4).

The strength of selection for “domestication syndrome” gene variants and their speed of fixation remains controversial. Although strong selection with rapid evolution of domestication traits within as little as 100–200 y has been suggested (8, 9, 29), recent archaeological studies have questioned these conclusions, at least for cereal domestication.

In wheat, barley, and rice, it took ~2,000–4,000 y to fix the nonshattering spikelet phenotype, a key indicator of cereal domestication (7). There are other indications in the Near East of long periods of cultivation without morphological evidence of domestication, including specific field weed flora associated with morphologically wild cereals and legumes, and large stores, suggesting reliance on cultivated production of morphologically wild species (30, 31). Doust et al. (32) show that factors previously underappreciated, such as G×E (gene-by-environment) and epistasis (gene-by-gene) interactions may have been important in slowing domestication rates. A comparison of rates of phenotypic evolution between wild and domesticated species also indicates that, contrary to expectations, evolutionary rates in domesticated species are not generally faster than those observed in wild species (7). Indeed, selection strengths for some traits are at the same level as the strength of natural selection acting on wild species, or even slightly lower (33).

The evidence for a slow pace of domestication implies a cultural period in agricultural origins called “predomestication cultivation” (PDC) (34). These periods lasted for many centuries before fully domesticated cereals appeared, as has been inferred from evidence in the Near East and China (7, 31, 35). Instances of PDC have also recently been documented in northwestern South America (36). Increasing evidence for PDC goes hand-in-hand with increasing indications of a nonsimultaneous development of the suite of traits that make up the domestication syndrome, in turn raising questions about when exactly to call archaeological remains “domesticated” and how and in what order the domestication syndrome was assembled.

These factors also make it more likely that crops were independently brought under cultivation more than once, even within a given “nuclear region,” then hybridized with cultivated or domesticated plants from other regions to become the domesticated versions we study today (37, 38). Neither genetic nor archaeobotanical studies can easily sort out these different activities, which has led to increased skepticism of the traditional models that purport rapid events taking place in a single location to explain transitions from wild to domesticated species (39, 40). In addition, the recent reevaluation of the speed of cereal domestication has led to a renewed discussion of unconscious vs. conscious selection. Charles Darwin was the first to explicitly articulate the difference between conscious selection during domestication, in which humans directly select for

desirable traits (called by Darwin “methodical” selection) (1), versus unconscious selection, where traits evolve as a by-product of growth and natural selection in field environments, or from selection on other traits. In rice, for example, glutinous grains most likely arose from conscious selection by certain Asian cultures for this cuisine-prized trait (41). In contrast, seed nonshattering in cereals is thought to have arisen as a by-product of stalk-harvesting by sickles or harvest knives, which select for seeds that do not readily fall off the stalk, rather than a result of a conscious strategy associated with beating seed heads into baskets (29). Other domestication traits in grasses are generally thought to result from unconscious selection, including seed size, seed dormancy, synchronous seed ripening, and apical dominance (27).

Most domesticated plants are not cereals, and other crops with different domestication syndromes may have had faster rates of domestication once humans targeted them for cultivation, and been more prone to have traits selected by conscious selection. The great cultural geographer Carl Sauer (42) insightfully noted that squashes, beans, and various root crops (along with maize, the premier cereal crop of the Americas) were not mass-harvested and mass-planted, nor likely mass-selected, as the Old World cereals were. Individual harvesting and selection by early farmers, who would be expected to choose and deliberately propagate the crop attributes most useful to them when they could distinguish the useful phenotypes, could foster conscious selection and result in faster fixation of crucial and preferred domestication traits, such as the loss of toxicity and increased size of starch storage organs in tubers and roots. Fruit nonbitterness in squashes and melons, major early domesticates in all regions of the Americas and parts of Asia and Africa may also have been rapidly and consciously selected. For example, botanical remains from human teeth indicate that the loss of fruit bitterness in the squash species *Cucurbita moschata* took place by at least 9200 B.P., only 800 y later than the first evidence for its domestication. In fact, the loss may have taken place even earlier because the seed traits used to document domestication do not inform fruit-flesh characteristics (36, 43). Arguments for relatively fast, conscious selection have also been made for the important seed dormancy trait in Old World lentils and peas (44).

What about conscious vs. unconscious selection in animals? Marshall et al. (12) make a compelling case that intentional breeding of females was largely absent during the early stages of domestication for a wide range

of species. This theory, along with what probably was considerable gene flow between wild and early managed animals (13), poses challenges to a number of commonly held assumptions about early domestication in some species relating to interpretations of genetic bottlenecks and molecular sequences more generally, the number of times a species was domesticated, and how various domestication traits emerged and were maintained in the long term. Clearly, many questions persist about the roles of directed vs. undirected selection across the spectrum of domesticated plants and animals.

Research over the past few decades has made it clear that prehistoric humans around the world significantly modified their environments, sometimes before and during the process of plant and animal domestication, and the role of humans in the enduring modification of environments is no longer underestimated (45–49). A uniquely important aspect of human environmental modification is the additional role cultural transmission plays in maintaining patterns of enduring local ecologies, resulting in a strongly enculturated ecological inheritance. Because they can often be traced archaeologically, cultural transmission processes have received increasing interest and mathematical modeling in the social sciences (50–52) and are embedded both in practice and in material settings (e.g., terraces, canals, mounding, soil management, lassos, penning, somatic modifications such as castration, food-processing tools). Although the process of cultural inheritance differs from that of genetics, it plays a crucial role in maintaining both cultural practices over generations and environments in which domestication and husbandry occurred and were maintained. Human intentionality and knowledge systems must have been key components among the interacting mechanisms within these bio/eco-cultural environments, and cultural transmission provided a basis for the maintenance of cumulative innovation. Traditional ecological knowledge over the longer term has maintained crop landrace diversity, and remains important for biodiversity distribution and ecosystem services more generally (53).

Genetic and Evolutionary Insights from Domestication

The study of domesticated species has led to increased interest in several important issues in genetics and evolutionary biology, including the underlying genetic architecture of adaptations and parallel evolution. Genetic research is increasingly identifying

domestication genes, especially in plants (4). By contrast, many fewer domestication genes have been identified in animals (13). With the exception of coat-color genes, genetic variants that can be unambiguously assigned to early stages in domestication in animals have not yet been revealed. There are several possible reasons for this. First, discovering the molecular basis of domestication traits is relatively easy and inexpensive in plants compared with animal populations because early animal selection likely focused on behavioral and other characteristics (such as tameness and altered reproduction), with complex genetic foundations that are more difficult to study than classic morphological traits (54). Second, there may simply be few domestication loci with major effects in animals. Early animal domestication may have happened by shifting the allele frequencies at many loci, each with small individual effects, thereby altering the phenotype. This scenario would be consistent with the observation that many domestic animals (e.g., pigs) can readily establish feral populations that in many aspects mimic the phenotype of their wild ancestors (55).

Thus, an important question for both plants and animals is whether the striking phenotypic changes seen during domestication are under the control of single or multiple genes. Thus far, separate studies have identified both single (or few) genes and combinations of numerous genes of small effect, depending on the approach and species in question (4). To some extent, different inferences concerning the genetic architecture of domestication can be because of different methodological approaches. Forward genetic approaches, such as quantitative trait loci (QTL) mapping and genome-wide association studies have the capability of finding multiple loci controlling phenotypic traits, and thus to interpret a domestication trait as under the control of multiple genes (4). Reverse genetic approaches concentrate on particular genes and cannot, by themselves, discover multiple loci for a particular phenotype. Genes in reverse genetic approaches are often chosen because their mutant phenotypes in model systems, such as chicken, mouse, *Arabidopsis*, maize, and rice, are analogous to phenotypic differences between wild and domesticated species. It is then possible to ask whether sequence changes in the locus explain phenotypic differences. An example of this approach involves a mutation of the transcription factor *ramosa1* (*ra1*) locus in maize that results in loss of floral branches (56). Differences in the *ra1* locus were later found to be correlated

with differences in floral branching in maize and other grasses (57). However, it is not the only gene involved, as shown by QTL studies that indicated up to five significant QTL regions controlling these traits (58). Finally, a recent study (59) demonstrated that the action of *sh4* in rice is not always sufficient to produce nonshattering phenotypes.

An additional question is whether the same genes underlie similar phenotypic shifts in numerous domesticated plants and animals. In other words, is there parallelism from the same underlying genetic and developmental pathways or convergent evolution of unrelated taxa using unrelated gene networks (60)? In grasses, such forms are particularly striking, and similar awned and awnless spikelets, hulled and free-threshing grain, black-, red-, and straw-colored seed coats are found in multiple domesticated cereals. The geneticist Vavilov termed this phenomenon the Law of Homologous Series (61), and the first phase of comparative mapping in the grasses, using restriction fragment-length polymorphism markers, inferred QTL for shattering in rice, sorghum, and maize to be at the same location (62). Further work has proved equivocal, since most genes for shattering in grasses are unique to each domesticated lineage (63), though a recent study has shown that the major locus for shattering in sorghum corresponds to minor loci in rice and maize (64). Nevertheless, some mutations in domesticates are in fact parallel mutations. For example, variants of the *MC1R* locus are responsible for independently derived pig coloration patterns (65). Moreover, mutations at this gene appear to be associated with difference in color patterns in numerous domestic animals (66). A similar example of parallel evolution is associated with the rise of sticky cereals in northeast Asia, where glutinous rice, millets, and barley, among others (41, 67), are the result of alternative mutations at the *Waxy* gene (68–70).

Whether mutations selected during domestication were novel or were present as standing genetic variation in ancestral wild populations is a question of increasing interest. It has traditionally been assumed that phenotypic change and new adaptations arise from new mutations, but recent research increasingly shows that standing genetic variation plays important roles in a variety of species (71). For example, traits present as variants in wild progenitors today include the gene for tomato fruit size (*fw2.2*) (72), maize plant architecture (e.g., *teosinte branched1*) (73), seasonality controls (74, 75), and seed size [usually polygenic (76)]. Fast morphological evolution in cultivated plant popula-

tions may have ensued as favorable phenotypes, including those initially exposed by genetic or external environmental perturbations in response to the new field conditions, may have been preferentially selected by farmers who were not constrained by mutation rates (77, 78). Having said that, several traits in domesticated plants, including those associated with the reduction of seed-shattering in legumes and grasses, are deleterious in the wild, and if present, are rarely expressed phenotypically. In animals, analyses of modern dog genomes have revealed a handful of mutations (not found in extant wolves) with large effects on morphological variation, although given the predominance of selection for novel and unusual characteristics in dogs, this pattern is likely the exception (79). In many other domestic animals, humans likely selected for trait variants that were already present in ancestral populations, thereby altering the frequencies of the standing genetic variation.

As success in isolating domestication-related genes proceeds, it should become easier to distinguish between standing and new genetic variation, as well as to recognize parallelism in *de novo* mutations among domesticated species. Additionally, as the availability of genome-wide sequence data for domesticated species increases, it is becoming increasingly feasible to use selective sweep mapping to identify genomic regions that have been targets of selection during domestication without a priori information on candidate domestication genes (e.g., ref. 80). Challenges associated with this approach include the fact that the trait or traits affected by the selected genes may not be known, that selection that favors a *de novo* mutation during the domestication process will generate a more conspicuous signature of a selective sweep than selection for mutations that were already segregating in populations of the wild progenitor, and that some demographic processes can mimic the effects of selection on patterns of genetic variation. Understanding the different genetic architecture of domestication across crop types and in animals remains a major challenge for genetic research.

One new promising direction is the study of ancient DNA. Our increasing ability to identify selected mutations for domestication-associated traits in archaeological plant and animal remains is providing a unique temporal trajectory of the evolution of domesticated species, and the selection strengths that acted upon selected genes. One such example tested claims that two different genes (*TSHR* and *BCDO2*) were involved in

early chicken domestication by typing the mutations in ancient European chickens. Because the wild-type alleles of both genes were segregating at a high frequency as recently as 500 y ago, the ancient DNA evidence demonstrated that the modern ubiquity of a mutation, even one that differentiates domestic and wild populations, cannot automatically be conflated with an ancient origin linked to early domestication (81).

Key Challenges for the Future

The enormous amount of empirical data compiled on domestication and associated human- and naturally driven circumstances during the past decades has naturally led to the generation of a number of questions, some of which pose key future challenges.

Filling in Gaps on Maps

One of the fundamental challenges of domestication research is filling the gaps that remain in both geographical and genomic maps. Genetic research provides a growing toolkit for elucidating the relationships between domesticates and their wild ancestors, and between the traits that make domesticates suited to anthropogenic environments and their underlying genetic architecture. The successes of genetics, touched upon above, at identifying domestication genes have been numerous and mostly recent. Expanding this repertoire remains a priority, but it is increasingly evident that we also need more evidence from ancient DNA, so that patterns found in modern populations can be compared with those of the past, and geographies and phylogeographic and adaptive hypotheses can be tested over the evolutionary time period of domestication.

In addition, archaeological research has many gaps on the chronology and regional sequences of domestication of plants and animals, and the contexts of agricultural origins. Recent research has shown that increased sampling and methodological developments have made it possible to clearly document cereal domestication [e.g., rice (82)], push back the earliest evidence for both the domestication of maize in southern Mexico (83) and the arrival of crops in northern Peru (36), and to recognize the likely independent processes of agricultural origins and domestication in New Guinea (45), parts of India (84), and Africa (85). These research successes within the past decade imply that more new information on more species from more regions and earlier periods can be expected and should be actively sought.

Related to this are important continuing challenges in determining why so few of the animal and plant species that were hunted

and gathered by ancestral human populations were ever domesticated (86), and whether most species were domesticated once or multiple times. We recognize that distinguishing these options is complicated, and it is increasingly clear that incomplete archaeological evidence and genetic data are open to conflicting interpretation. This aspect highlights the importance of explicit modeling and simulation of a range of hypotheses concerning the starting conditions and processes of domestication (14, 87). Factors potentially leading to confusion include the fact that multiple domestication episodes may be hidden from genetic view today as a result of both bottlenecks (in some cases leading to extinction) and introgression. Archaeobotany, for example, has increasingly recognized extinct morphotypes of domesticated wheat (88, 89), and ancient DNA can help to identify lost genetic lineages of crops. Introduced domesticates may introgress with local wild populations, thus capturing genetic and phenotypic variation that can later be misinterpreted as the independent domestication of distinct wild animal (13) and plant (e.g., rice) (90–92) populations. Resolving these issues requires more targeted ancient DNA research and more realistic and sophisticated modeling.

Environmental and Ecological Contexts of Agricultural Origins

Although climate change remains the prime landscape and ecological modifier at the origins of agriculture, human behavior and the activities of diverse cultural traditions must be better understood. Beyond simply collecting more archaeological and paleoecological evidence, there is a need to broaden the study of past landscapes and their related ecosystems for both naturally derived features and the legacies of past human action. For example, more research should systematically map local and regional distributions of enriched soils, created through human activities, which are well known from Amazonia and Europe but much less well documented elsewhere (e.g., refs. 93, 94). Vegetation formations studied by plant ecologists and environmental historians may also be anthropogenic legacies, as has been suggested for a number of regions including South Asia and throughout the Americas (45, 53, 95, 94).

New or underdeveloped fields, such as ecological developmental biology (eco-devo) (96) and epigenetics (97)—together with mechanisms, such as developmental plasticity (98, 99)—are assuming increasing importance in the study of diversification, the origin of novelties, and evolutionary change. These fields should be extended to the realm

of domestication research, in part because phenotypic and genetic responses to natural- and human-created environmental variability are among the most neglected issues in domestication studies. As recent work with teosinte has shown, field- and laboratory-controlled experiments are needed to better understand them (99). Another example is that although it has been inferred that large seed size was selected by soil disturbance and depth of burial (e.g., ref. 100), as presumably seeds with the largest mass were better able to emerge from deeper burial depths associated with cultivation practices, others have suggested that seed size increase may be a plastic phenotypic response to enriched soils of early cultivation (101). New experimental research (102) on different legume crops indicate seed mass was important for emergence in some species, including those predicted to conform to the burial hypothesis (60), but not in others, suggesting a common single mechanism for seed size increase was not at work. In another vein, Blumler's analysis (103), suggesting that the Near East was unusually well endowed with large seeded grasses preadapted to domestication, might explain the early and diverse domestication of plants in that region. In addition, Marshall et al. (12) make the point that epigenetic mechanisms should also be investigated in animal genetic responses during the domestication process.

A few scholars have discussed the potential role of climate shifts and atmospheric gas concentrations on biota at the transition between the late Pleistocene and early Holocene. More specifically, the authors have suggested that agriculture was a more favorable strategy in the Holocene as a result of these environmental shifts (e.g., refs. 99, and 104–107). Lower CO₂ and temperature reduced plant productivity, in part by reducing photosynthetic efficiency, thus exacerbating drought stress: effects that were more marked on C₃ plants but also present to a surprising degree in C₄ plants. Did the rapid increase of CO₂ and temperature at the onset of the Holocene make plants more attractive as a readily intensifiable resource and make cultivation more efficient? As plant productivity increased, why would some cultural traditions delay the shift in cultivation until the middle Holocene, and how can we connect the adoption of animal herding to changes in plant productivity? Although important global processes have doubtless impacted foragers and early cultivators, a great deal more research is necessary to unravel the causes, constraints, and exceptions to the early or middle Holocene transitions to farming.

Further experimental data on the impact of late-glacial and early Holocene temperatures

and CO₂ levels on the biological characteristics of wild progenitors of crops are needed to also understand how they may have influenced other phenotypic attributes of crop and animal progenitors on the eve of and during agricultural beginnings (99). Just as genetic studies of domestication have shown that conclusions drawn only from modern populations provide an incomplete and sometimes biased picture of the past (81, 108, 109), we need to better understand the interplay between past ecology, climate, plant phenotypic responses, and human activities.

Why Hunters and Gatherers Turned to Cultivation and Herding

Explaining the origins of agriculture is still one of the most contentious issues for social scientists. Few dispute that the interplay of climate, human demography, and social systems through time and space played a significant role (110). Although some consider the primary driving factors to be patterns of climatic and ecological change, others argue for the primacy of social imperatives and changes within social systems (23, 24, 111). More generally, some scholars have claimed that no explanations are likely to be universally applicable (112), whereas others have adopted an explicitly comparative approach, identifying parallel processes and exploring common underlying patterns (7, 15, 25). Further progress on this issue should focus not only on the acquisition of more data, but also on marshaling and discussing existing evidence, which may suggest which factors driving agricultural origins were of greater importance. In a number of nuclear centers there are now fewer disagreements about the cultural history of early agriculture (including the chronology and the organisms involved), which should make explanatory endeavors less complicated. As known instances of agricultural origins are further clarified, we will have more parallel histories of domestication from which to derive commonalities or process and patterns of causation.

Conclusions

The collection of papers presented in this Special Feature attempts to rise to the challenges outlined above. The articles illustrate a range of approaches to the study of domestication, including genetics, archaeological science, and anthropology, and raise new questions and hypotheses that are ripe for further testing. Even so, the new evidence and ideas presented here highlight a minority of the many species that were domesticated and subsequently improved by prehistoric cultures. Domest-

tication remains a vibrant research area in biology and archaeology 145 y after Darwin's seminal work (1), and the coming decade will no doubt generate satisfying and perhaps definitive answers to a wide range of outstanding questions.

- 1 Darwin C (1868) *The Variation of Animals and Plants under Domestication* (John Murray, London).
- 2 Gepts P, et al., eds (2012) *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability* (Cambridge Univ Press, Cambridge).
- 3 Price D, Bar-Yosef O eds (2011) *The Beginnings of Agriculture: New Data, New Ideas. Curr Anthropol* 52 (Supp 4. Wenner-Gren Symposium Series; Univ of Chicago Press, Chicago).
- 4 Olsen KM, Wendel JF (2013) A bountiful harvest: Genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol* 64:47–70.
- 5 Council MotNR (2009) *A New Biology for the 21st century. Committee on a New Biology for the 21st Century: Ensuring the United States Leads the Coming Biology Revolution* (National Academies, Washington, DC).
- 6 Wake M (2010) Development in the real world. Review of *Ecological Developmental Biology: Integrating Epigenetics, Medicine, and Evolution* by Gilbert, S.F., Epel, D. *Am Sci* 98(1): 75–78.
- 7 Fuller DQ, et al. (2014) Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record. *Proc Natl Acad Sci USA* 111:6147–6152.
- 8 Abbo S, Lev-Yadun S, Gopher A (2010) Agricultural origins: Centers and noncenters; a Near Eastern reappraisal. *Crit Rev Plant Sci* 29(5):317–328.
- 9 Innan H, Kim Y (2004) Pattern of polymorphism after strong artificial selection in a domestication event. *Proc Natl Acad Sci USA* 101(29):10667–10672.
- 10 Zeder MA (2012) Pathways to animal domestication. *Biodiversity in Agriculture: Domestication, Evolution and Sustainability*, eds Gepts P, et al. (Cambridge Univ Press, Cambridge).
- 11 Piperno DR (2011) The origins of plant cultivation and domestication in the New World tropics. *Curr Anthropol* 52(S4): S453–S470.
- 12 Marshall FB, Dobney K, Denham T, Capriles JM (2014) Evaluating the roles of directed breeding and gene flow in animal domestication. *Proc Natl Acad Sci USA* 111:6153–6158.
- 13 Larson G, Burger J (2013) A population genetics view of animal domestication. *Trends Genet* 29(4):197–205.
- 14 Gerbault P, et al. (2014) Storytelling and story testing in domestication. *Proc Natl Acad Sci USA* 111:6159–6164.
- 15 Barker G (2006) *The Agricultural Revolution in Prehistory: Why did Foragers Become Farmers?* (Oxford Univ Press, Oxford).
- 16 Fuller DQ (2010) An emerging paradigm shift in the origins of agriculture. *General Anthropology* 17(2):1–12.
- 17 Kraft KH, et al. (2014) Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annum*, in Mexico. *Proc Natl Acad Sci USA* 111:6165–6170.
- 18 Larson G, et al. (2012) Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA* 109(23):8878–8883.
- 19 McDougall I, Brown FH, Fleagle JG (2005) Stratigraphic placement and age of modern humans from Kibish, Ethiopia. *Nature* 433(7027):733–736.
- 20 Binford LR (1968) *Post-Pleistocene Adaptations*, eds Binford SR, Binford LR (Aldine, Chicago), pp 313–342.
- 21 Moore A, Hillman GC (1992) The Pleistocene to Holocene transition and human economy in Southwest Asia: The impact of the Younger Dryas. *Am Antiq* 57:482–494.
- 22 Cohen MN (2009) Introduction: Rethinking the origins of agriculture. *Curr Anthropol* 50(5):591–595.
- 23 Hayden B (2009) The proof is in the pudding: Feasting and the origins of domestication. *Curr Anthropol* 50(5):597–601.
- 24 Bowles S, Choi J-K (2013) Coevolution of farming and private property during the early Holocene. *Proc Natl Acad Sci USA* 110(22): 8830–8835.
- 25 Gremillion KJ, Barton L, Piperno DR (2014) Particularism and the retreat from theory in the archaeology of agricultural origins. *Proc Natl Acad Sci USA* 111:6171–6177.
- 26 Hammer K (1984) Das domestikationssyndrom. *Kulturpflanze* 32(1):11–34.
- 27 Harlan JR (1992) *Crops and Man* (American Society of Agronomy, Madison, WI), 2nd Ed.
- 28 Darwin C (1859) *On the Origin of Species by Means of Natural Selection, or the Preservation of Favoured Races in the Struggle for Life* (John Murray, London).
- 29 Hillman GC, Davies MS (1992) Domestication rates in wild wheats and barley under primitive cultivation. *Préhistoire de l'Agriculture: Nouvelles Approches Expérimentales et Ethnographiques, Monographie du CRA*, ed Anderson PC (Centre de Recherches Archéologiques, CNRS, Paris), pp 113–158.
- 30 Weiss E, Kislev ME, Hartmann A (2006) Anthropology. Autonomous cultivation before domestication. *Science* 312(5780): 1608–1610.
- 31 Willcox G, Stordeur D (2012) Large-scale cereal processing before domestication during the tenth millennium cal BC in northern Syria. *Antiquity* 86(331):99–114.
- 32 Doust AN, et al. (2014) Beyond the single gene: How epistasis and gene-by-environment effects influence crop domestication. *Proc Natl Acad Sci USA* 111:6178–6183.
- 33 Purugganan MD, Fuller DQ (2011) Archaeological data reveal slow rates of evolution during plant domestication. *Evolution* 65(1): 171–183.
- 34 Harris DR, Hillman GC (1989) *Foraging and Farming: The Evolution of Plant Exploitation* (Unwin Hyman, London).
- 35 Yang X, et al. (2012) Early millet use in northern China. *Proc Natl Acad Sci USA* 109(10):3726–3730.
- 36 Piperno DR, Dillehay TD (2008) Starch grains on human teeth reveal early broad crop diet in northern Peru. *Proc Natl Acad Sci USA* 105(50):19622–19627.
- 37 Allaby RG, Fuller DQ, Brown TA (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proc Natl Acad Sci USA* 105(37):13982–13986.
- 38 Bettinger R, Barton L, Morgan C (2010) The origins of food production in North China: A different kind of agricultural revolution. *Evol Anthropol* 19:9–21.
- 39 Brown TA, Jones MK, Powell W, Allaby RG (2009) The complex origins of domesticated crops in the Fertile Crescent. *Trends Ecol Evol* 24(2):103–109.
- 40 Fuller DQ, Willcox G, Allaby RG (2011) Cultivation and domestication had multiple origins: Arguments against the core area hypothesis for the origins of agriculture in the Near East. *World Archaeol* 43(4):628–652.
- 41 Fuller D, Rowlands M (2011) Ingestion and food technologies: Maintaining differences over the long-term in West, South and East Asia. *Interweaving Worlds: Systematic Interactions in Eurasia, 7th to 1st millennia BC. Essays from a Conference in Memory of Professor Andrew Sherratt*, eds Bennet J, Sherratt S, Wilkinson TC, (Oxbow Books, Oxford).
- 42 Sauer CE (1965) American agricultural origins: a consideration of nature and culture. *Land and Life. A Selection of the Writing of Carl Ortwin Sauer*, ed Leighy J (Univ of California Press, Berkeley).
- 43 Dillehay TD, Rossen J, Andres TC, Williams DE (2007) Pre-ceramic adoption of peanut, squash, and cotton in northern Peru. *Science* 316(5833):1890–1893.
- 44 Abbo S, et al. (2011) Experimental growing of wild pea in Israel and its bearing on Near Eastern plant domestication. *Ann Bot (Lond)* 107(8):1399–1404.
- 45 Denham TP, et al. (2003) Origins of agriculture at Kuk Swamp in the highlands of New Guinea. *Science* 301(5630):189–193.
- 46 Piperno DR (2006) Quaternary environmental history and agricultural impact on vegetation in Central America. *Ann Mo Bot Gard* 93(2):274–296.
- 47 Lentz DL, ed (2000) *Imperfect Balance: Landscape Transformations in the Pre-Columbian Americas* (Columbia Univ Press, New York).
- 48 Ellis EC, et al. (2013) Used planet: A global history. *Proc Natl Acad Sci USA* 110(20):7978–7985.
- 49 Arroyo-Kalin M (2012) Slash-burn-and-churn: Landscape history and crop cultivation in pre-Columbian Amazonia. *Quat Int* 249:4–18.
- 50 Richerson PJ, Boyd R (2005) *Not By Genes Alone* (Univ of Chicago Press, Chicago).
- 51 Shennan SJ (2002) *Genes, Memes and Human History: Darwinian Archaeology and Cultural Evolution* (Thames and Hudson, London).

- 52 Powell A, Shennan S, Thomas MG (2009) Late Pleistocene demography and the appearance of modern human behavior. *Science* 324(5932):1298–1301.
- 53 Ford A, Nigh R (2009) Origins of the Maya forest gardens. *J Ethnobiol* 29:213–236.
- 54 Albert FW, et al. (2009) Genetic architecture of tameness in a rat model of animal domestication. *Genetics* 182(2):541–554.
- 55 Albarella U, Dobney K, Rowley-Conwy P (2006) The domestication of the pig (*Sus scrofa*): New challenges and approaches. *Documenting Domestication: New Genetic and Archaeological Paradigms*, eds Zeder MA, Bradley DG, Emshwiller E, Smith BD (Univ of California Press, Berkeley), pp 209–227.
- 56 Vollbrecht E, Springer PS, Goh L, Buckler ES, 4th, Martienssen R (2005) Architecture of floral branch systems in maize and related grasses. *Nature* 436(7054):1119–1126.
- 57 Sigmon B, Vollbrecht E (2010) Evidence of selection at the *ramosa1* locus during maize domestication. *Mol Ecol* 19(7):1296–1311.
- 58 Upadayaya N, da Silva HS, Bohn MO, Rocheford TR (2006) Genetic and QTL analysis of maize tassel and ear inflorescence architecture. *Theor Appl Genet* 112(4):592–606.
- 59 Ishikawa R, et al. (2010) Allelic interaction at seed-shattering loci in the genetic backgrounds of wild and cultivated rice species. *Genes Genet Syst* 85(4):265–271.
- 60 Fuller DQ, Allaby RG (2009) Seed dispersal and crop domestication: Shattering, germination and seasonality in evolution under cultivation. *Fruit Development and Seed Dispersal, Annual Plant Reviews*, ed Ostergaard L (Wiley-Blackwell, Oxford), Vol 38, pp 238–295.
- 61 Vavilov NI (1951) Phytogeographic basis of plant breeding. The origin, variation, immunity and breeding of cultivated plants. *Chron Bot* 13:1–366.
- 62 Paterson AH, et al. (1995) Convergent domestication of cereal crops by independent mutations at corresponding genetic Loci. *Science* 269(5231):1714–1718.
- 63 Li W, Gill BS (2006) Multiple genetic pathways for seed shattering in the grasses. *Funct Integr Genomics* 6(4):300–309.
- 64 Lin ZW, et al. (2012) Parallel domestication of the Shattering1 genes in cereals. *Nat Genet* 44(6):720–U154.
- 65 Fang M, Larson G, Ribeiro HS, Li N, Andersson L (2009) Contrasting mode of evolution at a coat color locus in wild and domestic pigs. *PLoS Genet* 5(1):e1000341.
- 66 Linderholm A, Larson G (2013) The role of humans in facilitating and sustaining coat colour variation in domestic animals. *Semin Cell Dev Biol* 24(6–7):587–593.
- 67 Eriksson G (1969) The waxy character. *Hereditas* 63(1–2):180–204.
- 68 Ma J, et al. (2010) Molecular characterization and comparative analysis of two waxy alleles in barley. *Genes & Genomics* 32(6):513–520.
- 69 Araki M, Numaoka A, Kawase M, Fukunaga K (2012) Origin of waxy common millet, *Panicum miliaceum* L. in Japan. *Genet Resour Crop Evol* 59(7):1303–1308.
- 70 Olsen KM, Purugganan MD (2002) Molecular evidence on the origin and evolution of glutinous rice. *Genetics* 162(2):941–950.
- 71 Barrett RD, Schluter D (2008) Adaptation from standing genetic variation. *Trends Ecol Evol* 23(1):38–44.
- 72 Nesbitt TC, Tanksley SD (2002) Comparative sequencing in the genus *Lycopersicon*. Implications for the evolution of fruit size in the domestication of cultivated tomatoes. *Genetics* 162(1):365–379.
- 73 Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43(11):1160–1163.
- 74 Jones H, et al. (2008) Population-based resequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. *Mol Biol Evol* 25(10):2211–2219.
- 75 Cloutault J, et al. (2012) Evolutionary history of pearl millet (*Pennisetum glaucum* [L.] R. Br.) and selection on flowering genes since its domestication. *Mol Biol Evol* 29(4):1199–1212.
- 76 Gupta PK, Rustgi S, Kumar N (2006) Genetic and molecular basis of grain size and grain number and its relevance to grain productivity in higher plants. *Genome* 49(6):565–571.
- 77 Pickersgill B (2007) Domestication of plants in the Americas: Insights from Mendelian and molecular genetics. *Ann Bot (Lond)* 100(5):925–940.
- 78 Gremillion KJ, Piperno DR (2009) Human behavioral ecology, phenotypic (developmental) plasticity, and agricultural origins. *Curr Anthropol* 50(5):615–619.
- 79 Boyko AR, et al. (2010) A simple genetic architecture underlies morphological variation in dogs. *PLoS Biol* 8(8):e1000451.
- 80 Rubin C-J, et al. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464(7288):587–591.
- 81 Flink LG, et al. (2014) Establishing the validity of domestication genes using DNA from ancient chickens. *Proc Natl Acad Sci USA* 111:6184–6189.
- 82 Fuller DQ, et al. (2009) The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323(5921):1607–1610.
- 83 Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106(13):5019–5024.
- 84 Fuller DQ (2011) Finding plant domestication in the Indian subcontinent. *Curr Anthropol* 52(5):S347–S362.
- 85 Fuller DQ, Hildebrand L (2013) *Domesticating Plants in Africa. The Oxford Handbook of African Archaeology*, eds Mitchell P, Lane P (Oxford Univ Press, Oxford), pp 507–525.
- 86 Diamond J (2002) Evolution, consequences and future of plant and animal domestication. *Nature* 418(6898):700–707.
- 87 Allaby R (2010) Integrating the processes in the evolutionary system of domestication. *J Exp Bot* 61(4):935–944.
- 88 Jones G, Valamoti S, Charles M (2000) Early crop diversity: A “new” glume wheat from northern Greece. *Vegetation History and Archaeobotany* 9(3):133–146.
- 89 Fuller DQ, Willcox G, Allaby RG (2012) Early agricultural pathways: Moving outside the “core area” hypothesis in Southwest Asia. *J Exp Bot* 63(2):617–633.
- 90 Fuller DQ (2011) Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. *Rice* 4:78–92.
- 91 Molina J, et al. (2011) Molecular evidence for a single evolutionary origin of domesticated rice. *Proc Natl Acad Sci USA* 108(20):8351–8356.
- 92 Gross BL, Zhao Z (2014) Archaeological and genetic insights into the origins of domesticated rice. *Proc Natl Acad Sci USA* 111:6190–6197.
- 93 Baade J (2012) Irragic anthrosols—Artifacts of human adaptation to arid conditions: Examples from the high Himalaya and the coastal desert of Peru. *Climates, Landscapes, and Civilizations Geophysical Monograph Series* 198:203–208.
- 94 Fairhead J, Leach M (2009) *Amazonian Dark Earths in Africa? Amazonian Dark Earths: Wim Sombroek’s Vision* (Springer, New York), pp 265–278.
- 95 Woods WI, et al. (2009) *Amazonian Dark Earths: Wim Sombroek’s Vision* (Springer, New York).
- 96 Gilbert SF, Epel D (2009) *Ecological Developmental Biology: Integrating Epigenetics, Medicine and Evolution* (Sinauer Associates, Inc., Sunderland, MA).
- 97 Jablonka E, Raz G (2009) Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol* 84(2):131–176.
- 98 West-Eberhard MJ (2003) *Developmental Plasticity and Evolution* (Oxford Univ Press, New York).
- 99 Piperno DR, Holst I, Winter K, McMillan O (2014) Teosinte before domestication: Experimental study of growth and phenotypic variability in late Pleistocene and early Holocene environments. *Quat Int*, 10.1016/j.quaint.2013.12.049.
- 100 Fuller DQ, Harvey E, Qin L (2007) Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* 81(312):316–331.
- 101 Willcox G (2004) Measuring grain size and identifying Near Eastern cereal domestication: Evidence from the Euphrates valley. *J Archaeol Sci* 31(2):145–150.
- 102 Klyver TA, Charles M, Jones G, Rees M, Osborne CP (2013) Did greater burial depth increase the seed size of domesticated legumes? *J Exp Bot* 64(13):4101–4108.
- 103 Blumler MA (1998) *Evolution of Caryopsis Gigantism and Agricultural Origins* (Binghamton Univ, New York).
- 104 Piperno DR, Pearsall DM (1998) *The Origins of Agriculture in the Lowland Neotropics* (Academic, San Diego, CA).
- 105 Richerson PJ, Boyd R, Bettinger RL (2001) Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. *Am Antiq* 66(3):387–411.
- 106 Cunniff J, Charles M, Jones G, Osborne CP (2010) Was low atmospheric CO₂ a limiting factor in the origin of agriculture? *Environmental Archaeology* 15(2):113–123.
- 107 Sage RF (1995) Was low atmospheric CO₂ during the Pleistocene a limiting factor for the origin of agriculture? *Glob Change Biol* 1(2):93–106.
- 108 Asplund L, Hagenblad J, Leino MW (2010) Re-evaluating the history of the wheat domestication gene NAM-B1 using historical plant material. *J Archaeol Sci* 37(9):2303–2307.
- 109 Roullier C, Benoit L, McKey DB, Lebot V (2013) Historical collections reveal patterns of diffusion of sweet potato in Oceania obscured by modern plant movements and recombination. *Proc Natl Acad Sci USA* 110(6):2205–2210.
- 110 Belfer-Cohen A, Goring-Morris AN (2011) Becoming farmers. *Curr Anthropol* 52(5):S209–S220.
- 111 Hayden B, Canuel N, Shanse J (2013) What was brewing in the Natufian? An archaeological assessment of brewing technology in the Epipaleolithic. *J Archaeol Method Theory* 20(1):102–150.
- 112 Zeder MA, Smith BD (2009) A conversation on agricultural origins. *Curr Anthropol* 50(5):681–690.

Convergent evolution and parallelism in plant domestication revealed by an expanding archaeological record

Dorian Q Fuller^{a,1}, Tim Denham^b, Manuel Arroyo-Kalin^a, Leilani Lucas^a, Chris J. Stevens^a, Ling Qin^c, Robin G. Allaby^d, and Michael D. Purugganan^{e,f}

^aInstitute of Archaeology, University College London, London WC1H 0PY, United Kingdom; ^bSchool of Archaeology and Anthropology, Australian National University, Canberra, ACT 0200, Australia; ^cSchool of Archaeology and Museology, Peking University, Beijing 100871, China; ^dSchool of Life Sciences, University of Warwick, Warwick CV35 9EF, United Kingdom; ^eDepartment of Biology, New York University, New York, NY 10003; and ^fCenter for Genomics and Systems Biology, New York University Abu Dhabi Institute, Abu Dhabi, United Arab Emirates

Edited by Dolores R. Piperno, Smithsonian National Museum of Natural History and Smithsonian Tropical Research Institute, Fairfax, Washington, DC, and approved November 15, 2013 (received for review September 11, 2013)

Recent increases in archaeobotanical evidence offer insights into the processes of plant domestication and agricultural origins, which evolved in parallel in several world regions. Many different crop species underwent convergent evolution and acquired domestication syndrome traits. For a growing number of seed crop species, these traits can be quantified by proxy from archaeological evidence, providing measures of the rates of change during domestication. Among domestication traits, nonshattering cereal ears evolved more quickly in general than seed size. Nevertheless, most domestication traits show similarly slow rates of phenotypic change over several centuries to millennia, and these rates were similar across different regions of origin. Crops reproduced vegetatively, including tubers and many fruit trees, are less easily documented in terms of morphological domestication, but multiple lines of evidence outline some patterns in the development of vegetative systems across the New World and Old World tropics. Pathways to plant domestication can also be compared in terms of the cultural and economic factors occurring at the start of the process. Whereas agricultural societies have tended to converge on higher population densities and sedentism, in some instances cultivation began among sedentary hunter-gatherers whereas more often it was initiated by mobile societies of hunter-gatherers or herder-gatherers.

archaeobotany | Neolithic | agriculture | archaeology | vegeticulture

Domestication offers an ideal laboratory for understanding evolution because it is a recent phenomenon in terms of geological time scales and because the selection pressures that affect harvestability by humans are often known (1). Domestication is a product of human behaviors that regulate or increase food supply, but may also inadvertently lock humans into an increased reliance on managed taxa (2). Archaeological research provides a fossil record of past organisms undergoing domestication, often accompanied by cultural artifacts associated with habitat management or niche construction (3, 4). The effects of agriculture in terms of intensifying land productivity to support larger populations has been fundamental to the development of civilizations and the ongoing impact on and management of ecosystems (5, 6).

Domestications have occurred separately on different continents and in different cultural traditions, and thus represent a set of parallel experiments from which to infer recurrent processes (Fig. 1). In some cases this represents parallelism of phylogenetically related organisms that have been subjected to similar selection pressures and developed identical or similar adaptations in different places. In others, we can consider domestication as convergent evolution, in as much as similar adaptations have evolved across crops in different plant families. These parallel adaptations have been defined as the “domestication syndrome” (7, 8). A distinction can be made between true

convergence, in which analogous states have been reached from very different and unrelated starting points, versus parallelism, in which similar pathways of change follow on from similar starting points, for example, as with taxa that share the same underlying developmental ontogeny and orthologous genetic loci (9). Some domestication traits, such as seasonality controls, have evolved in parallel across many species on the basis of the same genetic and developmental mechanisms; other traits, such as loss of wild-type seed dispersal or changes to seed and fruit size, have been attained through homoplasy based on different genetic and developmental changes. Similarly, in terms of the trajectories of domestication and their cultural causes, we can consider whether these were truly parallel, as is the case for wheat, barley, and Chinese rice, or have converged on similarly domesticated forms through different pathways, as seen for domesticated pulses (Fabaceae) and probably African pearl millet (9, 10).

This paper develops the perspective of domestication as a laboratory to understand evolution in relation to different cultural contexts of domestication. It examines whether instances of early agriculture worldwide converged through very different processes and from different starting points, or were parallel in terms of working from the same behavioral and botanical materials. We offer an updated review on the archaeobotany of plant domestication, including key processes for both plant and cultural evolution. For the purposes of this paper we will use the term “cultivation” to refer to a group of behaviors aimed at modifying soil environments and the management of the plants that grow in them. “Domestication” will be restricted to phenotypic

Significance

Agriculture was a transformative development in the history of human societies and natural environments and drove the evolution of new domesticated species. Crop plants are the predominant domesticated species in all agricultural systems and are an essential component in most the food production systems that underpinned the development of urban societies. Archaeological plant remains provide a range of insights into the processes by which plants were domesticated in different parts of the world. The present paper provides a unique synthesis of evidence, including quantitative evidence on the trajectory and rate of domestication in seed crops and patterns in the development of tropical vegetatively propagated crops.

Author contributions: D.Q.F., R.G.A., and M.D.P. designed research; D.Q.F., T.D., M.A.-K., L.L., C.J.S., and L.Q. performed research; D.Q.F. analyzed data; and D.Q.F., T.D., M.A.-K., L.L., C.J.S., L.Q., R.G.A., and M.D.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: d.fuller@ucl.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1308937110/-DCSupplemental.

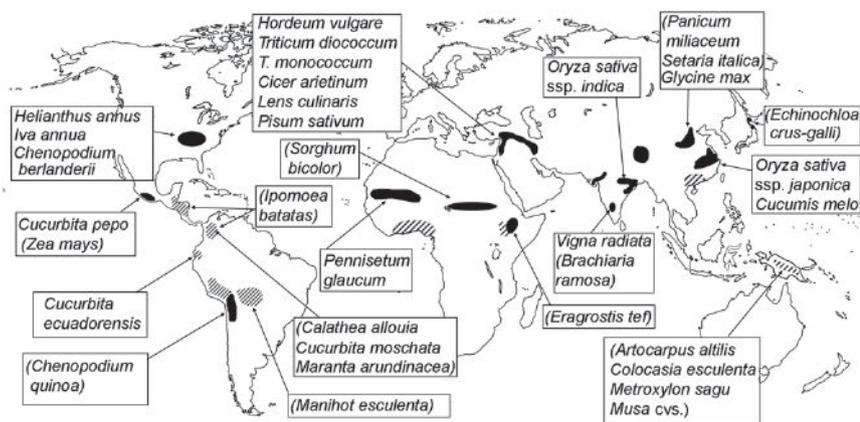


Fig. 1. Map of centers of domestication. Black areas indicate key areas of early seed crop domestication and hatched regions have an early focus on vegetative propagation. Species with quantified domestication rates are indicated, whereas others (species in parentheses) are discussed in the text.

changes in cultivars that make them different from unmanaged wild populations. Such phenotypic changes necessarily increase gradually at a population level, and therefore form a process, or an episode (which may take centuries or millennia) (2, 11). While selection on crops continues to occur and is behind varietal improvement and crop diversification, we interpret the “domestication episode” as the period in which key domestication syndrome traits underwent directional change and approached fixation within cultivated populations: these traits are now normally shared among all populations of a specific crop. “Agriculture” represents systems of land use in which cultivation behaviors became dominant, with domesticated species often the major cultivated taxa, and which took place at a large enough scale to become the primary economic activity of past populations. In this sense domestication emerges following a period of predomestication cultivation (12), with agriculture an outcome of both cultivation and domestication.

Advances in the Archaeobotany of Plant Domestication

Archaeobotany consists of the recovery and study of plant remains from archaeological sites. Although methods of recovery and analysis have improved, there has also been an increase in the number of sites and species studied and a broadening range of geographical regions subject to analysis (7, 9, 13–16). Along with botanical field studies of wild relatives of crops, this has led to recognition of a larger number of centers of agricultural origins, perhaps more than 20 (1, 8). For a number of crops, it is now becoming possible to compare domestication processes across species and geographical centers, thus allowing us to characterize similarities and differences. The present paper makes explicit comparisons of seed crop domestication and vegetative domestication processes from several regions.

Cereal crops are at the core of many agricultural systems and the seeds are highly visible in the archaeological record and, thus, lend themselves to quantitative studies on domestication and agricultural origins. Further, they may be directly dated by radiocarbon (15). The past decades have seen considerable methodological advances in identification criteria for crop subspecies and cereal varieties (e.g., refs. 16 and 17). Whereas the morphological distinction between wild and domesticated cereals has long been recognized (18), it is only in the last decade that substantial quantities of preserved remains have become available from Southwest Asia (7, 19, 20) and from Asian rice (21). In the New World, important pseudocereals, such as *Chenopodium* spp., have seen increased attention in terms of seed coat traits that relate to germination inhibition, an important target of selection during domestication (9, 15). In a wide range of taxa, metrical traits of seed size or phytolith size can be compared over time.

Where vegetative propagation was the focus of food production, poor preservation has made domestication traits harder to document. Advances have been made, however, in the study of phytoliths

and starches, most notably of tuber crops, as well as evidence of landscape modification (22, 23). In the Neotropics, the field has been revolutionized by microfossil evidence of domesticated crops that precede, in some cases by many millennia, empirical evidence for established agricultural practices (22, 24). Additional inferences have relied on cross-examining geochronological, genetic, and botanical evidence, suggesting, for example, the influence of cultivation practices on the development of varietal differences of manioc (*Manihot esculenta*) (25). In New Guinea, early human-managed habitats have been inferred. For instance, early agriculture based on the vegetative propagation of plants, including bananas, taro, and some yams, has been dated to 7000–6400 B.P. based on archaeological remains of former cultivation plots on old land surfaces, dramatic degradation of montane rainforest to grasslands, and microbotanical evidence for high frequencies of crop plants (26, 27). Plant microremains have also recently allowed recognition of palm (sago) starch consumption in tropical South China (28) and Borneo (23).

Where evidence for the presence of morphological domesticates is available alongside regional environmental modification, it appears that agriculture succeeds the establishment of domesticated crops. In early Neolithic Europe, archaeobotanical evidence for the presence of domesticated cereals in archaeological sites precedes palynological indicators of forest reduction and increases in arable pollen indicators (29). Early European and Anatolian weed floras and N^{15} isotope data from cereal grains indicate that these first crops were manured, suggesting intensively managed, small-scale fields, or grain “gardens” (30). The earliest preserved field systems for rice cultivation in China, ca. 6000 B.P., indicate small individual fields less than 2 m in diameter which allowed the careful management of soil and water conditions (31). As in Europe, regional deforestation in China proceeds gradually after this period (32). In the millet-dominated area of northern China, forest reduction occurred from 5000 B.P., millennia after domesticated millet production was widespread (33), whereas deforestation of the South Indian hills is evident around 3500 B.P., centuries after the first village-farming cultures (34).

Similar delays in agricultural systems are evident in the New World. In the Eastern Woodlands of North America, morphological changes document the domestication of several seed crops by 4500–4000 B.P. (35, 36), although a shift toward these crops over wild nuts only happened after 2,000 y ago (37). In the Balsas region of Mexico, starch and phytoliths indicate the exclusive processing of domesticated maize, without comparable finds of wild teosinte, at 8700 B.P., whereas lake core data in the region indicate landscape modification for slash-and-burn agriculture starting from 7600 B.P. (38, 39). In the Nanchoc Valley (northwestern Peru), starch from human teeth indicates the consumption of a range of crops, including introduced domesticates by 8000 B.P.; this is some 2,000 y before evidence for irrigation ditches indicate agricultural landscape modification in

the region (40). In the tropical lowlands of eastern South America, anthropogenic soils and raised fields occur millennia after the appearance of fossil remains of domesticated crops in the region (22, 25, 41). Thus agriculture itself as a system of major landscape modification was a convergent development that evolved after the establishment of cultivation and morphological domestication of crops.

Comparing Timing and Tempo of Domestication in Seed Crops

The domestication syndrome is likely to differ for various crop plants, according primarily to how they are reproduced (by seed or by cuttings) and according to which plant organ is the target of selection. The best-defined and -studied domestication syndrome is for grain crops, including cereals, pulses, and oilseeds (7, 9). Foremost are the traits selected by harvesting and a crop's growing reliance on humans for seed dispersal, including the loss of natural seed dispersal mechanisms. Second, there are traits related to the more friable soil conditions within tilled fields, leading to the deeper burial of seeds. The increase in seed size seen in most crops is hypothesized to aid seedling establishment, including from deeper burial (1, 7, 9), and is the most widely documented change in archaeobotanical evidence. Another key change is the loss of germination inhibition, in which germination occurs shortly after planting; this is regarded as the key domestication trait of pseudocereals like *Chenopodium* (7, 15, 35) and many pulses (9). For some taxa this is visible in preserved seed coat structure.

Nonshattering is often taken as a sine qua non of domesticated seed crops, making these species dependent on humans for reproduction by planting (1, 18, 19). In cereals this difference can be documented by the preserved abscission scar on the base of spikelets or rachis segments. Domestication in terms of this trait took at least 2,000–2,500 y (Fig. 2). Grain size change is more gradual over the same period as the shift from shattering (wild-type) to nonshattering seed dispersal (Fig. 2). After the episode of domestication, grain size becomes variable, fluctuating both up and down, suggesting processes of varietal differentiation and local adaptation; however, by the time of diversification, directional selection of the domestication episode is complete. Nonshattering becomes fixed at ~100% in wheat and barley, whereas percentages as low as 70% are returned from archaeological rice populations, due to the persistence of weedy rices as a major contaminant of fields. Fig. 3 summarizes the variation in

domestication rates and inferred coefficients of selection across 15 taxa and 18 traits (Tables S1–S4).

Evolutionary change in nonshattering is generally faster than grain size change, especially when measured in terms of haldanes (H). A haldane represents a change of one SD of a trait value per generation (11). For emmer wheat (*Triticum dicoccum*), the low estimate of H (Fig. 3) may be due to the relatively limited sample size available, as expanded datasets for *Triticum monococcum* and *Hordeum vulgare* have increased the estimates of rate somewhat over those made previously from a smaller dataset (11). Compared with reported phenotypic evolution rates in wild plant and animal studies that average around 0.03 H, our range of phenotypic evolution rates in domestication traits are generally similar. These rates of phenotypic change can be used to estimate the coefficient of selection, namely, the average increase per generation in gene(s) for a trait, which ranges from 0.0007 for nonshattering in *T. dicoccum* to ~0.1 in *T. monococcum* and *H. vulgare*.

Rates of phenotypic change for nonshattering were often one or two orders of magnitude greater than for grain size (Fig. 3), although this was still a protracted process. In the case of pearl millet (*Pennisetum glaucum*), domesticated in West Africa, chaff impressions in early Malian pottery indicate nonshattering types predominated by 4000 B.P., suggesting a preceding domestication episode of at least 1,000–2,000 y (10). In the case of maize (*Zea mays mays*), phytoliths indicative of nonshattering and absence of wild teosinte (*Zea mays parviglumis*) in the Rio Balsas region of Mexico at 8700 B.P. suggest that the domestication process and evolution of nonshattering occurred earlier (39). However, a sequence of evidence tracking the transition from nonshattering to shattering, or grain size changes, during this period is not yet available. For other cereals (various millets), archaeological evidence for nonshattering is similarly elusive. Remains of sorghum (*Sorghum bicolor*) from sites in the eastern Sahara indicate the consumption of wild (shattering) forms from 10000 to 5500 B.P. (42, 43), whereas the first domesticated forms are only documented following the crop's introduction to India after 4000 B.P. (44). This suggests that domestication took place in Africa between 6000 and 4000 B.P., but hard evidence for the process is lacking. Widespread use of cultivars suggests they had evolved by 8000–7500 B.P. for north Chinese millets (45, 46) and before 4500 B.P. for some Indian millets (44).

A more widely documented domestication trait in seed crops is increased seed size (2, 47). Most seeds increase by 20–60% in one or two dimensions, mainly thickness or breadth (Figs. S1 and

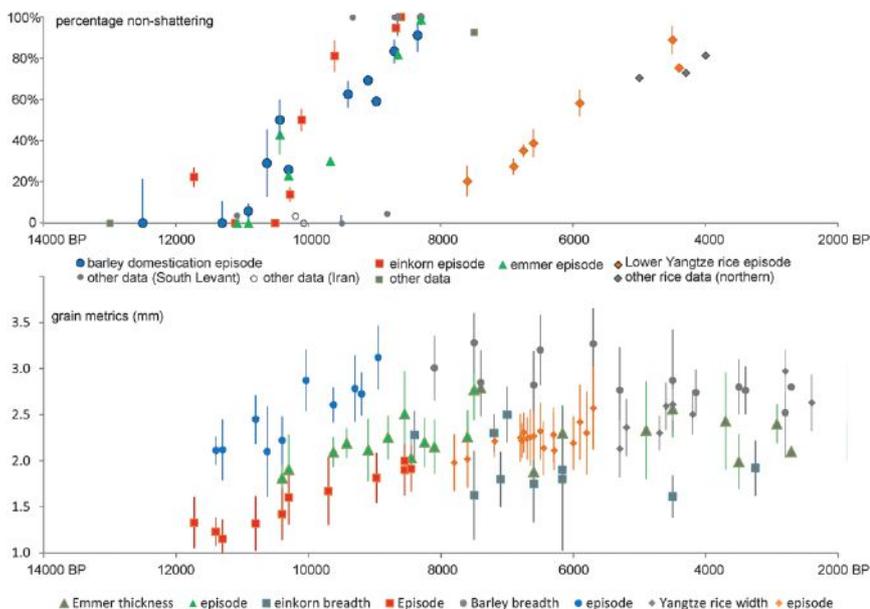


Fig. 2. Evidence for protracted domestication episodes in Old World cereals, including proportion of nonshattering spikelet scars (*Upper*) and grain size indices (*Lower*).

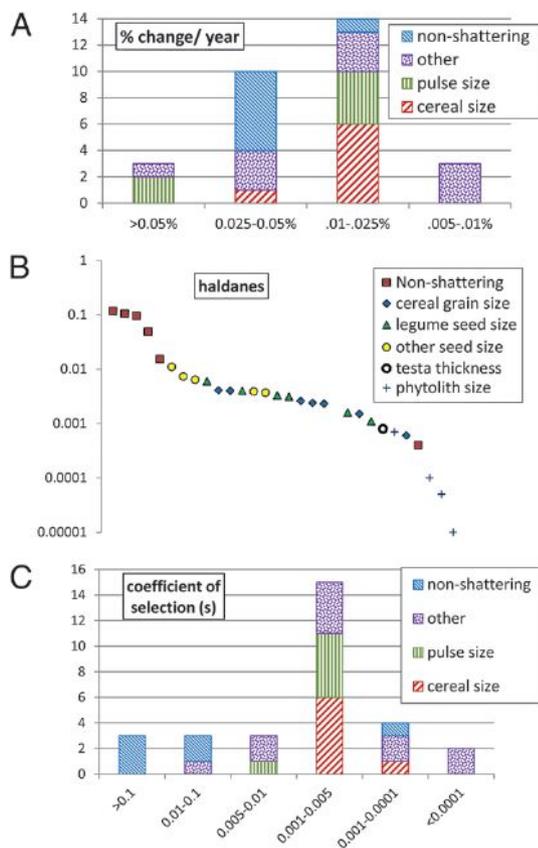


Fig. 3. Rates of change in domestication traits across the collected dataset. (A) Graph comparing frequency of rates in terms of percentage change in trait per year. (B) Scatter of all haldane rate estimates, indicating trait/crop type. (C) Frequency of estimated selection coefficients in the dataset.

S2 and Table S1). A 20% increase is reported for barnyard millet (*Echinochloa crus-galli*) domestication in northern Japan over 2,000–3,000 y (48), whereas the smallest change in the current dataset is the 19% increase in peas. At the upper end are reported thickness changes in some millets, including a 72% difference in grain thickness between modern wild and domesticated *Eragrostis tef* (49), and 76% in archaeological *P. glaucum*. A few taxa show much more dramatic increases on the order 100% or more, i.e., doubling seed dimensions, such as in Indian mungbean (*Vigna radiata*), Chinese soybean (*Glycine max*) (50), and North American *Iva annua* (51). The rate of change ranges between 0.0006 and 0.11 H. Evidence for einkorn wheat at early occupations on Cyprus suggests the rate of change in domestication traits was accelerated in an island context, perhaps due to genetic isolation from the wild population or bottleneck effects (52). Here einkorn grain size increased at 0.06 H as opposed to 0.006 H on the Eurasian mainland (52).

Cucurbit seeds appear to be closer to the lower end of size increases, with about 15% increase found with the domestication of *Cucurbita pepo* and recorded through the squash seeds of Guila Naquitz in southern Mexico 10000–8000 B.P. (15). In the case of Chinese melons, mean seed sizes increase by about 22% over 500 y, with a moderately fast 0.0073 H. Phytolith size in squash rinds also correlates with increasing fruit size (Fig. S3) showing a significant size increase in Ecuadorean assemblages (>40%) over the Early Holocene (53), but with a low H (1×10^{-5}). A more rapid increase (1×10^{-4} H) is associated with an initial domestication that took place by 11000 B.P.

Seed size increase may not always be an indicator of domestication. In some starchy fruits, reduction of seeds correlates with increase in starch content, as was clearly the case in the development

of domesticated bananas (*Musa* cvs.) (54). In some cases, regional selection trajectories have been divergent. For example, breadfruit (*Artocarpus altilis*) has been selected for seedless cultivars with larger fruits (consumed for edible pulp), as well as seeded cultivars in which seeds are eaten. The degree of seediness decreases eastwards and away from the New Guinea region (55).

Comparing Timing of Domestication in Vegetational Crops

It has generally been thought that vegetative reproduction made the domestication of tuberous plants possible through piecemeal replication of the characteristics of parent clones followed by selection and multiplication of useful phenotypic variations arising in planted stock (e.g., larger and smooth-skinned tubers, or less toxic/bitter forms). Unlike domesticated seed crops, some vegetative crops are not dependent on human efforts to reproduce and spread (23). However, sexual reproduction cycles and cooption of volunteer wild seedlings provide an important source of genetic diversity and local adaptations in at least some tubers (56). Indeed, it is clear that domesticated forms have been genetically altered from their wild progenitors on the basis of differing functional traits, such as those relating to toxicity, as well as on evidence for genetic bottlenecks between cultivars and wild relatives (57). Although increase in tuber size may be expected to correlate with an increase in the size of individual parenchyma cells, archaeologically recovered tuber fragments tend to preserve few morphological attributes relevant to phenotypic change. However, some research suggests that microremains such as starch grains have increased in size with tuber domestication (14, 58) and banana phytoliths have increased in size between diploid and triploid cultivars (59). Notwithstanding, compared with seed crops, it is harder to document archaeologically phenotypic changes in crops that are vegetatively propagated, such as tubers and some fruits, e.g., banana, grape, and olive (60, 61).

In some cases, cultivation practices may induce phenotypic alteration without genotypic change. Larger tubers develop in yams replanted in loosened, prepared soil as opposed to harder unprepared soils (62, 63). Thus some tuber crops could be cultivated for long periods without undergoing morphological domestication. Recent work on SSR molecular markers of live germplasm of *Ipomoea batatas* (64) and *Manihot esculenta* (65) permit disentangling complex histories of domestication. The latter study provides support for the argument (25) that one of the two manioc macro varieties (sweet manioc, the lower starch yielder that is less resistant to pests and poor soils) was domesticated earlier than bitter manioc (which grows well on poor soils, is resistant to pests, and yields more starch). The sequence most likely reflects an initial selection of plants in dump heaps with subsequent cultivation and range expansion. Subsequently in Amazonia, detoxification and cultivation techniques were innovated and led to the relaxation of selective pressures against cyanogenic glucosides to the point where a more toxic, but high starch-yielder evolved (66). Recent research has identified traits that facilitate vegetative propagation of cropped manioc, such as pronounced parenchymatous swellings at the nodes leading to brittle stems that can be readily broken for replanting, are absent in wild relatives (67).

Although it is not possible yet to provide quantitative data on domestication rates for vegetatively reproduced crops, we can provide comparisons of some general trajectories. In both New Guinea and Central/South America, evidence for consumption of starchy plants has been found back to the start of the Holocene or into the Pleistocene, whereas pollen evidence indicates human-disturbed forests, often including intentional burning (22, 24, 27, 68). This indicates human management of the landscape, within which some plants could be encouraged and planted. Artificial cultivation mounds in New Guinea (26, 27) suggest cultivation of tubers by the Mid-Holocene (7000–6400 B.P.), whereas microfossils beyond the range of the wild species, as well as dedicated lithic tools for cultivation and processing of roots, suggest tuber cultivation in Central and South America by 7000 B.P. (22, 24). In New Guinea, agricultural dependence and

intensification of production can be suggested from 5000 to 3500 B.P., with ditched fields, wooden spades, and probable sedentism (68). A similar timing is suggested for the introduction of rice and possible spread of banana in Southeast Asia (28, 69). In northern South America, sedentism associated with tuber crops and other cultivars (squashes, fruit trees, and maize) dates to the Mid-Holocene in different regions (70). These data suggest a shift from initial management of forest gaps and edges, where gap-colonizing species were exploited for food, to the creation of patches of food plants through planting. Vicariance induced by deliberate diffusion beyond the range of the wild species before the adoption of dedicated agricultural practices, along with the formalization of systematic plots where selection pressures and genetic isolation from wild populations increased, both played a role in the trajectories of domestication of vegetational crops.

Trajectories to Agriculture: Parallelism and Convergence in Cultural Evolution

The morphological changes of domestication are only one aspect of documenting the origins of agriculture. A domestication episode can be regarded as providing a species-specific time scale against which evidence for cultivation and management practices can be charted to reveal the interplay of human action and domestication. Although the evolution of domestication traits tends to increase the efficiency of harvests and yields, it also requires adjustments in human activities. Some of these may be characterized as labor traps, such as the additional requirement of threshing and winnowing as nonshattering rose to dominance, or the need to add nutrients to soils as erect crop growth habits packed more plants into the same units of soil, or the relocation of plots following nutrient depletion of soils (2). Thus, over the course of domestication there would have been fluctuations toward efficiency alternating with increased labor demands, but with an overall direction toward increasing yields and intensification of cultivation activities.

Taken at a comparative global level, the long-term impact of agricultural origins has been to support denser human populations through intensification of land use (5), including sedentism and fostering a greater reliance on a limited range of domesticated food stuffs. Although this represents convergence at a global level, the different domestication processes that can be documented across crops suggest that we should also look for multiple cultural patterns of agricultural origins, for example, in terms of the mobility of past societies and nature of crop reproduction.

Pristine domestications of crops have often occurred within mobile societies, which might include either hunter-gatherers or noncultivating pastoralists. In the case of vegetational origins, both in New Guinea and Central America, sedentism dates back only 4,000–5,000 y, long after the earliest inferred cultivation in these regions. Similarly seed crop cultivation of maize in Mesoamerica precedes settled villages by ~5,000 y. In northern China, settled village farmers of millets date from at least 6500 B.P., whereas millet exploitation occurred by 9500 B.P., with clear cultivation by 8000 B.P. (45, 71). Both maize and Chinese millet cultivation correlate with periods of climatic amelioration in the Early Holocene (39, 46). In the case of the Old World savannah millet domestications, mobile gatherer-pastoralists with domesticated ungulates entered the Sahel of western Africa and then cultivated millet, whereas in peninsular India seasonally mobile herder-hunters precede sedentism or crops (44). In these cases of Mid-Holocene grain domestication, there was climatic aridification shortly before evidence for cultivation, including

desertification processes of the Sahara or savannah expansion in South India (10, 44). In eastern North America, cultivation began among seasonally mobile Late Archaic hunter-fisher-gatherers ca. 5000 B.P. (15, 35). Among the key factors suggested for domestication by mobile groups are risk avoidance and seasonal conflicts in resource availability, leading to cultivation to make such resources readily available when seasonally needed (44, 72) or to buffer risks in wild food availability (46, 73).

Although sedentism based on agricultural economies was to become universal, only in a few instances is there possible evidence for sedentary foragers involved in the initial cultivation of crops. In China's Yangtze Valley, substantial settlements at ~8000 B.P. indicate sedentism before morphological domestication of rice and alongside early cultivation (31, 32). In Southwest Asia, the Late Pleistocene Natufian culture is often regarded as sedentary or nearly so, by 13,500 y ago, and a precursor of the cultivating villages during the Pre-Pottery Neolithic from ca. 11500 B.P. (74). Although the extent of year-round sedentism at villages focused on farming can be queried (74), it is nevertheless the case that architecturally permanent sites became increasingly important for cultivation, storage, and consumption of early crops like wheats and barley. Early cultivation by the Jomon of Japan was in the context of sedentism (48). The adoption of crops in the Nanchoc Valley of northwestern Peru after 9000 B.P. is associated with sedentism (76). In these cases, pressure on resources and the need to support growing, and sedentary, populations are often considered part of the explanation of the origins of cultivation (77).

Conclusion

Agriculture is increasingly recognized as the coalescence of human activities and genetically transformed species that extends the widespread proclivity of *Homo sapiens* for niche construction (4–6) into a more intensive coevolutionary relationship that enhances the fitness, population size, and density of both humans and their crop plants. The pathways to agriculture were prolonged episodes of coevolution, genetic adaptations on the part of the plants, and cultural shifts and innovations on the part of people. These processes demand long-term and interregional comparative study. For seed crops, domestication trajectories are increasingly documented by quantitative patterns in archaeological plant assemblages, whereas for vegetative only some general outlines have begun to emerge. The development of cultivation among mobile forager societies focusing on tubers as well as seeds indicates some parallelism across the seed crop and vegetative crop divide, whereas sedentary collectors turned cultivators was a less common development. In terms of cultural history, the domestication of wild plant species has been a process of convergence from different regional, environmental, and economic starting points.

ACKNOWLEDGMENTS. This paper emerged from a meeting in Durham, NC (April 2011), supported by the National Evolutionary Synthesis Center, National Science Foundation (NSF) EF-0905606. Current research on Comparative Pathways to Agriculture (ComPAg) by D.Q.F., C.J.S., and L.L. is supported by European Research Council Grant 323842; research on early rice cultivation by D.Q.F. and L.Q. was supported by UK Natural Environment Research Council (London) Grants NE/G005540/1 and NE/K003402/1; research on early agriculture in the highlands of New Guinea was supported by Monash and Australian Research Council postdoctoral fellowships (to T.D.); research by M.A.-K. is supported by a British Academy postdoctoral fellowship; and research on the evolution of crop species by M.D.P. is supported by the US NSF Plant Genome Research Program and the New York University Abu Dhabi Institute.

- Purugganan MD, Fuller DQ (2009) The nature of selection during plant domestication. *Nature* 457(7231):843–848.
- Fuller DQ, Allaby RG, Stevens C (2010) Domestication as innovation: The entanglement of techniques, technology and chance in the domestication of cereal crops. *World Archaeol* 42(1):13–28.
- Laland KN, Odling-Smee J, Feldman MW (2000) Niche construction, biological evolution, and cultural change. *Behav Brain Sci* 23(1):131–146, discussion 146–175.
- Smith BD (2011) General patterns of niche construction and the management of 'wild' plant and animal resources by small-scale pre-industrial societies. *Philos Trans R Soc Lond B Biol Sci* 366(1566):836–848.
- Ellis EC, et al. (2013) Used planet: A global history. *Proc Natl Acad Sci USA* 110(20):7978–7985.
- Smith BD, Zeder MA (2013) The onset of the Anthropocene. *Anthropocene*, 10.1016/j.anucene.2013.05.001.

7. Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates: Recent archaeobotanical insights from the Old World. *Ann Bot (Lond)* 100(5):903–924.
8. Meyer RS, DuVal AE, Jensen HR (2012) Patterns and processes in crop domestication: An historical review and quantitative analysis of 203 global food crops. *New Phytol* 196(1):29–48.
9. Fuller DQ, Allaby R (2009) *Fruit Development and Seed Dispersal*. *Annual Plant Reviews*, ed Østergaard L (Wiley-Blackwell, Oxford), Vol 38, pp 238–295.
10. Manning K, Pelling R, Higham T, Schwenniger J-L, Fuller DQ (2011) 4500-year old domesticated pearl millet (*Pennisetum glaucum*) from the Tilemsi Valley, Mali: New insights into an alternative cereal domestication pathway. *J Archaeol Sci* 38(2): 312–322.
11. Purugganan MD, Fuller DQ (2011) Archaeological data reveal slow rates of evolution during plant domestication. *Evolution* 65(1):171–183.
12. Hillman G, Davies MS (1990) Domestication rates in wild wheats and barley under primitive cultivation. *Biol J Linn Soc Lond* 39(1):39–78.
13. Denham TP, Haberle SG, Pierret A (2009) *New Directions in Archaeological Science (Terra Australis 28)*, eds Fairbairn A, O'Connor S, Marwick B (ANU E Press, Canberra, Australia), pp 139–154.
14. Piperno D (2012) New archaeobotanical information on early cultivation and plant domestication. *Biodiversity in Agriculture. Domestication, Evolution, and Sustainability*, eds Gepts P, et al. (Cambridge Univ Press, Cambridge, UK), pp 110–135.
15. Smith BD (2006) Eastern North America as an independent center of plant domestication. *Proc Natl Acad Sci USA* 103(33):12223–12228.
16. Jones G, Valamoti S, Charles M (2000) Early Crop diversity: A “new” glume wheat from northern Greece. *Veg Hist Archaeobot* 9(3):133–146.
17. Fuller DQ, Willcox G, Allaby RG (2012) Early agricultural pathways: Moving outside the ‘core area’ hypothesis in Southwest Asia. *J Exp Bot* 63(2):617–633.
18. Helbaek H (1966) Commentary on the phylogenesis of *Triticum* and *Hordeum*. *Econ Bot* 20(4):350–360.
19. Tanno KI, Willcox G (2012) Distinguishing wild and domestic wheat and barley spikelets from early Holocene sites in the Near East. *Veg Hist Archaeobot* 21(2):107–115.
20. Riehl S, Zeidi M, Conard NJ (2013) Emergence of agriculture in the foothills of the Zagros Mountains of Iran. *Science* 341(6141):65–67.
21. Fuller DQ, et al. (2009) The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323(5921):1607–1610.
22. Piperno DR, Pearsall D (1998) *The Origins of Agriculture in the Lowland Neotropics* (Academic, San Diego).
23. Barton H, Denham T (2011) *Anthropological and Archaeological Approaches to Foraging-Farming Transitions in Southeast Asia*, McDonald Institute Monographs, eds Barker G, Janowski M (McDonald Institute for Archaeological Research, Cambridge, UK), pp 17–25.
24. Piperno DR (2011) The origins of plant cultivation and domestication in the New World tropics: Patterns, process, and new developments. *Curr Anthropol* 52(54): 53–470.
25. Arroyo-Kalin M (2010) The Amazonian Formative: Crop domestication and anthropogenic soils. *Diversity* 2(4):473–504.
26. Denham TP, et al. (2003) Origins of agriculture at Kuk Swamp in the highlands of New Guinea. *Science* 301(5630):189–193.
27. Haberle S, Lentfer C, O'Donnell S, Denham T (2012) The palaeoenvironments of Kuk Swamp from the beginnings of agriculture in the highlands of Papua New Guinea. *Quat Int* 249:129–139.
28. Yang X, et al. (2013) Sago-type palms were an important plant food prior to rice in southern subtropical China. *PLoS ONE* 8(5):e63148.
29. Willis KJ, Bennett KD (1994) The Neolithic transition—fact or fiction? Palaeoecological evidence from the Balkans. *Holocene* 4(3):326–330.
30. Bogaard A, et al. (2013) Crop manuring and intensive land management by Europe's first farmers. *Proc Natl Acad Sci USA* 110(31):12589–12594.
31. Fuller DQ, Qin L (2009) Water management and labour in the origins and dispersal of Asian rice. *World Archaeol* 41(1):88–111.
32. Fuller DQ, Qin L (2010) Declining oaks, increasing aristry, and cultivating rice: The environmental and social context of the emergence of farming in the Lower Yangtze Region. *Environmental Archaeology* 15(2):139–159.
33. Ren G (2007) Changes in forest cover in China during the Holocene. *Veg Hist Archaeobot* 16(2–3):119–126.
34. Kingwell-Banham E, Fuller DQ (2012) Shifting cultivators in South Asia: Expansion, marginalisation and specialisation over the long-term. *Quat Int* 249:84–95.
35. Smith BD (1989) Origins of agriculture in eastern North America. *Science* 246(4937): 1566–1571.
36. Smith BD (2006) *Documenting Domestication*, eds Zeder MA, Bradley D, Emshwiller E, Smith BD (Univ of California Press, Berkeley, CA), pp 25–31.
37. Milner GR (2004) *The Moundbuilders: Ancient Peoples of Eastern North America* (Thames and Hudson, London).
38. Piperno DR, et al. (2007) Late Pleistocene and Holocene environmental history of the Iguale Valley, Central Balsas Watershed of Mexico. *Proc Natl Acad Sci USA* 104(29): 11874–11881.
39. Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106(13):5019–5024.
40. Piperno DR, Dillehay TD (2008) Starch grains on human teeth reveal early broad crop diet in northern Peru. *Proc Natl Acad Sci USA* 105(50):19622–19627.
41. McKey DS, et al. (2010) Pre-Columbian agricultural landscapes, ecosystem engineers, and self-organized patchiness in Amazonia. *Proc Natl Acad Sci USA* 107(17): 7823–7828.
42. Stemler AB (1990) A scanning electron microscopic analysis of plant impressions in pottery from the sites of Kadero, El Zakiab, Um Direiwa and El Kadada. *Archéologie du Nil Moyen* 4:87–105.
43. Wasylkowska K, et al. (1995) Archaeobotany of the early Neolithic site E-75-6 at Nabta Playa, Western Desert, South Egypt (preliminary results). *Acta Palaeobotanica* 35(1): 133–155.
44. Fuller DQ (2011) Finding plant domestication in the Indian subcontinent. *Curr Anthropol* 52(4):S347–S362.
45. Liu X, Hunt HV, Jones MK (2009) River valleys and foothills: Changing archaeological perceptions of North China's earliest farms. *Antiquity* 83(319):82–95.
46. Bettinger RL, Barton L, Morgan C (2010) The origins of food production in north China: A different kind of agricultural revolution. *Evol Anthropol* 19(1):9–21.
47. Harlan JR, De Wet JMJ, Price EG (1973) Comparative evolution of cereals. *Evolution* 27(2):311–325.
48. Crawford G (2011) Advances in understanding early agriculture in Japan. *Curr Anthropol* 52(54):S331–S345.
49. D'Andrea AC (2008) T'ef (*Eragrostis tef*) in ancient agricultural systems of highland Ethiopia. *Econ Bot* 62(4):547–566.
50. Lee G-A, Crawford GW, Liu L, Sasaki Y, Chen X (2011) Archaeological soybean (*Glycine max*) in East Asia: Does size matter? *PLoS ONE* 6(11):e26720.
51. Asch DL, Asch NB (1985) Prehistoric food production in North America. *Anthropological Papers* 75, ed Ford RI (Museum of Anthropology, Univ of Michigan, Ann Arbor, MI), pp 149–204.
52. Lucas L, Colledge S, Simmons A, Fuller DQ (2012) Crop introduction and accelerated island evolution: Archaeobotanical evidence from 'Ais Yiorkis and Pre-Pottery Neolithic Cyprus. *Veg Hist Archaeobot* 21(2):117–129.
53. Piperno DR, Stothert KE (2003) Phytolith evidence for early Holocene *Cucurbita* domestication in southwest Ecuador. *Science* 299(5609):1054–1057.
54. Perrier X, et al. (2011) Multidisciplinary perspectives on banana (*Musa* spp.) domestication. *Proc Natl Acad Sci USA* 108(28):11311–11318.
55. Zerega NJC, Ragone D, Motley TJ (2004) Complex origins of breadfruit (*Artocarpus altilis*, Moraceae): Implications for human migrations in Oceania. *Am J Bot* 91(5): 760–766.
56. McKey D, Elias M, Pujol B, Duputié A (2012) *Biodiversity in Agriculture. Domestication, Evolution, and Sustainability*, eds Gepts P, et al. (Cambridge Univ Press, Cambridge, UK), pp 377–406.
57. Olsen KM, Schall BA (2006) *Documenting Domestication: New Genetic and Archaeological Paradigms*, eds Zeder MA, et al. (Univ of California Press, Berkeley, CA), pp 123–133.
58. Perry L (2002) Starch granule size and the domestication of manioc (*Manihot esculenta*) and sweet potato (*Ipomoea batatas*). *Econ Bot* 56(4):335–349.
59. Vrydaghs L, et al. (2009) Differentiating the Volcaniform Phytoliths of Bananas: *Musa acuminata*. *Ethnobotanical Research and Applications* 7:239–246.
60. Hather JG (1996) *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, ed Harris DR (UCL Press, London), pp 538–551.
61. Zohary D, Hopf M, Weiss E (2012) *Domestication of Plants in the Old World: The Origin and Spread of Domesticated Plants in Southwest Asia, Europe, and the Mediterranean Basin* (Oxford Univ Press, Oxford).
62. Chikwendu VE, Okezie CEA (1989) *Foraging and Farming: The Evolution of Plant Exploitation*, eds Harris DR, Hillman GC (Unwin Hyman, London), pp 344–357.
63. Hildebrand EA (2007) *Rethinking Agriculture. Archaeological and Ethnoarchaeological Perspectives*, eds Denham T, Vrydaghs L, Iriarte J (Left Coast Press, Walnut Creek, CA), pp 273–298.
64. Roullier C, et al. (2013) Disentangling the origins of cultivated sweet potato (*Ipomoea batatas* (L.) Lam.). *PLoS ONE* 8(5):e62707.
65. Mühlen GS, Alves-Pereira AA, Clement CR, Valle TR (2013) Genetic diversity and differentiation of Brazilian bitter and sweet manioc varieties (*Manihot esculenta* Crantz, Euphorbiaceae) based on SSR molecular markers. *Tipiti* 11(2):66–73.
66. Arroyo-Kalin M (2012) Slash-burn-and-churn: Landscape History and Crop cultivation in pre-Columbian Amazonia. *Quat Int* 249:4–18.
67. Ménard L, McKey D, Mühlen GS, Clair B, Rowe NP (2013) The evolutionary fate of phenotypic plasticity and functional traits under domestication in manioc: Changes in stem biomechanics and the appearance of stem brittleness. *PLoS ONE* 8(9):e74727.
68. Denham TP, *The Cambridge World Prehistory*, eds Bahn P, Renfrew C (Cambridge Univ Press, Cambridge, UK).
69. Fuller DQ, Madella M (2009) Banana cultivation in South Asia and East Asia: A review of the evidence from archaeology and linguistics. *Ethnobot Res Applications* 7:333–351.
70. Raymond JS (2008) The process of sedentism in northwestern South America. *Handbook of South American Archaeology*, eds Silverman H, Isbell W (Springer, New York), pp 79–90.
71. Yang X, et al. (2012) Early millet use in northern China. *Proc Natl Acad Sci USA* 109(10):3726–3730.
72. Marshall F, Hildebrand EA (2002) Cattle before crops: The beginnings of food production in Africa. *J World Prehist* 16(2):99–143.
73. Clement CR (2006) *Time and Complexity in the Neotropical Lowlands: Studies in Historical Ecology*, eds Balée W, Erickson CL (Columbia Univ Press, New York), pp 165–185.
74. Asouti E, Fuller DQ (2013) A contextual approach to the emergence of agriculture in Southwest Asia. *Curr Anthropol* 54(3):299–345.
75. Willcox G, Stordeur D (2012) Large-scale cereal processing before domestication during the tenth millennium cal BC in northern Syria. *Antiquity* 86(331):99–114.
76. Dillehay TD, Rossen J, Andres TC, Williams DE (2007) Preeramic adoption of peanut, squash, and cotton in northern Peru. *Science* 316(5833):1890–1893.
77. Cohen MN (2009) Rethinking the origins of agriculture. Introduction. *Curr Anthropol* 50(5):591–595.

Evaluating the roles of directed breeding and gene flow in animal domestication

Fiona B. Marshall^{a,1}, Keith Dobney^b, Tim Denham^c, and José M. Capriles^{d,e}

^aDepartment of Anthropology, Washington University in St Louis, St Louis, MO 63130; ^bDepartment of Archaeology, University of Aberdeen, Aberdeen AB24 3UF, Scotland, United Kingdom; ^cSchool of Archaeology and Anthropology, Australian National University, Canberra, ACT 0200, Australia; ^dInstituto de Alta Investigación, Universidad de Tarapacá, Casilla 6-D, Arica, Chile; and ^eCentro de Investigaciones del Hombre en el Desierto, Arica, Chile

Edited by Greger Larson, Durham University, Durham, United Kingdom, and accepted by the Editorial Board October 8, 2013 (received for review July 18, 2013)

For the last 150 y scholars have focused upon the roles of intentional breeding and genetic isolation as fundamental to understanding the process of animal domestication. This analysis of ethnoarchaeological, archaeological, and genetic data suggests that long-term gene flow between wild and domestic stocks was much more common than previously assumed, and that selective breeding of females was largely absent during the early phases of animal domestication. These findings challenge assumptions about severe genetic bottlenecks during domestication, expectations regarding monophyletic origins, and interpretations of multiple domestications. The findings also raise new questions regarding ways in which behavioral and phenotypic domestication traits were developed and maintained.

reproductive isolation | selected breeding | zooarchaeology | donkey | pig

Domestication resulted in diverse phenotypic and behavioral changes to wild animals, including decreased flight responses, increased sociality, earlier reproduction, and modification of endocrine and metabolic systems (1–4). Darwin's (5) seminal research, heavily influenced by European animal breeding practices during the 19th century, led subsequent scholars studying animal domestication to prioritize the central roles of human intentionality, directed or controlled breeding of individuals, and genetic isolation of captive herds from wild relatives (6). This anthropocentric legacy is evident in various widely used definitions of domestication that emphasize isolation of captive animals from wild species and total human control over breeding and animal care (6–8). However, a growing body of archaeological, genetic, and ethnohistorical evidence discussed here shows that neither reproductive isolation nor intentional breeding of individuals was as significant as traditionally thought. Our findings indicate long-term gene flow between managed and wild animal populations, and little control of breeding of domestic females. These findings challenge assumptions about severe genetic bottlenecks during domestication and interpretations of genetic variability in terms of multiple instances of domestication. The findings also raise questions about ways in which behavioral and phenotypic domestication traits were maintained.

Research into dog and pig domestication over the last several decades has drawn attention to the roles of nonhuman drivers in the domestication process (9, 10) with

early domestication routes for these taxa now widely viewed as commensal (3). Prey pathways provided other trajectories to domestication for goats, sheep, and cattle (11), whereas more directed routes to domestication have been proposed for animals such as donkeys (3). Despite these new emphases on varied human–animal relations, most models still rely on human-directed breeding over generations (3, 12, 13) and reproductive isolation to delineate all but the very earliest phases of domestication (14). The creation of separate breeding populations of animals, wholly isolated from their wild progenitors, persists as a fundamental assumption of classic speciation-based models (14, 15).

To date, there has been little discussion of how variabilities in the biology and behavior of captive animals, human environments, management regimes, and migration and dispersal of domestic animals affected directed breeding and gene flow between domestic and wild populations. These processes are explored here through archaeological, biological, ethnographic, and genetic evidence, focusing on large ungulates (Table 1).

Management and Gene Flow

Equids, Camelids, and Yaks. Humans have relied heavily on horses, donkeys, camelids, and yaks for transport, food, fiber, and ritual practices over the millennia. Physiologically well adapted to extreme environments, these animals enable mobile herders to survive in cold steppe, desert, and mountainous regions. With the exception of horses and yaks, transport animals are territorial and challenging

to manage; they are also large-bodied with correspondingly slow gestation and herd growth rates that do not permit high levels of culling. These biological influences on human management mean herders value the adaptations of wild relatives of their domestic animals, manage animals lightly, cull at low levels, and grow herds through capture of more wild animals. Consequently, transport animals reflect low levels of directed selection resulting from intentional human management, including breeding, culling, or castration of selected animals, and high levels of gene flow.

Donkey's desert adaptations, lack of sociability, long gestation rates, and use by mobile herders for long-distance movement have resulted in particularly low levels of management, little directed breeding, and constant gene flow with their wild and feral relatives, at least within their wild range. Much like cats, donkeys have often been treated as an exception to the accepted rules for domestication and, by definitions that focus on reproductive isolation (6, 8), they

Author contributions: F.B.M. designed research; F.B.M. and K.D. performed research; F.B.M., K.D., T.D., and J.M.C. analyzed data; and F.B.M., K.D., T.D., and J.M.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. G.L. is a guest editor invited by the Editorial Board.

¹To whom correspondence should be addressed. E-mail: fmarshall@wustl.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1312984110/-DCSupplemental.

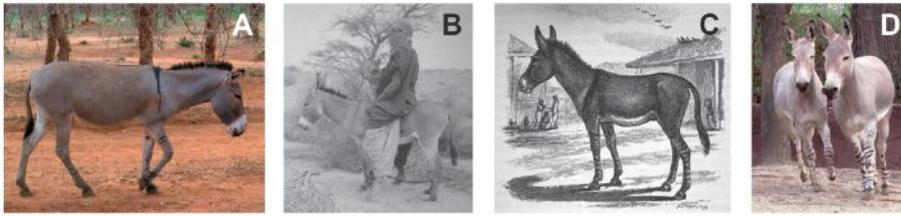


Fig. 1. Intentional capture and out-crossing of donkeys, wild asses, and hybrids. (A) African donkey with shoulder cross (Image courtesy of Lior Weissbrod). (B) Tuareg taming captured Saharan wild ass or feral donkey, 1951 (21) (Image courtesy of Ida Nicolaisen and the Carlsberg Foundation). (C) Donkey-Somali wild ass hybrid with cross and striped legs, Berbera 1900s. Donkeys were tied outside the village to breed with Somali wild asses (20). (D) Somali wild asses with striped legs.

could, perhaps, not even be considered a domestic animal.

African wild asses (*Equus africanus*) were the ancestors of domestic donkeys (16, 17) (Fig. 1 and Table 1). Today, African pastoralists rely on donkeys for transport and they are rarely slaughtered for food. As a result, drought and disease are the principal causes of donkey mortality. Herders value individual animals for strength and hardiness (18) and castrate difficult males, but prefer uncastrated ones for transport-use. The presence of multiple breeding males reduces directed selection (18). Moreover, because they are challenging to herd, donkeys range widely in search of mates and donkey-owners do little to manage reproduction (18). Slow herd growth and the value placed on the size, strength, and hardiness of transport donkeys led historic pastoralists (and Romans in North Africa) to capture feral donkeys and African wild asses, and to encourage interbreeding with wild males (19–21) (Fig. 1 and Table S1).

Modern pastoral use of donkeys presents a picture of weak directed selection principally resulting from castration and strong environmental selection. Environmental

selection (15) primarily relates to unconscious or natural selection, resulting from mortality because of the effects of events, such as drought, disease, and predation on managed animals (Table S2). In regions where wild asses existed historically, continued gene flow resulted from managed or inadvertent breeding of domestic donkeys with wild asses (19–21). These aspects of the recent past are relevant to understanding ancient processes (22) because they reflect consistent mechanisms, biology, and transport use.

Archaeological and genetic data support conclusions that donkeys were domesticated in arid environments, bred with a variety of wild populations, and were used for transport and trade over long distances (Fig. S1). Archaeological evidence for specialized hunting of territorial desert asses goes back *ca.* 16,000 y in northeast Africa (18). However, desert assemblages are rare and evidence is lacking for the likely period of earliest management 9000–6000 B.P. (all dates are reported in calibrated years before present). The presence of two divergent mitochondrial lineages in donkeys has been interpreted as evidence for more than one domestication,

but may be equally consistent with recurrent recruitment of females into domestic herds from genetically divergent Nubian wild ass populations (16, 17). A reduction in the size of some asses, often accepted as indicative of domestication, is first documented at Maadi in Egypt *ca.* 6000 B.P. (23). A thousand years later, despite expectations for significantly smaller animals, metacarpals from equids ritually buried at Abydos still fall within the size range of wild asses (19). Nevertheless, pathologies indicative of loading demonstrate that these morphologically wild animals were used for transport (19). Size decrease appears slow and inconsistent through time, with variability within and between archaeological sites indicating a nonlinear process of phenotypic change.

Herder reliance on donkeys for transport, the behavior of donkeys, and the long-term presence of wild asses near the Nile suggest that weak directed selection, continued recruitment of animals from the wild, and gene flow with wild asses contributed significantly to phenotypic variability among Predynastic and Early Dynastic donkeys in Egypt over at least a 2,500-y period. The value that donkey herders placed on strength is demonstrated by donkey-onager and subsequent donkey-horse hybrids (mules) bred in the ancient Near East (7, 24). Uncontrolled breeding among village donkeys and along trade routes also contributed to gene flow between founder populations and mitigated genetic drift (17, 18).

Zooarchaeological evidence, ethnographic observations, and genetic data suggest herd management has always been *laissez faire* and characterized by intentional and unintentional interbreeding with wild asses and feral donkeys, as well as by environmental selection for animals that survived in pastoral settlements. Together, these processes resulted in a prolonged and complicated process of domestication for donkeys.

Ethnographic and archaeological data for horses, Bactrian camels, dromedaries, llamas, alpacas, and yaks provide further insights into biological and human social factors affecting selective breeding and gene flow during the domestication of transport animals. Extinct *Equus ferus* from central Asia was the wild ancestor of domestic horses (Table 1 and Fig. S1). Evidence for biting, milking, corralling, and size decrease documents domestication by horse-hunters at Botai in Kazakhstan *ca.* 5500 B.P. (25, 26). As with other species, mitochondrial DNA lineages were often initially interpreted in terms of multiple origins (25, 27), whereas genetic modeling now suggests domestication in a restricted region

Table 1. Domestic animals, key archaeological sites, and domestication time-ranges

Animal	Domestication	Sites	Sources
Donkey, <i>Equus asinus</i>	6000–3500 B.P.	Maadi, Abydos, Uan Muhuggiag	17, 19, 23
Horse, <i>Equus caballus</i>	5500 B.P.	Botai	25, 26, 28
Bactrian camel, <i>Camelus bactrianus</i>	6000–4000 B.P.	Anau	29–31
Dromedary, <i>Camelus dromedarius</i>	4000–3000 B.P.	Shahr-i-Sokhta	35–37
Llama, <i>Lama glama</i>	6000–4000 B.P.	Pikimachay, Tulan, Inca Cueva	39–42, 44
Alpaca, <i>Vicugna pacos</i>	5000–3000 B.P.	Telarmachay	39–42, 44
Pig, <i>Sus scrofa</i>	12000–8300 B.P.	Çayönü Tepesi, Jiahu	10, 48–52, 54
Goat, <i>Capra hircus</i>	11000–9000 B.P.	Asiab, Ganj Dareh, Ali Kosh	58, 63
Sheep, <i>Ovis aries</i>	12000–10500 B.P.	Cafer Hüyük, Zawi Chemi Shanidar	56–58
Taurine cattle, <i>Bos taurus</i>	10500–10000 B.P.	Dja'de, Çayönü	66, 67
Zebu cattle, <i>Bos indicus</i>	8000–7500 B.P.	Mehrgarh	68, 69
Yak, <i>Bos grunniens</i>	?	Tibetan Plateau	45, 46

Wild-domestic gene-flow occurred among all taxa. Large transport animals were subject to low culling and high out-crossing potentials.

with subsequent incorporation of many different wild lineages into domestic stocks (28). Horse herds grow slowly and are subject to die-offs in severe storms, so the hardiness of wild horses is advantageous to herders. Accordingly, it has been argued that difficulties in maintaining domestic horse herd sizes during pastoral migrations led directly to restocking through the capture of wild females (25, 28).

Another transport animal subject to long-term gene flow is the Bactrian camel. Evidence is sparse, but ancient populations related to *Camelus ferus* are thought to have been domesticated in cold desert regions of Central Asia (Table 1 and Fig. S1). The presence of Bactrian camels found outside their likely wild range suggests domestication ca. 6000–4000 B.P. (29), with a geographically restricted domestication indicated by genetic data (30, 31). Extinction of their closest wild relatives (30) is thought to have resulted from both hunting and introgression with domestic camels (32). Historically herders have relied heavily on the strength of domestic Bactrian-dromedary crosses (33). Possibilities for increased strength and resilience may also have led nomads to encourage breeding of early domestic and wild camels, with chance admixture more likely occurring within their natural range (34). The domestication of a related camelid—the dromedary—also indicates both intentional and chance breeding of domestic and wild camels. Dromedaries are adapted to hot deserts and were domesticated in Arabia (35). Their wild ancestor (*Camelus* sp.) is now extinct (36) but increased frequencies of dromedaries at archaeological sites suggest domestication ca. 4000 B.P. (36, 37). Ethnographic data show that herders select bulls based on factors including size, color, family milk yields, and environmental adaptations (38), but all females are bred. Culling takes place at low levels and principally affects males, therefore directed selection is low. In contrast, high environmental selection on domestic camel herds is indicated by camelid genetics (30, 35). As shown by Bactrian-dromedary crosses, strength and hardiness were important to ancient herders and admixture is thought to have played a role in wild camelid extinctions.

There is also strong evidence for wild-domestic admixture and weak directed selection among domestic South American llama and alpaca and their wild relatives, guanaco (*Llama guanicoe*) and vicuña (*Vicuna vicuna*). These camelids are adapted to Andean high-altitude environments (Table 1 and Fig. S1). Zooarchaeological research suggests multiple processes of domestication

by hunters and possibly early cultivators in the central and south central Andes ca. 6000–4000 B.P. (39, 40). Archaeological and ethnographic data indicate that, although initially used for meat, herders have increasingly relied on larger llamas for transport and managed alpacas for fiber production. In the Lake Titicaca basin, the zooarchaeological record documents increasingly intensified and controlled herding, continued hunting, and gene flow among camelids 3500–900 B.P. Evidence for continuous morphological variation implies long-term cross-breeding within and between South American camelids (41).

An extremely complex history of interbreeding, even blurring the taxonomy of these species, is indicated by the occurrence of maternal mitochondrial DNA (mtDNA) haplotypes from vicuñas and guanacos in both domesticated llamas and alpacas. Recent mtDNA-based research documents early divergences within the guanaco clade, interpreted as evidence for multiple centers of llama domestication (42). However, the nature of connections among early herders is not well known and these genetic and morphological patterns could, once again, simply reflect recurrent recruitment of individuals from diverse wild populations. Adaptations of wild ancestors to extreme environmental conditions may have contributed to intentional breeding of wild and domestic camelids. Because of the unpredictability of animals surviving extreme weather events and disease, contemporary herders prefer diverse herds, retaining rather than culling individuals with a wide variety of characters (43). In the southern Andes there are records of wild guanacos being tamed and hybridized with llamas (44). Chance breeding of wild and domestic animals also occurs when llamas and alpacas graze unsupervised in the same pastures and most hybrid offspring are fertile (44). Given prolonged interspecific and intraspecific gene flow among Andean camelids, an ancient chimera species is likely.

Low levels of selection and high levels of gene flow among transport animals are also indicated by ethnographic data for yak management on the Tibetan plateau, where limited archaeological data suggest domestication by sheep-herders some 5000–4000 B.P. (45) (Table 1 and Fig. S1). Because wild yaks (*Bos mutus*) are adapted to high-altitude environments (32), human reliance on them for transport and food allowed herders to survive year-round on the high plateau. Genetics show two mtDNA lineages in domestic yaks (45), which are now interpreted in terms of recurrent recruitment of diverse wild yak lineages into

domestic herds (46). Ethnographic data show that breeding of wild and domestic animals is encouraged because domestic yaks are subject to frequent mortality during winter storms. These crosses have strong flight responses but are desired by herders because of their adaptation to the harsh plateau environment, size, and superior ability to protect herds from wolves (45, 47). Wild bulls move to lower elevations to mate with domestic females, where both encouraged and accidental breeding occurs (45, 47). Castration and limited culling are the only forms of directed breeding (47), but environmental selection on herded animals in pastoral camps and landscapes is strong (47).

These cases involving animals from extreme environments, primarily used for transport, all show relatively low levels of directed selection resulting from limited culling and castration, but strong environmental selection within the human niche. The examples also demonstrate practical difficulties for mobile herders of breeding selected animals and maintaining genetic isolation from wild relatives, and the advantages of wild adaptations. Given the demands placed on transport animals and their domestication history, it could be argued that this scenario is unlikely to hold more broadly. However, current evidence suggests that gene flow between domestic and wild populations is not unique to animals used for transport, but may well be true for most other domestic taxa, including animals kept for meat and secondary products, such as milk and wool.

Pigs. Research into the domestication of wild boar provide some of the most comprehensive evidence for out-crossing and gene flow during and after initial domestication, as well as significant variability in these processes across Eurasia (Table 1 and Fig. S1). Wild boar (*Sus* sp.) are social animals, adapted to temperate or subtropical climates. Pigs are multiparous, with rapid gestation and herd growth rates leading to culling at much higher levels than equids, camelids, or bovines, and consequently to higher levels of selection. Unlike animals principally used for transport, intentional interbreeding of pigs with wild relatives confers no productive advantage. Gene flow is most likely to result from wild-capture as a herd-building strategy, or from chance breeding of domestic pigs with wild relatives (Table S1).

Zooarchaeological research indicates a long and complex process, possibly involving two different but related stages: initial commensalism followed by direct human involvement/control and resultant selection (10). Morphometric studies at early Neolithic sites

dating to 9500–8600 B.P. in eastern Anatolia (10) and central China (48) indicate at least two separate domestications of *Sus scrofa*.

Genetic research over the last decade on both ancient and modern *Sus* reveals at least six phylogeographically distinct wild boar lineages have contributed mtDNA to domestic pig populations across the Old World, as well as clear evidence for out-crossing of domestic pigs and wild boar. Evidence also exists for the introduction and dispersal throughout Europe of several Near Eastern mtDNA *S. scrofa* haplotypes with early Neolithic farmers (49). Subsequent recruitment of European wild boar mtDNA lineages into these introduced domesticated swineherds led to the rapid replacement of Near Eastern lineages, first in Europe and then, during the late Bronze Age/Early Iron Age, eastwards across Anatolia (49, 50).

The story of pig domestication in East and Southeast Asia is quite different from that of southwestern Asia and Europe. Here, mtDNA from both ancient and modern *S. scrofa* show that most contemporary Chinese lineages were never incorporated into domestic herds, nor exterminated as a result of hunting or introgression with feral pigs (51), suggesting control (even penning) of pigs from an early stage in the domestication process. Early agriculturalists moving into southeastern Asia deliberately or accidentally recruited local wild boar lineages into their domestic stock, with the result that ancient mainland and island southeastern Asian, New Guinea, and remote Oceanic domestic pigs share their maternal ancestry with lineages recruited from southeastern Asian wild boar populations (49, 52–54), and not with the earliest central Chinese domestic pigs. However, neutral markers, such as mtDNA, can themselves be rapidly replaced during the hybridization process between incoming domestic and local wild stock (53). The nuclear genome retains introgression signatures over longer evolutionary timescales and is now the principal focus for ancient DNA research (53).

These new Eurasian datasets for *S. scrofa* reveal significant introgression and gene flow between wild boar and domestic pig populations after domestication, indicating a rather different domestication process than traditionally purported: one involving initial domestication of a limited number of wild boar from discrete local populations, leading to a degree of genetic isolation. Extensive and mobile swineherding practices, along with subsequent migration/dispersal of early stock-keepers, led to introgression with new local wild boar lineages, which rapidly replaced “founding” lineages.

Historical and modern-day ethnographic observations of traditional pig keeping in, for example, the Mediterranean and Europe, point to the common practice of rather loose and extensive management of domestic pigs, along with long-distance mobility patterns linked with the search for summer and winter feeding (55). Such traditional pig husbandry was likely to have been the norm across Europe millennia earlier than the historical period, and in such circumstances it is likely that out-crossing of domestic pigs with wild boar was common.

Sheep, Goats, and Cattle. Unlike pigs, domestic bovids were widely used for meat, milk, and fiber. Ancient populations of *Capra aegagrus* and *Ovis aries* are the southwestern Asian ancestors of domestic goats and sheep (Table 1 and Fig. S1). Zooarchaeological data document early culling or managed herds of both species by settled hunter-gatherers and early cultivators in eastern Anatolia and the Zagros mountains ca. 11,000–10,000 y ago (56, 57), with goats already displaying morphological changes by ca. 9400–8900 B.P. (11, 58). Compared with pigs, sheep and goat produce only one or two offspring at a time, altering the dynamics of herd management and culling. Traditional pastoralists today manage sheep and goats principally for growth, maximizing females in herds with male-offtake sustained up to 8–16% a year (59). Herders’ decisions regarding males spared for breeding or new stock acquisition (male or female) are informed by family histories of growth potential, color, milk production, and resilience (60–62). Nevertheless, acting primarily on males, directed selection remains weak.

Six wild bezoar lineages found in domestic goats suggest long-term recruitment of wild females to domestic herds (63). Long-distance pastoral movements of flocks through the Zagros provided continual opportunities for unintentional admixture within the natural range of sheep and goats. Morphological change, traditionally associated with domestication, may not have occurred in ancient goats until gene flow was reduced by the dispersal of managed herds outside the range of their wild relatives (58). Any decline in domestic herd size would have provided incentives for wild-capture with periodic weather events, drought, and disease strongly influencing pastoral herd dynamics and viability (59). Similar instability is implied in the case of pigs and goats introduced to Cyprus during the mid-11th millennium B.P. (13, 64). Once secondary products—such as milk or wool—became important, domestic traits, such as productivity and docility, would have

become highly desirable, increasing the influence and intensity of directed selection.

Because of their large size, diverse use, and broad environmental adaptations, relations between humans and cattle differ greatly from those of sheep and goats. Cattle, native to temperate or semiarid subtropical environments, were principally used for meat, and at times depended on heavily for milk, traction, and ceremonial use. *Bos primigenius*, ancestral to taurine cattle, was domesticated in Anatolia 10,500–10,000 B.P. (65–67), whereas *Bos namadicus*, ancestral to zebu cattle, was domesticated in South Asia by ca. 8000–7500 B.P. (68, 69) (Table 1 and Fig. S1). The size of cattle, low growth, and culling rates, as well as early use for milk (70) or traction, implies lower levels of directed selection than even those experienced by pigs or sheep and goat. When selecting herd bulls today, African pastoralists consider similar factors to those discussed for camels, sheep, and goats (59, 71), although cattle are seldom culled at higher than 4–8%. Productive females are not culled, multiple bulls are kept, and natural mortality is often higher than that resulting from culling (72), which results in weak directed selection and strong environmental selection. Slow herd growth promotes gene flow, as does lightly supervised grazing.

The zooarchaeological record indicates a protracted process of domestication of taurine cattle (66) but genetic data suggest small numbers of wild cattle contributed to initial domestication in Anatolia (73), and that diverse wild populations were not incorporated into domestic herds. In contrast to pigs, there is no genetic support for interbreeding of domestic taurine cattle with wild cattle as herders moved across Europe (74), the one exception being data from Italy, where ancient mtDNA suggests female aurochs may have been recruited into domestic herds. The picture is different for South Asia, where high autosomal diversity indicates repeated crossing of domestic zebu cattle with wild males and females (75). Multiple mitochondrial lineages represent either two separate domestications or, again, recruitment of wild animals into domestic zebu herds (68). This variability highlights the roles of regional differences in management practices or herd viability in promoting gene flow. The debate over the question of local domestication of cattle in northeast Africa (76) versus interbreeding of Near Eastern cattle with African wild cattle indicates the extent to which scholars are grappling with the significant role of gene flow in patterning genetic data.

Despite differences in environments, biology, and husbandry practices between taxa,

there is strong evidence for gene flow between pigs, sheep, goat, and cattle and their wild relatives in areas of common distribution. Set against the whole history of domestication, complete separation between wild and domestic populations was relatively late and region-specific. Regional variability in gene flow is demonstrated for pigs and cattle, which took several domestication “pathways” with different degrees of admixture in western, southern, and eastern Eurasia. These patterns of gene flow suggest regionally different approaches to management, with animals closely herded or provisioned in some settings and extensively ranging in others. Variability in herd size and viability was a contributory factor leading to admixture in some—but not all—regions.

Implications of Widespread Gene Flow

Because the role of gene flow in the domestication of large herbivores has, until now, largely been considered minor or peripheral to more dominant processes, drivers of gene flow have not been systematically investigated. Ethnographic and ethnoarchaeological data clearly demonstrate that admixture is not simply an occasional or accidental process. Recent and historic herders intentionally captured wild relatives of their domestic animals and encouraged directed breeding between them. Both herders’ goals and unintended circumstances influenced the extent of gene flow between wild and domestic animals (Table S1). At the same time as discounting gene flow as a significant component of early domestication history, the primacy of strong directional selection in the process has often been assumed (15). It appears that under most historic and prehistoric management regimes, weak directed selection was driven primarily by culling or castration of male surplus to the growth needs of herds. Environmental selection was also a key factor for domestication histories in human-influenced environments.

These findings have significant implications for our interpretation of the archaeological record, determinations of the timing and location of initial domestication, and interpretations of genetic data on domestication. Trends in the extent of directed selection and in gene-flow potentials reinforce many of the distinctions proposed among commensal, prey, and directed pathways to domestication (11, 13), and point to additional selective mechanisms that differentiate them. Culling rates were lower and out-crossing potentials higher for larger transport

animals, horses, donkeys, camelids, and yaks. Correspondingly, higher rates of culling, and therefore of directed selection characterized sheep, goats, and pigs, or more rapidly maturing animals domesticated and managed in less extreme environments.

Interbreeding among domestic, feral, and wild animals, augmented by the opportunities afforded by migrations and trade, has created long and complex evolutionary and domestication histories that challenge assumptions regarding genetic isolation and long-held definitions of domestication. Given differences of degree between domestic and wild animals, some might question whether domestication remains a useful concept. We consider it is essential to treat changing human–animal relations as a continuum, specifying domestication traits that vary with taxon and context—animal–human relationship, place, and time—rather than focusing on general expectations or arbitrary boundaries. This is the direction in which recent archaeological research has been moving (11, 13, 77).

Current assumptions regarding severe domestication bottlenecks and monophyletic origins have complicated attempts by zooarchaeologists and geneticists alike to study the domestication histories of animals such as South American camelids (41), or to interpret coalescence data and estimate domestication time-frames for cats (15). Recurrent gene flow makes wild and domestic animals more similar and the perceived time of divergence more recent. The same assumptions have resulted in widespread (mis-)interpretation of mitochondrial variability in terms of multiple instances of domestication. Recognition of the extent of long-term gene flow within and between wild and domestic animals better reconciles archaeological and genetic data for many species and suggests longer and more complex domestication processes (53). Long-term gene flow also undermines the ability of modern genetic data derived from highly developed modern-day breeds to shed light on the earliest phases of domestication (78).

If gene flow resulting from breeding between wild and domestic animals was common during domestication and has not ceased until recent historic times, it raises many fascinating questions regarding ways in which behavioral and phenotypic domestication traits were maintained, and just what a domestic population was. To address these issues, we need better characterization of animal–human relationships through time, including

better integration of multiple scales of analysis: from the molecular level, to whole animals, to the social contexts and landscapes within which domestication occurs. Diverse zooarchaeological, biochemical, and geoarchaeological approaches to documenting changes in herd sizes, penning, milking and feeding strategies, as well as culling and castration across ancient sites, offer promise for eliciting temporal and site-specific data on selection processes and gene flow. We need to know, for example, exactly where and when out-crossing was common or directed selection high before we can begin to evaluate the respective importance of these processes in the domestication of particular species or to understand regional variability. Other questions, such as the amount of gene flow required to counter directed selection at different levels of culling or natural mortality in human environments, are amenable to modeling (79).

We identify environmental selection under human management as an important force in animal domestication, an area that genomic studies are currently exploring (4) (Table S2). Understanding epigenetic mechanisms, such as patterns of DNA methylation that cause genes to express themselves differently in human compared with wild settings or under varying management regimes (e.g., under stress), promise to provide new insights into ways in which selection was maintained (80, 81). Finally, landscape genetic studies of how small-scale social and biological processes, such as household mobility and exchange or captive animal breeding rates affect movement, interbreeding, and gene flow at large scales, have the potential to integrate anthropological, behavioral, and genetic data (82).

Instead of assuming strong intentional and directional selection during the early stage of animal domestication, the challenge is to investigate sources of selection more critically, bearing in mind the complex interplay of human and environmental selection and the likelihood of long-term gene flow from the wild. These insights on gene flow and unintentional breeding provide new perspectives on early animal domestication, alter current sets of assumptions and questions, and enhance our understanding of domestication as a complex biocultural process.

ACKNOWLEDGMENTS. We thank Ida Nicolaisen and Lior Weissbrod for providing images, and the reviewers for their exceptionally thoughtful comments. This manuscript resulted from a catalysis meeting entitled “Domestication as an Evolutionary Phenomenon: Expanding the Synthesis” that was awarded and hosted by the National Evolutionary Synthesis Centre (National Science Foundation EF-090560) in 2011.

- 1 Trut L, Oskina I, Kharamova A (2009) Animal evolution during domestication: The domesticated fox as a model. *Bioessays* 31(3): 349–360.
- 2 Hemmer H (1990) *Domestication* (Cambridge Univ Press, Cambridge, UK).
- 3 Zeder MA (2011) Pathways to animal domestication. *Harlan II: Biodiversity in Agriculture: Domestication, Evolution and Sustainability*, eds Damania A, Gepts P (Univ California Press, Davis, CA), pp 227–259.
- 4 Andersson LS, et al. (2012) Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. *Nature* 488(7413):642–646.
- 5 Darwin C (1868) *The Variation of Animals and Plants Under Domestication* (John Murray, London).
- 6 Clutton-Brock J (1992) The process of domestication. *Mammal Rev* 22(2):79–85.
- 7 Clutton-Brock J (1999) *A Natural History of Domesticated Animals* (Cambridge Univ Press, Cambridge, UK).
- 8 Bökönyi S (1989) Definitions of animal domestication. *The Walking Larder: Patterns of Domestication, Pastoralism and Predation*, ed Clutton-Brock J (Unwin Hyman, London), pp 22–27.
- 9 Morey D (1994) The early evolution of the domestic dog. *Am Sci* 82(4):336–347.
- 10 Erynyck A, Dobney K, Hongo H, Meadow R (2002) Born free? New evidence for the status of “*Sus scrofa*” at Neolithic Çayönü Tepesi (Southeastern Anatolia, Turkey). *Paleorient* 27(2):47–73.
- 11 Zeder M (2012) The domestication of animals. *J Anthropol Res* 68(2):161–190.
- 12 Price E (2002) *Animal Domestication and Behavior* (CABI Publishing, New York).
- 13 Vigne J-D (2011) The origins of animal domestication and husbandry: A major change in the history of humanity and the biosphere. *C R Biol* 334(3):171–181.
- 14 O’Connor T (2007) Wild or domestic? Biometric variation in the cat *Felis silvestris* Schreber. *Int J Osteoarchaeol* 17:581–595.
- 15 Driscoll CA, Macdonald DW, O’Brien SJ (2009) From wild animals to domestic pets, an evolutionary view of domestication. *Proc Natl Acad Sci USA* 106(Suppl 1):9971–9978.
- 16 Beja-Pereira A, et al. (2004) African origins of the domestic donkey. *Science* 304(5678):1781.
- 17 Kimura B, et al. (2011) Ancient DNA from Nubian and Somali wild ass provides insights into donkey ancestry and domestication. *Proc Biol Sci* 278(1702):50–57.
- 18 Marshall F, Weissbrod L (2011) Domestication processes and morphological change through the lens of the donkey and African pastoralism. *Curr Anthropol* 52(5):S397–S413.
- 19 Rossel S, et al. (2008) Domestication of the donkey: Timing, processes, and indicators. *Proc Natl Acad Sci USA* 105(10):3715–3720.
- 20 Bland-Sutton J (1911) *Man and Beast in Eastern Ethiopia* (McMillan, London), pp 189.
- 21 Nicolaisen J, Nicolaisen I (1997) *The Pastoral Tuareg* (The Carlsberg Foundations Nomad Research Project. Thames and Hudson, New York), p 163.
- 22 Wylie A (2002) *Thinking from Things* (Univ California Press, Berkeley).
- 23 Boessneck J, von den Driesch A, Ziegler R (1989) Die Tierreste von Maadi und Wadi Digla [The animal remains of Maadi and Wadi Digla]. *Maadi III*, eds Rizkana I, Seeher J (Phillipp von Zabern, Mainz), pp 87–128 German.
- 24 Weber J (2008) Elite equids: Redefining equid burials of the mid- to late 3rd millennium BC from Umm el-Marra. *Archaeozoology of the Near East VIII*, eds Vila E, Gourichon L, Choyke MA, Buitenhuis H (Travaux de la Maison de l’Orient et de la Méditerranée, Lyon), pp 499–520.
- 25 Olsen S (2006) Early horse domestication on the Eurasian steppe. *Documenting Domestication: New Genetic and Archaeological Paradigms*, eds Zeder MA, Bradley DG, Emschwiller E, Smith BD (Univ California Press, Berkeley, CA), pp 245–269.
- 26 Outram AK, et al. (2009) The earliest horse harnessing and milking. *Science* 323(5919):1332–1335.
- 27 Vilà C, et al. (2001) Widespread origins of domestic horse lineages. *Science* 291(5503):474–477.
- 28 Warmuth V, et al. (2012) Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proc Natl Acad Sci USA* 109(21):8202–8206.
- 29 Peters J, von den Driesch A (1997) The two-humped camel (*Camelus bactrianus*): New light on its distribution, management and medical treatment in the past. *J Zool (Lond)* 242(4):651–679.
- 30 Ji R, et al. (2009) Monophyletic origin of domestic bactrian camel (*Camelus bactrianus*) and its evolutionary relationship with the extant wild camel (*Camelus bactrianus ferus*). *Anim Genet* 40(4):377–382.
- 31 Trinks A, Burger P, Benecke N, Burger J (2012) Ancient DNA reveals domestication process: The case of the two-humped camel. *Camels in Asia and North Africa: Interdisciplinary Perspectives on Their Past and Present Significance*, eds Knoll E-M, Burger P (Austrian Academy of Sciences Press, Vienna), pp 79–86.
- 32 Schaller GB (1998) *Wildlife of the Tibetan Steppe* (Univ Chicago Press, Chicago).
- 33 Potts DT (2004) Camel hybridization and the role of *Camelus bactrianus* in the ancient Near East. *J Econ Soc Hist Orient* 47(2): 143–165.
- 34 Yadamsuren A, Dulamtsuren E, Reading R (2012) The conservation status and management of wild camels in Mongolia. *Camels in Asia and North Africa: Interdisciplinary Perspectives on Their Past and Present Significance*, eds Knoll E-M, Burger P (Austrian Academy of Sciences Press, Vienna), pp 45–54.
- 35 Al-Swailem AM, et al. (2010) Sequencing, analysis, and annotation of expressed sequence tags for *Camelus dromedarius*. *PLoS ONE* 5(5):e10720.
- 36 Peters J (1997) [The dromedary: Ancestry, history of domestication and medical treatment in early historic times]. *Tierarztl Prax Ausg G Grosstiere Nutztiere* 25(6):559–565.
- 37 Uerpmann H-P, Uerpmann M (2002) The appearance of the domestic camel in south-east Arabia. *J Oman Stud* 12:235–260.
- 38 Köhler-Rollefson I, Rathore HS (2004) Indigenous versus official knowledge, concepts and institutions: Raika pastoralists and the outside world. *Nomad People* 8(2):150–167.
- 39 Mengoni-Gohalons GL (2008) Camels in ancient Andean societies: A review of the zooarchaeological evidence. *Quat Int* 185(1):59–68.
- 40 Mengoni-Gohalons G, Yacobaccio HD (2006) The domestication of South American camelids: A view from the south-central Andes. *Documenting Domestication: New Genetic and Archaeological Paradigms*, eds Zeder M, Bradley DG, Emschwiller E, Smith BD (Univ California Press, Berkeley), pp 228–244.
- 41 Moore KM (2011) Grace under pressure: Responses to changing environments by herders and fishers in the Formative Lake Titicaca Basin, Bolivia. *Sustainable Livelihoods: Cultural Persistence in an Ever-Changing Environment*, eds Miller NF, Moore KM, Ryan K (Univ Pennsylvania Museum of Archaeology and Anthropology, Philadelphia), pp 244–272.
- 42 Barreta J, et al. (2013) Analysis of mitochondrial DNA in Bolivian llama, alpaca and vicuna populations: A contribution to the phylogeny of the South American camelids. *Anim Genet* 44(2): 158–168.
- 43 Bolton M (2006) Genetic defects or generative prototypes? Competing models for livestock improvement in southern Bolivia. *J R Anthropol Inst* 12(3):531–549.
- 44 Bonavia D (2008) *The South American Camelids* (Cotsen Institute of Archaeology, Univ of California, Los Angeles).
- 45 Rhode D, Madsen DB, Brantingham PJ, Dargye T (2007) Yaks, yak dung, and prehistoric human habitation of the Tibetan Plateau. *Developments in Quaternary Sciences* 9:205–224.
- 46 Wang Z, et al. (2010) Phylogeographical analyses of domestic and wild yaks based on mitochondrial DNA: New data and reappraisal. *J Biogeogr* 37(12):2332–2344.
- 47 Wiener G, Han J, Long R (2003) *The Yak* (FAO, Bangkok) 2nd. Ed.
- 48 Cucchi T, et al. (2011) Early Neolithic pig domestication at Jiahu, Henan Province, China: Clues from molar shape analyses using geometric morphometric approaches. *J Archaeol Sci* 38(1):11–22.
- 49 Larson G, et al. (2007) Ancient DNA, pig domestication, and the spread of the Neolithic into Europe. *Proc Natl Acad Sci USA* 104(39):15276–15281.
- 50 Ottoni C, et al. (2013) Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. *Mol Biol Evol* 30(4):824–832.
- 51 Larson G, et al. (2010) Patterns of East Asian pig domestication, migration, and turnover revealed by modern and ancient DNA. *Proc Natl Acad Sci USA* 107(17):7686–7691.
- 52 Larson G, Cucchi T, Dobney K (2011) Genetic aspects of pig domestication. *The Genetics of the Pig*, eds Rothschild MF, Ruvinsky A (CAB International, Wallingford), pp 14–37.
- 53 Larson G, Burger J (2013) A population genetics view of animal domestication. *Trends Genet* 29(4):197–205.
- 54 Dobney K, Cucchi T, Larson G (2008) The pigs of Island Southeast Asia and the Pacific: New evidence for taxonomic status and human-mediated dispersal. *Asian Perspect* 47(1):59–74.
- 55 Albarella U, Manconi F, Trentacoste A (2011) A week on the plateau: Pig husbandry, mobility and resource exploitation in Central Sardinia. *Ethnozoarchaeology: The Present and Past of Human-animal Relationships*, eds Albarella U, Trentacoste A (Oxbow Books, Oxford), pp 143–159.
- 56 Hiendler S, Kaup B, Wassmuth R, Janke A (2002) Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc Biol Sci* 269(1494):893–904.
- 57 Peters J, von den Driesch A, Helmer D (2005) The upper Euphrates-Tigris Basin: Cradle of agro-pastoralism? *The First Steps of Animal Domestication. New Archaeological Approaches*, eds Vigne JD, Peters J, Helmer D (Oxbow Books, Oxford), pp 96–124.
- 58 Zeder MA (2008) Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc Natl Acad Sci USA* 105(33):11597–11604.
- 59 Dahl G, Hjort A (1976) *Having Herds: Pastoral Herd Growth and Household Economy* (Department of Social Anthropology, Univ Stockholm, Stockholm).
- 60 Kebede T, Haile A, Dadi H (2012) Smallholder goat breeding and flock management practices in the central rift valley of Ethiopia. *Trop Anim Health Prod* 44(5):999–1006.
- 61 Tabbaa M, Al-Atiyat R (2009) Breeding objectives, selection criteria and factors influencing them for goat breeds in Jordan. *Small Rumin Res* 84(1):8–15.
- 62 Mbuku SM, Kosgey IS, Kahi AK (2010) Identification systems and selection criteria for small ruminants among pastoralist communities in northern Kenya: Prospects for a breeding programme. *Trop Anim Health Prod* 42(7):1487–1492.
- 63 Naderi S, et al. (2007) Large scale mitochondrial DNA analysis of the domestic goat reveals six maternal lineages with high haplotype diversity. *PLoS ONE* 2(10):e1012.
- 64 Vigne J-D (2013) Domestication Process and Domestic Ungulates: New Observations from Cyprus. *The Origins and Spread of Domestic Animals in Southwest Asia and Europe*, eds Colledge S, Conolly J, Dobney K, Manning K, Shennan S (Left Coast Press, Walnut Creek), pp 115–128.
- 65 Hanotte O, et al. (2002) African pastoralism: Genetic imprints of origins and migrations. *Science* 296(5566):336–339.
- 66 Hongo H, et al. (2009) The process of ungulate domestication at Çayönü, Southeastern Turkey: A multidisciplinary approach focusing on *Bos* sp. and *Cervus elaphus*. *Anthropozoologica* 44(1):63–73.
- 67 Loftus RT, MacHugh DE, Bradley DG, Sharp PM, Cunningham P (1994) Evidence for two independent domestications of cattle. *Proc Natl Acad Sci USA* 91(7):2757–2761.
- 68 Chen S, et al. (2010) Zebu cattle are an exclusive legacy of the South Asia neolithic. *Mol Biol Evol* 27(1):1–6.
- 69 Meadow RH (1996) The origins and spread of agriculture and pastoralism in northwestern South Asia. *The Origins and Spread of Agriculture and Pastoralism in Eurasia*, ed Harris DR (Univ College London Press, London), pp 390–412.
- 70 Evershed RP, et al. (2008) Earliest date for milk use in the Near East and southeastern Europe linked to cattle herding. *Nature* 455(7212):528–531.
- 71 Kugonza DR, Nabasiye M, Hanotte O, Mpairwe D, Okeyo AM (2012) Pastoralists’ indigenous selection criteria and other breeding practices of the long-horned Ankole cattle in Uganda. *Trop Anim Health Prod* 44(3):557–565.
- 72 Mutundu K (2005) Domestic stock age profiles and herd management practices: Ethnoarchaeological implications from Maasai settlements in southern Kenya. *Archaeofauna* 14:83–92.
- 73 Bollongino R, et al. (2012) Modern taurine cattle descended from small number of near-eastern founders. *Mol Biol Evol* 29(9): 2101–2104.
- 74 Edwards CJ, et al. (2007) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc Biol Sci* 274(1616):1377–1385.
- 75 Murray C, Huerta-Sanchez E, Casey F, Bradley DG (2010) Cattle demographic history modelled from autosomal sequence variation. *Philos Trans R Soc Lond B Biol Sci* 365(1552):2531–2539.
- 76 Stock F, Gifford-Gonzalez DP (2013) Genetics and African cattle domestication. *Af Arch Rev* 30(1):51–72.
- 77 Denham T, Iriarte J (2007) *Rethinking Agriculture: Archaeological and Ethnoarchaeological Perspectives*, ed Vrydachs L (Left Coast Press, Walnut Creek).
- 78 Larson G, et al. (2012) Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA* 109(23):8878–8883.
- 79 Gerbault P, et al. (2014) Storytelling and story testing in domestication. *Proc Natl Acad Sci USA* 111:6159–6164.
- 80 Jablonka E (2013) Behavioral epigenetics in ecological context. *Behav Ecol* 24:325–326.
- 81 Nätt D, et al. (2012) Heritable genome-wide variation of gene expression and promoter methylation between wild and domesticated chickens. *BMC Genomics* 13:59.
- 82 Segelbacher G, et al. (2010) Application of landscape genetics in conservation biology: Concepts and challenges. *Conserv Genet* 11(2):375–385.

Storytelling and story testing in domestication

Pascale Gerbault^a, Robin G. Allaby^b, Nicole Boivin^c, Anna Rudzinski^a, Ilaria M. Grimaldi^c, J. Chris Pires^d, Cynthia Climer Vigueira^e, Keith Dobney^f, Kristen J. Gremillion^g, Loukas Barton^h, Manuel Arroyo-Kalinⁱ, Michael D. Purugganan^{j,k}, Rafael Rubio de Casas^l, Ruth Bollongino^m, Joachim Burger^m, Dorian Q. Fuller^j, Daniel G. Bradleyⁿ, David J. Balding^o, Peter J. Richerson^p, M. Thomas P. Gilbert^q, Greger Larson^r, and Mark G. Thomas^{a,1}

^aResearch Department of Genetics, Evolution, and Environment and ^oUniversity College London Genetics Institute, University College London, London WC1E 6BT, United Kingdom; ^bSchool of Life Sciences, Gibbet Hill Campus, University of Warwick, Coventry CV4 7AL, United Kingdom; ^cResearch Laboratory for Archaeology and the History of Art, School of Archaeology, Oxford OX1 3QY, United Kingdom; ^dDivision of Biological Sciences, University of Missouri, Columbia, MO 65211; ^eDepartment of Biology, Washington University in St. Louis, St. Louis, MO 63130; ^fDepartment of Archaeology, University of Aberdeen, Aberdeen AB24 3UF, United Kingdom; ^gDepartment of Anthropology, Ohio State University, Columbus, OH 43210; ^hDepartment of Anthropology, Center for Comparative Archaeology, University of Pittsburgh, Pittsburgh, PA 15260; ⁱInstitute of Archaeology, University College London, London WC1H 0PY, United Kingdom; ^jDepartment of Biology, New York University, New York, NY 10003-6688; ^kCenter for Genomics and Systems Biology, New York University Abu Dhabi Research Institute, Abu Dhabi, United Arab Emirates; ^lDepartamento de Ecología, Universidad de Granada, E-18071 Granada, Spain; ^mInstitute of Anthropology, Johannes Gutenberg University, D-55099 Mainz, Germany; ⁿSmurfit Institute of Genetics, Trinity College, Dublin 2, Ireland; ^oDepartment of Environmental Science and Policy, University of California, Davis, CA 95616; ^pCentre for GeoGenetics, Natural History Museum of Denmark, University of Copenhagen, 1350 Copenhagen, Denmark; and ^qDurham Evolution and Ancient DNA, Department of Archaeology, Durham University, Durham DH1 3LE, United Kingdom

Edited by Andrea Manica, Cambridge University, Cambridge, United Kingdom, and accepted by the Editorial Board March 18, 2014 (received for review January 30, 2014)

The domestication of plants and animals marks one of the most significant transitions in human, and indeed global, history. Traditionally, study of the domestication process was the exclusive domain of archaeologists and agricultural scientists; today it is an increasingly multidisciplinary enterprise that has come to involve the skills of evolutionary biologists and geneticists. Although the application of new information sources and methodologies has dramatically transformed our ability to study and understand domestication, it has also generated increasingly large and complex datasets, the interpretation of which is not straightforward. In particular, challenges of equifinality, evolutionary variance, and emergence of unexpected or counter-intuitive patterns all face researchers attempting to infer past processes directly from patterns in data. We argue that explicit modeling approaches, drawing upon emerging methodologies in statistics and population genetics, provide a powerful means of addressing these limitations. Modeling also offers an approach to analyzing datasets that avoids conclusions steered by implicit biases, and makes possible the formal integration of different data types. Here we outline some of the modeling approaches most relevant to current problems in domestication research, and demonstrate the ways in which simulation modeling is beginning to reshape our understanding of the domestication process.

model | inference | evolution | agriculture | Neolithic

The emergence of agriculture beginning some 10,000 y ago marked more than a change in human patterns of subsistence. The beginnings of food production ushered in an era of radically new relationships between humans and other species, dramatic new evolutionary pressures, and fundamental transformations to the earth's biosphere. The evolutionary process of plant and animal domestication by humans led to morphological, physiological, behavioral, and genetic differentiation of a wide range of species from their wild progenitors (1, 2). The selection pressures that were placed on such species continue today, sometimes through direct genetic modification, and both the processes and their outcomes are accordingly of significant broader interest. Domestication is also part of a cultural evolutionary process (3, 4), and some human genes have evolved in response to cultural innovations (5–8), much as the genes of domesticated species have changed under the impact of human artificial selection. The study of domestication today is a multidisciplinary enterprise in which archaeologists and agricultural scientists have been joined by evolutionary biologists and population geneticists (2, 9).

At least five major sets of questions tend to reoccur in the domestication literature. The first three are demographic: (i) When, where, and in how many geographic locations was a given species domesticated? (ii) What were the dispersal routes from the original domestication centers? (iii) To what extent did hybridization between domesticates and local wild relatives occur? The remaining questions relate to adaptation: (iv) To what extent, and how rapidly, were domestic traits fixed? (v) How well did domesticates adapt to diverse anthropogenic environments?

Most of these questions can be at least partially addressed using population genetic data from both ancient and modern samples. This is because variation across the genome is shaped by—and thus reflects—past demography, whereas genetic variation in and around particular genes determining key phenotypic traits is shaped by adaptation history. These principles, in combination with the availability of increasing quantities of ancient and modern genetic data, have led to a profusion of studies on particular domestication scenarios (e.g., refs. 10 and 11). However, the relationship between genetic data and the demographic or adaptation history that shaped it is noisy and often difficult to predict. This difficulty is primarily because: (i) in any evolving system that includes stochastic processes, patterns in genetic (or archaeological) data could have been generated under a range of

Significance

Our knowledge of the domestication of animal and plant species comes from a diverse range of disciplines, and interpretation of patterns in data from these disciplines has been the dominant paradigm in domestication research. However, such interpretations are easily steered by subjective biases that typically fail to account for the inherent randomness of evolutionary processes, and which can be blind to emergent patterns in data. The testing of explicit models using computer simulations, and the availability of powerful statistical techniques to fit models to observed data, provide a scientifically robust means of addressing these problems. Here we outline the principles and argue for the merits of such approaches in the context of domestication-related questions.

Author contributions: P.G., R.G.A., P.J.R., M.T.P.G., G.L., and M.G.T. designed research; and P.G., R.G.A., N.B., A.R., I.M.G., J.C.P., C.C.V., K.D., K.J.G., L.B., M.A.-K., M.D.P., R.R.d.C., R.B., J.B., D.Q.F., D.G.B., D.J.B., P.J.R., M.T.P.G., G.L., and M.G.T. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. A.M. is a guest editor invited by the Editorial Board.

¹To whom correspondence should be addressed. E-mail: m.thomas@ucl.ac.uk.

different histories (equifinality); (ii) any particular history can potentially give rise to a wide range of different patterns in data [evolutionary variance (12)]; and (iii) certain demographic or adaptation histories can give rise to counter intuitive patterns in data (emergence) (e.g., refs.13–16).

Evolutionary histories are rarely directly “revealed” by looking only at patterns in data, such as the distribution of particular markers (e.g., morphological traits, material culture, genetic variants). This is because such data may be only weakly constrained by those histories; many different histories may explain the same data equally well. Thus, instead of simply providing narratives based on interpretations, implicit assumptions, and preconceived ideas, domestication histories need to be tested to identify those scenarios that best explain observed data; to do this, domestication histories must be modeled explicitly. We outline a range of modeling techniques that can be used in domestication research and provide examples that illustrate their utility. Although most of the discussion and examples given in this paper are based on population genetic data, most of the principles and approaches can also be applied to other datasets used to explore domestication processes.

Types of Modeling Approaches

A model is an explicit and simplified representation of the underlying causative mechanisms in a system and is used to make predictions about the observed outcomes (data) of that system. We consider two classes of models: discriminative models, which fit directly observed data to predicted relationships (e.g., linear regression), and generative models, which are intended to capture the main real-world mechanisms that generate data, and are typically used to produce artificial datasets. Discriminative models make assumptions, sometimes but not always explicitly, on the ways aspects of the data are correlated without specifying the actual mechanisms that generate those correlations (e.g., refs. 17–21). Generative models aim to explicitly replicate key hypothesized (i.e., assumed) processes that generate the data. Because all evolutionary processes include stochastic elements, a range of different outcomes—or patterns in empirical data—can be generated from any particular scenario or model. For this reason, when using generative models, it is often necessary to produce many datasets by simulation.

In population genetics, a powerful means of simulating data is the “retrospective” approach of coalescent simulation (22), where the joining (or coalescence) of lineages is simulated backward in time under specific assumptions about such variables as population size, structure, migration, and admixture. This approach is highly efficient because it only simulates the lineage history of the sample, not of the whole population, so simulation can be very fast. However, coalescent approaches are limited in the demographic and selection scenarios that can be modeled, and some researchers instead favor a more flexible, but computationally demanding forward-in-time simulation (e.g., ref. 23), or simulations with a combination of forward- and backward-in-time elements (e.g., ref. 24).

Generative models can be agent-based, whereby agents with prescribed interaction behaviors are simulated as individual units. However, agent-based models tend to be computationally demanding, so their application is usually restricted to revealing some emergent, sometimes counter-intuitive, properties of a modeled system (e.g., ref. 25) rather than making inferences by fitting to existing data (e.g., ref. 26). An additional level of resolution in evolutionary modeling can be achieved through spatially explicit simulation, sometimes reducing continuous space to a number of cells (or “demes”) with defined neighbor relationships. These spatial refinements can be computationally challenging, particularly when geographic features (e.g., elevation, climate) or population dynamics (e.g., varying carrying capacities) are introduced.

A shared characteristic of these modeling approaches is that they are made up of a number of components reflecting the real-world processes that are hypothesized to have shaped the data. These components can be explicitly modified and combined with one another.

Models and Data

To be useful in evolutionary inference, models should be fitted to observed data. Often the most important aspect of model-fitting is deciding how to deal with unknown parameters, such as migration rates or selection coefficients.

Frequentist approaches to hypothesis testing or estimation treat the unknown parameters as fixed; the model specifies imaginary random repetitions of the data generation process (e.g., refs. 13, 25, 27, 28). There is therefore no probability distribution for the parameters, but instead statements are made about the frequency of future datasets satisfying certain conditions given assumed parameter values. Often the data are reduced to summary statistics (e.g., means, variances) intended to capture the most important information about the processes modeled. One of the simplest forms of inference is to consider the distribution of a summary statistic under a given model in comparison with the observed value of that statistic. This procedure can be used to reject models or parameter values as implausible, but is not useful for more quantitative comparisons.

Given some assumed—or known—model of the processes at play, the parameter values that maximize the probability of observing the data can be obtained (maximum likelihood). The main requirement of maximum-likelihood approaches is a likelihood function: a mathematical formula that specifies the probability of the data as a function of the parameter values. This function can be used in a frequentist setting, but is more commonly used directly to identify ranges of plausible parameter values (e.g., refs. 29, 30). Likelihood-based (and full-Bayesian, see below) approaches usually use the full information content of data and not just some aspect of it, such as summary statistics. However, (i) the likelihood function can be difficult to formulate for anything but the simplest models, (ii) if there are many parameters, maximizing the likelihood can be computationally demanding, and (iii) there can be multiple maxima of the likelihood function.

Full Bayesian methods also make use of the likelihood function but they allow the incorporation of “prior information” about the model parameters, which can help to focus on the most plausible regions of parameter space. Computational techniques, such as Markov chain Monte Carlo (MCMC), have made Bayesian methods more tractable and more popular. A wide variety of MCMC techniques exist. They are all “samplers” because they sample parameter values at random from their prior or posterior (i.e., target) distributions. More specifically, the aim of MCMC techniques in Bayesian inference is to use prior probabilities and the likelihood function to condition a random walk through parameter space. This process results in the distribution of parameter values “visited” numerically approximating the posterior distribution (i.e., the updated knowledge of the parameters given the data). However, for large parameter spaces MCMC can still be computationally very expensive (31, 32).

Because likelihood functions are only workable for relatively simple models, there can be a tension between fitting more elegant and powerful statistical methods assuming simplistic models, and assuming more general models that only permit crude modes of inference. For example, in population genetics a likelihood-based method may only be available for a single population model (e.g., ref. 33), and researchers may accept this limitation uncritically even when the data are clearly from multiple populations, or subsets of the data are selected to fit the model [e.g., selecting DNA sequences from only one branch of a phylogeny (34, 35)]. This uncritical acceptance will lead to

misleading inferences unless the data-selection step is incorporated in the model specification.

The difficulty of computing likelihoods for all but the simplest models has led to the development of a family of techniques known as approximate Bayesian computation (ABC) (36, 37). In its simplest form, ABC works by simulating data from a generative model with parameter values chosen at random from their prior distributions. A simulation is accepted if the simulated data resemble the observed data, and rejected otherwise, where the “resemblance” of datasets is measured using one or more summary statistics. The proposed parameter values that are accepted in this algorithm form a sample from an approximation to the posterior distribution. Thus, the approach is very similar to MCMC, with the exception that the latter samples are from the true posterior distribution. Several variants of this process have been developed to improve accuracy and computational efficiency (38–40). ABC provides a framework for estimating parameters of interest and comparing relative support for different models based on the same data (37, 41).

The big advantage of ABC is modeling flexibility because almost any generative model can be used, but this comes at the cost of only approximate answers, because ABC does not use all of the information in the data. In addition, the accuracy of the resulting approximation is hard to assess and choosing appropriate summary statistics can be difficult (42, 43). Although the development of full-likelihood Monte Carlo methods, particularly those addressing issues of statistical intractability, continues apace (e.g., ref. 44, 45), ABC provides a useful adjunct to these approaches (e.g., ref. 46), permitting currently intractable problems to be side-stepped, even if only temporarily. In addition, methodologies such as ABC allow for the integration of distinct sources of data (47). This integration will become increasingly important as new types of data (e.g., paleoclimatic, archaeological, genetic) accumulate and demand statistically informed comparison and integration. Given the many factors that are important in shaping data patterns in domesticates, ABC provides the most promising means of democratizing simulation modeling for the domestication research community.

Modeling Domestication History

Explicit modeling-based studies of domestication are relatively new and mostly confined to inference from population genetic data, but have nonetheless begun to transform our understanding of the five major domestication questions outlined above. Here we highlight some examples in which initial inference about domestication processes based on direct interpretation of patterns in data were later demonstrated problematic when tested using modeling approaches.

When, Where, and in How Many Geographic Locations Was a Given Species Domesticated? A common but questionable interpretation of lineage divergence date estimates for Y chromosome or mitochondrial DNA (mtDNA) data are that they represent founding events in species or populations. However, the choice of which lineages to estimate divergence dates for can be arbitrary, and there is little reason to expect demographic processes, such as domestication, to correlate with lineage ages, unless those founder events involved very small numbers of individuals; population genetic models show that lineage coalescent dates can predate major demographic episodes. This finding is well-illustrated with domestic dogs; an early estimate of 135,000 y for the coalescence age of the major mtDNA lineage (clade I, see ref. 48) was interpreted as indicating a domestication founding event around that time. More recently, modeling approaches based on diffusion approximations (45) and the generalized phylogenetic coalescent sampler (49), both conditioned on whole-genome sequence data, estimated domestic dog-wolf divergence between 32,000 y ago (50) and 11,000–16,000 y ago (51). Although these model-based

date estimates differ (most probably because of the assumed evolutionary rate), they have concordance with those from fossil canids currently considered morphologically more similar to dogs than wolves (e.g., refs. 52 and 53).

Goat domestication has also been reevaluated. mtDNA sequences in domesticated goat have been assigned to five major haplogroups (54), the first three of which have expansion age estimates in the range of 10,000–841 y ago, based on DNA sequence mismatch distributions (55). The coalescent date estimates between these haplogroups are considerably older (103,000 and 597,800 y ago) (54). Initially, these haplogroups were interpreted as representing independent domestication events (34, 55), and the overall patterns of mtDNA divergence as only being consistent with an implausibly high number of initial domestications [38,000–82,000 females (55)]. However, application of coalescent simulation and ABC fitting to published ancient and modern mtDNA data indicated that these data could be equally well explained by a single domestication episode of smaller size, or successive founding events as domestic goat populations expanded into Europe (56).

The extraordinary phenotypic range of the common bean (*Phaseolus vulgaris*) has made it a particularly interesting target for domestication research (57). A range of genetic studies (e.g., ref. 58) indicates two highly diverged gene pools, one hypothesized to originate in Mesoamerica and the other in the Andes. Within the Mesoamerican gene pool, random amplified polymorphic DNA (59) and chloroplast data (60) have been interpreted as indicating independent domestication events. However, using coalescent simulation and ABC, Mamidi et al. (57) showed that single domestication episodes for both the Mesoamerican and Andean gene pools, with strong bidirectional gene flow between domesticated and wild species, provided the best fit to data on 13 loci.

Modeling approaches have also altered our views of the domestication of rice (30). Previous genetic studies had inferred that rice was domesticated twice, in China and in India, giving rise to the *japonica* and *indica* cultivars, respectively (61). When demographic modeling using a diffusion approximation-based approach (45) was applied to SNP data from three rice chromosomes, however, only one domestication was indicated. In conjunction with archaeological data, a more nuanced view of rice domestication has emerged, suggesting that *japonica* was domesticated in China, and that *indica* arose possibly as a result of subsequent introgression of *japonica* into wild rice or proto-*indica* populations in India (62).

What Were the Dispersal Routes from the Original Domestication Centers? Traditionally, investigating the geographical origin and subsequent dispersal pathways of a domesticate rely heavily on identifying the genetically closest wild progenitor populations. However, shifts through time in the location of such ancestral populations, extinctions, and undersampling can all weaken such an approach. Although application of modeling techniques to the delineation of pathways of dispersal is in its infancy, preliminary geospatial modeling has been conducted to infer the dispersal of maize in the Americas (20). This approach accounted for landscape, radiocarbon dates of crop remains, and genetic diversity. The model-fitting was performed by multiple-criteria regression analysis over archaeobotanical and genetic data. This allowed tensions between archaeological and genetic data to be explicitly modeled and explored, and points the way forward for the more systematic and statistically informed examination of dispersal pathways, particularly in domesticates like rice, for which the geography of both domestication and subsequent dispersal remains a source of significant debate (63, 64).

To What Extent Did Hybridization Between Domesticates and Local Wild Relatives Occur? Gene flow from wild populations can have important effects on patterns of genetic variation. Low Y chromosomal diversity in modern horses has been interpreted as the result of a single geographically restricted area of domestication (65), whereas the high diversity and low phylogeographic structure in mtDNA has been interpreted as support for multiple origins of domesticated horses (66, 67). Using a spatially explicit forward simulation model, conditioned on autosomal genotype data using ABC (68), the Western Eurasian steppe has been identified as the most likely origin for modern horses, with a model of repeated introgression from local wild to domesticated horses offering the best fit to observed patterns of diversity.

Genetic interactions can be particularly complex when introgression between domesticates and several related wild taxa is possible. The cultivated apple (*Malus domestica*), for example, has been proposed to derive from multiple wild relatives: most notably *Malus sieversii* and *Malus sylvestris*, with potential contribution of other taxa, such as *Malus orientalis*. A recent study used demographic modeling with ABC to compare the introgression scenarios between the different taxa (69). Results supported *M. sieversii* as the primary source of the domesticated apple, but also with frequent and widespread introgression from *M. sylvestris*, potentially contributing characters relevant for the adaptation to novel environments and human use. In this case modeling was essential to disentangle the biological complexity of the domestication process, which involved at least two different species and repeated hybridization events over a long period (69).

To What Extent, and How Rapidly, Were Domestic Traits Fixed? Monophyletic patterns in phylogenetic analyses of genome-wide markers, such as amplified fragment-length polymorphisms, have been interpreted as indicating a rapid fixation of domestication traits and spread of domesticates from one center (11, 70). Rapid fixation once seemed to find confirmation in experimental field studies (71). Increasingly, however, the archaeological record reveals that some traits require centuries or millennia to reach fixation (19, 72), as would be expected if those traits were determined by dominant advantageous alleles. These monophyletic patterns were investigated using an individual-based modeling approach conditioned on amplified fragment-length polymorphism data from crops (13). Researchers found that both multiple and single origins of domestic crops could explain the data, and that the multiple-origin model produced the monophyletic signal more rapidly. Simulation also demonstrated that a monophyletic signal alone need not indicate rapid fixation.

How Well Did Domesticates Adapt to Diverse Anthropogenic Environments? An additional area of interest in domestication studies has been how outbreeding plant systems have adapted to the human environment in situations when there is continuous gene-flow between wild and domesticated populations (73). Le Thierry d'Ennequin et al. (74) used an individual-based model to demonstrate the increased number of genes likely to be under selection with linkage because of genome architecture and mating strategy. Thus, some genome architectures may be more adaptable to the domesticated environment than others. Artificial and natural selection can pull traits in opposite directions [e.g., seed size (75)], resulting in weak selection in net effect. Such adaptation complexity is consistent with the large number of genes (i.e., 27–70 genes) thought to underlie domestication traits in wheat, maize, and sunflower (76–79). Furthermore, because the number of genes that can be under selection simultaneously is constrained (80), it is necessary to consider gene interactions in models of adaptation (75, 81). A case in point is the adaptation of crops to

higher latitudes as they were moved out of their location(s) of origin (82–84). A reduction in crop adaptability may have led to population collapse in mid-Holocene Europe (85), resulting into regional agricultural abandonment (86).

In addition to looking at the adaptation of domesticates to anthropogenic environments, researchers interested in the cultural impact of the domestication process can also use models to evaluate human response to culturally determined selection pressures (87). The dietary changes, population growth, increased sedentism, and new diseases that accompanied the domestication of plants and animals appear to have triggered a wave of genetically based adaptations in our own immune and digestive systems (5, 88). The rapid increase in the availability of human gene-sequence data are making model-driven data analysis increasingly feasible and attractive in studies of human cultural and genetic response to agriculture-associated innovations (e.g., ref. 47).

Concluding Remarks

Model-based statistical approaches are an essential tool in domestication research. When inferring past processes, explicit models are particularly important as, typically, data are the result of a single experiment (the past), and it is necessary to explore a landscape of hypotheses to test which could have given rise to those observed data. Because the hypothesis landscape is effectively infinite, it will always be the case that some unjustifiably complex model can be found to explain the data well (overfitting), so it is necessary to make explicit assumptions, to consider simple models first, and to penalize complexity (89, 90).

These requirements are often seen by nonstatisticians as key drawbacks of modeling approaches. However, assumptions are always present when inferring past processes, and making them explicit enables their recognition and evaluation. In addition, the advantages of using simplified explicit models—particularly statistical tractability and the avoidance of overfitting—outweigh their drawbacks. Furthermore, simple does not mean easy. To quote George Box (91): “Just as the ability to devise simple but evocative models is the signature of the great scientist so overelaboration and overparameterization is often the mark of mediocrity.”

Because of the central role clear hypothesis formulation and testing play in scientific research, we suggest that the arguments presented here apply not only to the field of domestication research or population genetics, but to any discipline involving historical inference (85, 92). Modeling is not the only way to proceed and does not guarantee the right answers [indeed, “models are always wrong, and sometimes useful” (91)]. Likewise, interpretative approaches can be valuable in the scientific process and may lead to the correct, or nearly correct, explanation. However, interpretative inference is better thought of as a means of generating hypotheses (storytelling), whereas explicit models permit those hypotheses (or stories) to be tested. With advances in statistical modeling techniques and increases in computer power, the approaches discussed in this article are set to transform our understanding of domestication processes.

ACKNOWLEDGMENTS. This manuscript resulted from a catalysis meeting entitled “Domestication as an Evolutionary Phenomenon: Expanding the Synthesis” that was awarded and hosted by the National Evolutionary Synthesis Centre (National Science Foundation EF-0905606) in 2011. P.G. is funded by Leverhulme Trust; D.G.B. is funded by Science Foundation Ireland (09/IN.1/B2642); A.R. and R.R.d.C. are funded by European Union funding (PITN-GA-2011- 289966 “BEAN,” MC-IF-2011-300026 “TEE-OFF”); M.T.P.G. is funded by the Danish Council for Independent Research Grant 10-081390; and I.M.G. is funded by the European Research Council as part of Grant Agreement 206148 “SEALINKS” (to N.B.).

1. Darwin C (1868) *The Variation of Animals and Plants Under Domestication* (John Murray, London), Vol. I.
2. Larson G, Burger J (2013) A population genetics view of animal domestication. *Trends Genet* 29(4):197–205.
3. Cavalli-Sforza LL, Feldman MW (1981) *Cultural Transmission and Evolution: A Quantitative Approach* (Princeton Univ Press, Princeton, NJ).
4. Odling-Smee FJ, Laland KN, Feldman MW (2003) *Niche Construction: The Neglected Process in Evolution* (Princeton Univ Press, Princeton).
5. Gerbault P, et al. (2011) Evolution of lactase persistence: An example of human niche construction. *Philos Trans R Soc Lond B Biol Sci* 366(1566):863–877.
6. Laland KN, Odling-Smee J, Myles S (2010) How culture shaped the human genome: Bringing genetics and the human sciences together. *Nat Rev Genet* 11(2):137–148.
7. Leach HM (2003) Human domestication reconsidered. *Curr Anthropol* 44(3):349–368.
8. Richerson PJ, Boyd R, Henrich J (2010) Colloquium paper: Gene-culture coevolution in the age of genomics. *Proc Natl Acad Sci USA* 107(Suppl 2):8985–8992.
9. Zeder MA (2006) *Documenting Domestication: New Genetic and Archaeological Paradigms* (Univ of California Press, Berkeley, CA).
10. Edwards CJ, et al. (2007) Mitochondrial DNA analysis shows a Near Eastern Neolithic origin for domestic cattle and no indication of domestication of European aurochs. *Proc Biol Sci* 274(1616):1377–1385.
11. Heun M, et al. (1997) Site of Einkorn wheat domestication identified by DNA fingerprinting. *Science* 278(5341):1312–1314.
12. Wilson IJ, Balding DJ (1998) Genealogical inference from microsatellite data. *Genetics* 150(1):499–510.
13. Allaby RG, Fuller DQ, Brown TA (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proc Natl Acad Sci USA* 105(37):13982–13986.
14. Edmonds CA, Lillie AS, Cavalli-Sforza LL (2004) Mutations arising in the wave front of an expanding population. *Proc Natl Acad Sci USA* 101(4):975–979.
15. Excoffier L, Foll M, Petit RJ (2009) Genetic consequences of range expansions. *Annu Rev Ecol Syst* 40:481–501.
16. François O, et al. (2010) Principal component analysis under population genetic models of range expansion and admixture. *Mol Biol Evol* 27(6):1257–1268.
17. Fuller DQ, et al. (2009) The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323(5921):1607–1610.
18. Novembre J, Stephens M (2008) Interpreting principal component analyses of spatial population genetic variation. *Nat Genet* 40(5):646–649.
19. Purugganan MD, Fuller DQ (2011) Archaeological data reveal slow rates of evolution during plant domestication. *Evolution* 65(1):171–183.
20. van Etten J, Hijmans RJ (2010) A geospatial modelling approach integrating archaeobotany and genetics to trace the origin and dispersal of domesticated plants. *PLoS ONE* 5(8):e12060.
21. Vigne JD, Carrère I, Guilaine J (2003) Unstable status of early domestic ungulates in the near east: The example of Shillourokambos (Cyprus, IX–VIIIth millennia cal. B.C.). *Bulletin Correspondance Hellénique Suppl* 43:239–251.
22. Kingman JFC (1982) The coalescent. *Stochastic Process Appl* 13(3):235–248.
23. Hoggart CJ, et al. (2007) Sequence-level population simulations over large genomic regions. *Genetics* 177(3):1725–1731.
24. Currat M, Excoffier L (2004) Modern humans did not admix with Neanderthals during their range expansion into Europe. *PLoS Biol* 2(12):e421.
25. Kitchen JL, Allaby RG (2012) The limits of mean-field heterozygosity estimates under spatial extension in simulated plant populations. *PLoS ONE* 7(8):e43254.
26. Cillière K, Blum MG, Gaggiotti OE, François O (2010) Approximate Bayesian Computation (ABC) in practice. *Trends Ecol Evol* 25(7):410–418.
27. Boyd R, Richerson PJ, Henrich J (2011) Rapid cultural adaptation can facilitate the evolution of large-scale cooperation. *Behav Ecol Sociobiol* 65(3):431–444.
28. Bramanti B, et al. (2009) Genetic discontinuity between local hunter-gatherers and central Europe's first farmers. *Science* 326(5949):137–140.
29. Chikhi L, Nichols RA, Barbuani G, Beaumont MA (2002) Y genetic data support the Neolithic demic diffusion model. *Proc Natl Acad Sci USA* 99(17):11008–11013.
30. Molina J, et al. (2011) Molecular evidence for a single evolutionary origin of domesticated rice. *Proc Natl Acad Sci USA* 108(20):8351–8356.
31. Beaumont MA, Rannala B (2004) The Bayesian revolution in genetics. *Nat Rev Genet* 5(4):251–261.
32. Marjoram P, Tavaré S (2006) Modern computational approaches for analysing molecular genetic variation data. *Nat Rev Genet* 7(10):759–770.
33. Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* 7:214.
34. Naderi S, et al. (2008) The goat domestication process inferred from large-scale mitochondrial DNA analysis of wild and domestic individuals. *Proc Natl Acad Sci USA* 105(46):17659–17664.
35. Brotherton P, et al.; Genographic Consortium (2013) Neolithic mitochondrial haplogroup H genomes and the genetic origins of Europeans. *Nat Commun* 4:1764.
36. Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian computation in population genetics. *Genetics* 162(4):2025–2035.
37. Bertorelle G, Benazzo A, Mona S (2010) ABC as a flexible framework to estimate demography over space and time: Some cons, many pros. *Mol Ecol* 19(13):2609–2625.
38. Marjoram P, Molitor J, Plagnol V, Tavaré S (2003) Markov chain Monte Carlo without likelihoods. *Proc Natl Acad Sci USA* 100(26):15324–15328.
39. Sousa VC, Fritz M, Beaumont MA, Chikhi L (2009) Approximate Bayesian computation without summary statistics: The case of admixture. *Genetics* 181(4):1507–1519.
40. Wegmann D, Leuenberger C, Excoffier L (2009) Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics* 182(4):1207–1218.
41. Wegmann D, Leuenberger C, Neuenschwander S, Excoffier L (2010) ABCtoolbox: A versatile toolkit for approximate Bayesian computations. *BMC Bioinformatics* 11:116.
42. Joyce P, Marjoram P (2008) Approximately sufficient statistics and Bayesian computation. *Stat Appl Genet Mol Biol* 7(1):26.
43. Nunes MA, Balding DJ (2010) On optimal selection of summary statistics for approximate Bayesian computation. *Stat Appl Genet Mol Biol* 9:34.
44. Keith JM, Spring D (2013) Agent-based Bayesian approach to monitoring the progress of invasive species eradication programs. *Proc Natl Acad Sci USA* 110(33):13428–13433.
45. Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet* 5(10):e1000695.
46. Aandahl RZ, Stadler T, Sisson SA, Tanaka MM (2014) On the reliability of approximate Bayesian computation: Epidemiological parameter estimates from *Mycobacterium tuberculosis* genotypes. *Genetics*, 10.1534/genetics.113.158808.
47. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG (2009) The origins of lactase persistence in Europe. *PLoS Comput Biol* 5(8):e1000491.
48. Vilà C, et al. (1997) Multiple and ancient origins of the domestic dog. *Science* 276(5319):1687–1689.
49. Gronau I, Hubisz MJ, Gulko B, Danko CG, Siepel A (2011) Bayesian inference of ancient human demography from individual genome sequences. *Nat Genet* 43(10):1031–1034.
50. Wang GD, et al. (2013) The genomics of selection in dogs and the parallel evolution between dogs and humans. *Nat Commun* 4:1860.
51. Freedman AH, et al. (2014) Genome sequencing highlights the dynamic early history of dogs. *PLoS Genet* 10(1):e1004016.
52. Pionnier-Capitan M, et al. (2011) New evidence for Upper Palaeolithic small domestic dogs in South-Western Europe. *J Archaeol Sci* 38(9):2123–2140.
53. Larson G, et al. (2012) Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA* 109(23):8878–8883.
54. Naderi S, et al.; Econogene Consortium (2007) Large-scale mitochondrial DNA analysis of the domestic goat reveals six haplogroups with high diversity. *PLoS ONE* 2(10):e1012.
55. Luikart G, et al. (2001) Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci USA* 98(10):5927–5932.
56. Gerbault P, Powell A, Thomas MG (2012) Evaluating demographic models for goat domestication using mitochondrial DNA sequences. *Anthropozoologica* 47(2):65–78.
57. Mamidi S, et al. (2011) Investigation of the domestication of common bean (*Phaseolus vulgaris*) using multilocus sequence data. *Funct Plant Biol* 38(12):953–967.
58. Kwak M, Gepts P (2009) Structure of genetic diversity in the two major gene pools of common bean (*Phaseolus vulgaris* L., Fabaceae). *Theor Appl Genet* 118(5):979–992.
59. Beebe S, et al. (2000) Structure of genetic diversity among common bean landraces of Middle American origin based on correspondence analysis of RAPD. *Crop Sci* 40(1):264–273.
60. Chacón S MI, Pickersgill B, Deboucq DG (2005) Domestication patterns in common bean (*Phaseolus vulgaris* L.) and the origin of the Mesoamerican and Andean cultivated races. *Theor Appl Genet* 110(3):432–444.
61. Vitte C, Ishii T, Lamy F, Brar D, Panaud O (2004) Genomic paleontology provides evidence for two distinct origins of Asian rice (*Oryza sativa* L.). *Mol Genet Genomics* 272(5):504–511.
62. Fuller DQ (2011) Pathways to Asian civilizations: Tracing the origins and spread of rice and rice cultures. *Rice* 4(3–4):78–92.
63. Fuller DQ, et al. (2010) Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeological and Anthropological Sciences* 2(2):115–131.
64. Vaughan DA, Lu B-R, Tomooka N (2008) The evolving story of rice evolution. *Plant Sci* 174(6):394–408.
65. Lindgren G, et al. (2004) Limited number of patrilineal horse domestication. *Nat Genet* 36(4):335–336.
66. Jansen T, et al. (2002) Mitochondrial DNA and the origins of the domestic horse. *Proc Natl Acad Sci USA* 99(16):10905–10910.
67. Vilà C, et al. (2001) Widespread origins of domestic horse lineages. *Science* 291(5503):474–477.
68. Warmuth V, et al. (2012) Reconstructing the origin and spread of horse domestication in the Eurasian steppe. *Proc Natl Acad Sci USA* 109(21):8202–8206.
69. Cornille A, et al. (2012) New insight into the history of domesticated apple: Secondary contribution of the European wild apple to the genome of cultivated varieties. *PLoS Genet* 8(5):e1002703.
70. Zohary D (1999) Monophyletic versus polyphyletic origin of crops found in the Near East. *Genet Resour Crop Evol* 46(2):133–142.
71. Hillman GC, Davies MS (1990) Domestication rates in wild-type wheats and barley under primitive cultivation. *Biol J Linn Soc Lond* 39(1):39–78.
72. Tanno K, Willcox G (2006) How fast was wild wheat domesticated? *Science* 311(5769):1886.
73. Pernès J (1986) L'allogamie et la domestication des céréales: L'exemple du maïs (*Zea mays* L.) et du mil (*Pennisetum americanum* L.). *Sci Hum Bull Soc Bot* 133:27–34.
74. Le Thierry D'Ennequin M, Toupance B, Robert T, Godelle B, Gouyon PH (1999) Plant domestication: A model for studying the selection of linkage. *J Evol Biol* 12(6):1138–1147.
75. Allaby R (2010) Integrating the processes in the evolutionary system of domestication. *J Exp Bot* 61(4):935–944.
76. Chapman MA, et al. (2008) A genomic scan for selection reveals candidates for genes involved in the evolution of cultivated sunflower (*Helianthus annuus*). *Plant Cell* 20(11):2931–2945.
77. Peleg Z, Fahima T, Korol AB, Abbo S, Saranga Y (2011) Genetic analysis of wheat domestication and evolution under domestication. *J Exp Bot* 62(14):5051–5061.
78. Peng J, et al. (2003) Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proc Natl Acad Sci USA* 100(5):2489–2494.

79. Wright SI, et al. (2005) The effects of artificial selection on the maize genome. *Science* 308(5726):1310–1314.
80. Haldane JBS (1957) The cost of selection. *J Genet* 55:511–524.
81. Bekaert M, Edger PP, Hudson CM, Pires JC, Conant GC (2012) Metabolic and evolutionary costs of herbivory defense: Systems biology of glucosinolate synthesis. *New Phytol* 196(2):596–605.
82. Cockram J, et al. (2007) Control of flowering time in temperate cereals: Genes, domestication, and sustainable productivity. *J Exp Bot* 58(6):1231–1244.
83. Fu D, et al. (2005) Large deletions within the first intron in VRN-1 are associated with spring growth habit in barley and wheat. *Mol Genet Genomics* 273(1):54–65.
84. Wu W, et al. (2013) Association of functional nucleotide polymorphisms at DTH2 with the northward expansion of rice cultivation in Asia. *Proc Natl Acad Sci USA* 110(8):2775–2780.
85. Shennan S, et al. (2013) Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nat Commun* 4:2486.
86. Stevens CJ, Fuller DQ (2012) Did Neolithic farming fail? The case for a Bronze Age agricultural revolution in the British Isles. *Antiquity* 86(333):707–722.
87. Richerson PJ, Boyd R, Bettinger RL (2001) Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. *Am Antiq* 66(3):387–411.
88. Hawks J, Wang ET, Cochran GM, Harpending HC, Moyzis RK (2007) Recent acceleration of human adaptive evolution. *Proc Natl Acad Sci USA* 104(52):20753–20758.
89. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19(6):716–723.
90. Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90(430):773–795.
91. Box GEP (1976) Science and statistics. *J Am Stat Assoc* 71(356):791–799.
92. Allen TFH, Tainter JA, Pires JC, Hoekstra TW (2001) Dragnet ecology, “Just the facts ma’am”: The privilege of science in a postmodern world. *Bioscience* 51(6): 475–485.

Multiple lines of evidence for the origin of domesticated chili pepper, *Capsicum annuum*, in Mexico

Kraig H. Kraft^{a,1}, Cecil H. Brown^b, Gary P. Nabhan^c, Eike Luedeling^d, José de Jesús Luna Ruiz^e, Geo Coppens d'Eeckenbrugge^f, Robert J. Hijmans^g, and Paul Gepts^{a,2}

^aDepartment of Plant Sciences, Section of Crop and Ecosystem Sciences, University of California, Davis, CA 95616-8780; ^bDepartment of Anthropology, Northern Illinois University, Pensacola, FL 32503-6634; ^cSouthwest Center, University of Arizona, Tucson, AZ 85721-0185; ^dWorld Agroforestry Centre, Nairobi, Kenya; ^eCentro de Ciencias Agropecuarias, Universidad Autónoma de Aguascalientes, Aguascalientes, Mexico CP 20131; ^fCentre de coopération Internationale en Recherche Agronomique pour le Développement, Unité Mixte de Recherche 5175 Centre d'Ecologie Fonctionnelle et Evolutive, Campus Centre National de la Recherche Scientifique, Montpellier Cedex 5, France; and ^gDepartment of Environmental Science and Policy, University of California, Davis, CA 95616

Edited by Dolores R. Piperno, Smithsonian National Museum of Natural History and Smithsonian Tropical Research Institute, Fairfax, Washington, DC, and approved December 4, 2013 (received for review September 6, 2013)

The study of crop origins has traditionally involved identifying geographic areas of high morphological diversity, sampling populations of wild progenitor species, and the archaeological retrieval of macroremains. Recent investigations have added identification of plant microremains (phytoliths, pollen, and starch grains), biochemical and molecular genetic approaches, and dating through ¹⁴C accelerator mass spectrometry. We investigate the origin of domesticated chili pepper, *Capsicum annuum*, by combining two approaches, species distribution modeling and paleobiolinguistics, with microsatellite genetic data and archaeobotanical data. The combination of these four lines of evidence yields consensus models indicating that domestication of *C. annuum* could have occurred in one or both of two areas of Mexico: northeastern Mexico and central-east Mexico. Genetic evidence shows more support for the more northern location, but jointly all four lines of evidence support central-east Mexico, where preceramic macroremains of chili pepper have been recovered in the Valley of Tehuacán. Located just to the east of this valley is the center of phylogenetic diversity of Proto-Otomanguean, a language spoken in mid-Holocene times and the oldest protolanguage for which a word for chili pepper reconstructs based on historical linguistics. For many crops, especially those that do not have a strong archaeobotanical record or phylogeographic pattern, it is difficult to precisely identify the time and place of their origin. Our results for chili pepper show that expressing all data in similar distance terms allows for combining contrasting lines of evidence and locating the region(s) where cultivation and domestication of a crop began.

The analysis of plant macroremains, morphological variation in crop varieties, and identification of wild progenitor species (as determined through their ability to hybridize with the crop) constitute traditional methods for studying crop origins (1, 2). Currently, analysis of microremains such as starch grains, accelerator mass spectrometry (AMS) ¹⁴C radiocarbon dating, along with biochemical and molecular genetic analyses of wild and domesticated populations are also used to date and locate geographic areas of domestication (3, 4).

This set of approaches is extended here with two additional methods, species distribution modeling and paleobiolinguistics, integrating these in a comprehensive study of the origin of domesticated chili pepper, *Capsicum annuum* L., the world's most widely grown spice. *C. annuum* is one of five domesticated pepper species, which also include *Capsicum baccatum* L., *Capsicum chinense* Jacq., *Capsicum frutescens* L., and *Capsicum pubescens* Ruiz & Pav. The ~30 species of *Capsicum* are all native to the Americas (5). Comparing karyotypes of wild and domesticated *C. annuum* (var. *glabriusculum* and var. *annuum*, respectively), Pickersgill (6) identified Mexico as the general

region of domestication of this pepper. Loaiza-Figueroa et al. (7) used allozyme similarity to identify putative wild ancestral populations for chili pepper in a larger collection of wild and domesticated populations. They narrowed the likely domestication area to the eastern Mexican states of Tamaulipas, Nuevo León, San Luis Potosí, Veracruz, and Hidalgo. Since these investigations, others have sought to determine genetic relationships among wild and domesticated populations of chili pepper (8, 9).

The oldest macroremains unambiguously identified as *Capsicum* pepper were retrieved from preceramic strata of dry caves in two states of Mexico: Puebla (Tehuacán Valley; refs. 10, 11) and Tamaulipas (Ocampo caves; ref. 12) (Fig. 1A). These were found with macroremains of maize (*Zea mays*), squash (*Cucurbita* spp.), and other species used by humans, all of which, at both sites, were indirectly dated through associations in archaeological strata, suggesting a rough date for the chili pepper macroremains of around 9000–7000 B.P. (13). Subsequently, remains of maize from Tehuacán were dated directly by AMS and found to be more recent, 5600 y calibrated B.P. (14). AMS dating applied to bottle gourd and squash from Ocampo also yielded more recent ages, 6400–6000 y calibrated B.P. (15). Whereas no AMS dates

Significance

The novelty of the information of this manuscript resides in the addition of species distribution modeling and paleobiolinguistics data, combined with genetic and existing archaeobotanical data, to trace back the geographic origin of a crop, namely domesticated pepper, *Capsicum annuum*. Furthermore, the utilization of a geographic framework of reference for the four types of data has allowed us to combine these independent data types into a single hypothesis about the origin of this crop. Our results suggest that food crops in Mexico had a multiregional origin with chili pepper originating in central-east Mexico, maize in the Balsas River Basin and common bean in the Lerma–Santiago River Basin, resembling similar finds for the Fertile Crescent and China.

Author contributions: K.H.K., C.H.B., J.d.J.L.R., R.J.H., and P.G. designed research; K.H.K., C.H.B., J.d.J.L.R., G.C.d., R.J.H., and P.G. performed research; K.H.K., C.H.B., E.L., J.d.J.L.R., R.J.H., and P.G. contributed new reagents/analytic tools; K.H.K., C.H.B., G.P.N., E.L., G.C.d., R.J.H., and P.G. analyzed data; and K.H.K., C.H.B., G.P.N., E.L., J.d.J.L.R., G.C.d., R.J.H., and P.G. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹Present address: Committee on Sustainability Assessment, Philadelphia, PA 19147.

²To whom correspondence should be addressed. E-mail: plgepts@ucdavis.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1308933111/-DCSupplemental.

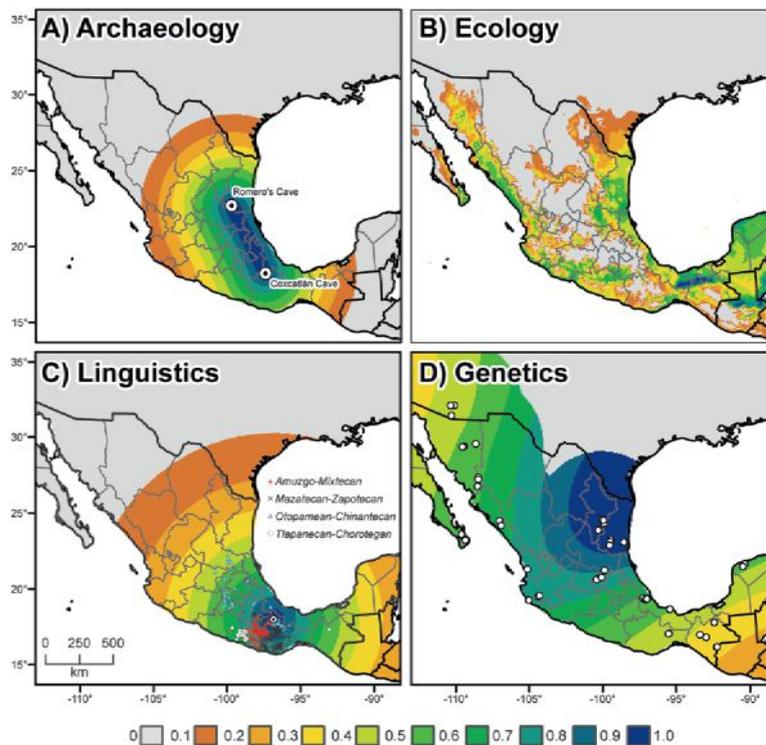


Fig. 1. Possible area of Mexico for *Capsicum annuum* domestication based on (A) archaeological, (B) paleoclimatic, mid-Holocene, (C) linguistic, and (D) genetic data. In addition to the strength of evidence (between 0 and 1), the maps show: (A) Location of the oldest archaeological remains of chili: Romero's Cave, Ocampo, Tamaulipas; Coxcatlán Cave, Tehuacán Valley, Puebla. (C) Location of the homeland of Proto-Otomanguean (dotted circle) and of the four subgroups of current Otomanguean languages (see Fig. S5 legend and Table 1; for a more detailed view of the subgroup distribution, see Fig. S5). Open circles represent approximate locations of protolanguages with a reconstructed word for chili. (D) Open circles indicate the location of the wild chili samples used in the genetic distance analysis (29). For explanation of values, see *Materials and Methods*.

have been recorded for the Tehuacán and Ocampo remains of chili pepper, remains from Guilá Naquitz and Silvia's Caves in the arid eastern valley of Oaxaca state were dated indirectly by AMS to 1400–500 B.P. (16). Rock shelters in the seasonally dry tropical forest of the Central Balsas watershed (state of Guerrero) have produced phytoliths and starch grain residue for domesticated maize and squash (*Cucurbita* sp.) dated by association to around 9000 B.P. (17). However, no remains of *Capsicum* pepper have been found at that site.

Species distribution modeling (SDM) can be used to predict areas that are environmentally suitable for a species from the sites where it is known to occur (18). In SDM, locations of the known current distribution of a species are compiled; values for climatic predictor variables at these locations and a large set of random (background) locations are extracted from spatial databases; and the climatic values are used to fit a model that estimates the similarity of the climate in any location to climatic conditions at known occurrence locations, using a machine-learning algorithm such as MaxEnt (19). The model is then used to predict the climatic suitability for a species across an area of interest. This prediction can be made using current climate data, but the model can also be “transferred” in time, by using past or future climate data simulated by global climate models (GCMs). This approach has been used for many purposes, including to predict the effect of climate change on the geographic distribution of crop wild relatives (20) and to successfully locate unknown *Capsicum* populations (21).

Crop origins can also be studied using paleobiolinguistics (PBL), which employs the comparative method of historical linguistics to reconstruct the biodiversity known to human groups of the remote, unrecorded past (22–24). By comparing words for a species in modern languages, terms for plants and animals in

ancestral languages can be retrieved. The presence of words for a species in an ancestral language is an indication of the species' significance to speakers of that language (25, 26), if not their status as domesticated plants. PBL uses Automated Similarity Judgment Program (ASJP) chronology for estimating the latest date at which a protolanguage was spoken based on lexical similarity (27). Lexical similarity found among related languages is calibrated with historical, epigraphic, and archaeological divergence dates for 52 language groups. In addition, the general area in which an ancestral language was spoken, i.e., the protolanguage homeland, can be approximately determined by locating the area where its modern descendant languages are found to be most diverse (28).

In this paper, we complement existing archaeobotanical data with ecological, paleobiolinguistic, and molecular diversity data to identify the region of initial intensification of human interest in chili pepper that led to crop domestication. The novelty of our approach resides in the addition of SDM and PBL to this type of analysis and the expression of all lines of evidence in comparable spatially explicit units (distance to the area of origin) that allows for their integration into a single prediction.

Results

Archaeological Evidence. The remains from Tehuacán and Ocampo constitute at present the oldest macrobotanical evidence for preceramic chili pepper in the New World. Although these chili specimens cannot be identified as cultivated or domesticated, their archaeological association with domesticated remains of important crops, such as maize and squash, is strongly suggestive of ancient intensive human interaction with chili in these areas. Based on this evidence, we assumed that the nearer a place may be to either of these sites, the more likely the location was part of

the region where the crop was first grown and domesticated (Fig. 1A).

Ecological Evidence. Wild chili pepper (*C. annuum* var. *glabriusculum*), the ancestor of domesticated *C. annuum* (6), is a perennial shrub that produces dozens of erect, globular, pea-sized fruits. The fruits are consumed and dispersed by frugivorous birds, which pass the seed through their digestive system. Generally found in the northern half of Mexico, the wild chili pepper is associated with a nurse plant—often a hackberry (*Celtis pallida* Torey), a mesquite (*Prosopis* sp.), or columnar cacti. As one moves further southwards, wild chili pepper is found more frequently in human-disturbed landscapes—fence rows, home gardens, and roadsides (29). Based on our own collecting localities and those of herbarium specimens and gene bank accessions (29), we estimate that wild chili peppers grow currently in environments with a median annual average temperature of 24 °C and between 20 °C and 26 °C for 90% of the locations. The coldest locations with known wild pepper populations are mostly in the central Mexican highlands, the warmest locations in the southern coastal regions of Mexico and Guatemala. The median annual rainfall of these locations is 907 mm, and between 495 and 2,253 mm for 90% of the locations, with the driest locations in the northwestern part of the distribution (e.g., Baja California and the Sonoran Desert) and the wettest locations in southeastern Mexico.

The MaxEnt species distribution model had an internal (training) fit area under the curve (AUC) of the receiver operating characteristic (ROC) curve of 0.89. The average cAUC (bias corrected) obtained with fivefold cross-validation was 0.80, which suggests that the model has very good predictive power (30). The two most important predictor variables (based on permutation

importance) were mean temperature of the coldest quarter (53%), followed by annual precipitation (14%).

Under the climate conditions of the mid-Holocene (about 6000 B.P.), the regions predicted to be most suitable for wild chili pepper include areas along the western and eastern coasts of Mexico, southeast Mexico and northern Guatemala (Fig. 1B). The central highlands were clearly unsuitable for this species during this period. The correlation coefficient between the predicted suitability for the current climate (Fig. S1) and the mid-Holocene climate was 0.92. Despite this overall similarity, there were important differences between these predictions, with areas in the southeast of Mexico more suitable and areas in the northeast less suitable during the mid-Holocene (Fig. S2).

Paleobiolinguistic Evidence. Brown (22) surveyed the reconstructed vocabularies of 30 protolanguages of Mesoamerica (southern half of Mexico and northern Central America) and abutting areas for terms for 41 different crops, including chili pepper. His survey presented for each protolanguage the estimated date it was spoken at the latest, making it possible to stratify reconstructed words for crops chronologically (Table 1) (27, 28).

Proto-Otomanguean is the oldest (~6500 B.P.) protolanguage of the New World for which a word for chili pepper reconstructs (31). All daughter languages of Proto-Otomanguean, as defined by Kaufman (32), show reconstructed terms for chili pepper (Table 1). Given that estimated dates are to be understood as the latest dates at which ancestral languages were spoken, it is plausible that speakers of Proto-Otomanguean actually had a word for chili pepper hundreds, if not thousands, of years before ~6500 B.P. The oldest protolanguage of Table 1 not belonging to the Otomanguean family is Proto-Totozoquean (~4300 B.P.), for which a term for chili pepper does not reconstruct. Non-Otomanguean languages for

Table 1. Reconstruction of terms for *Capsicum* in selected protolanguages of Mesoamerica and abutting areas

Years before present	Protolanguage [†]	Reconstructed word for chili [‡]	Location of modern descendant languages	Genetic affiliation
6591	Otomanguean	*ʔki ³	Mexico	Otomanguean
5976	Eastern Otomanguean	*(h)saH ³ , *ki	Mexico	Otomanguean
5498	Popolocan-Zapotecan	*ki	Mexico	Otomanguean
5357	Amuzgo-Mixtecan	*nʔsah ³	Mexico	Otomanguean
4542	Mixtecan	*(H)yaʔ, Hyah, H0aʔ ²	Mexico	Otomanguean
4274	Totozoquean	NR	Mexico	Totozoquean
4018	Uto-Aztecan	NR	US Southwest, Mexico, Central America	Uto-Aztecan
3654	Otopamean	*(m)ʔi	Mexico	Otomanguean
3472	Southern Uto-Aztecan	NR	Mexico, Central America	Uto-Aztecan
3434	Kiowa-Tanoan	NR	US Southwest	Kiowa-Tanoan
3149	Zapotecan	*ki:ʔnaʔ	Mexico	Otomanguean
3036	Popolocan	*hña	Mexico	Otomanguean
3000	Lencan	NR	Central America	Lencan
2774	Misumalpan	kuma	Central America	Misumalpan
2576	Northern Uto-Aztecan	NR	US Southwest	Uto-Aztecan
2445	Chiapanec-Mangue	*ni:- ⁿ giʔ	Mexico	Otomanguean
2400	Sonoran	*koʔokoLi	Mexico	Uto-Aztecan
2220	Mayan	*i:hk	Mexico, Central America	Mayan
1935	Chinantecan	*ʔu ^{HL}	Mexico	Otomanguean
1865	Yuman	NR	US Southwest	Yuman
1737	Numic	NR	US Southwest	Uto-Aztecan
1587	Takic	NR	US Southwest	Uto-Aztecan
1509	General Aztec	*či:l	Mexico	Uto-Aztecan
1435	Totonacan	*pi'n	Mexico	Totozoquean
1407	Mixe-Zoquean	*ni:wi	Mexico	Totozoquean

NR, not reconstructable.

[†]Sources for each language are listed in *SI Materials and Methods* under *Paleobiolinguistics*.

[‡]Explanations for phonetic representation of pepper words are listed in *SI Materials and Methods* under *Paleobiolinguistics*.

which a term for chili pepper reconstructs are Proto-Misumalpan (~2800 B.P.), Proto-Sonoran (~2400 B.P.), and Proto-Mayan (~2200 B.P.). Thus, the earliest non-Otomanguean dates for *Capsicum* in Mesoamerica and abutting regions are over 3,700 y more recent than the oldest date, suggesting that speakers of a prehistoric Otomanguean language or languages may have been among the first cultivators or domesticators of chili pepper. Note that the current word—chili—is derived from the General Aztec language, Nahuatl, which reconstructs to a much more recent date (~1500 B.P.; Table 1).

The area of maximum diversity of a language family has been viewed traditionally by linguists as suggestive of the location of a family's ancestral language (e.g., ref. 28). We use this phylogenetic diversity information in locating the Otomanguean homeland by identifying where languages of the four subgroups of the family—Mazatecan-Zapotecan, Amuzgo-Mixtecan, Tlapanecan-Chorotegan, and Otopamean-Chinantecan (32)—are currently spoken in closest proximity (Fig. 1C).

Genetic Evidence. During the fall of 2006 and 2007, expeditions were conducted in the southern United States and throughout Mexico to sample populations of wild *C. annuum* (29). This provided the most complete set of wild *C. annuum* from Mexico available to date. Based largely on this set, 139 wild types distributed over the entire exploration area were chosen as were 49 domesticated types that are endemic landraces (*ancho*, *puya*, and *guajillo*) (33). This collection was screened with 17 simple sequence repeat (SSR) DNA markers (34, 35). These markers were chosen for this study because of their consistency of amplification and polymorphism within our sample. For each wild plant, a distance was calculated to the domesticated group based on the average proportion of shared SSR alleles. These distances were then spatially interpolated to produce in each grid cell an estimated genetic similarity between wild pepper populations (if any occurred in the cell) and the group of domesticated chili peppers (regardless of where they occurred). This molecular-marker-based analysis of genetic similarity between wild and domesticated types revealed a broad area of high similarity in the northeastern quadrant of Mexico (Fig. 1D), including the states of Tamaulipas, Nuevo León, San Luis Potosí, and Veracruz. In contrast, genetic similarity between wild and cultivated types was generally low in southern and northwest Mexico, confirming earlier results (7).

Consensus Model. The four lines of evidence—archaeological, ecological, paleobiolinguistic, and genetic—were all expressed as

a spatial model and they can therefore be combined into a single consensus model represented geographically through mapping. Each type of evidence has its particular strengths and weaknesses, discussed below, which need to be taken into consideration when producing a consensus model. Because these merits and demerits are difficult to quantify (some are simply unknown), assigning differential weights to each line of evidence is problematic. Our solution is to present a number of different consensus maps based on several different weighting combinations (Fig. 2 and Fig. S3).

The first map, Fig. 2A, was established using equal weighting for each type of evidence (each weighted as making a 1/4 contribution). According to this model, areas in central-east Mexico and northeastern Mexico are the most likely area of origin of chili pepper. The second model assigned a high weight to genetic evidence (weighted 1/2) and equal but lower weights to the other three lines of evidence (each weighted 1/6). This assumes that genetic data might be superior to one or more of the other lines of evidence used because, for example, it might suffer less from sampling bias. This results in primary support for northeastern Mexico and only secondary support for central-east Mexico (Fig. 2B). The third approach assigned a low weight to archaeology (1/10) and equal higher weights to the other three lines of evidence (each weighted 1/3). This weighting was motivated by the observation that the current archaeological data are assembled from macroremains of only two sites in Mexico. This weighting produces a consensus model resembling the equal weighting of Fig. 2A because both central-east Mexico and northeastern Mexico result as equally plausible geographic candidates for chili pepper domestication (Fig. 2C). Additional information, from other sites and microremains, yet to be discovered, would justify a stronger weighting for archaeobotanical data.

Another weighting strategy produces different models based on randomly assigning combinations of weights for the four types of evidence. This approach allows us to explore the universe of possible weight combinations given different interpretations of the individual lines of evidence. Fig. S3 shows the percentile distribution obtained for this approach. The resulting maps suggest again that either central-east Mexico or northeastern Mexico or, conceivably, both areas were locations of the domestication of *C. annuum*.

Discussion

We have embraced the template of multidisciplinary approaches to study crop origins proposed first by de Candolle (36) and later

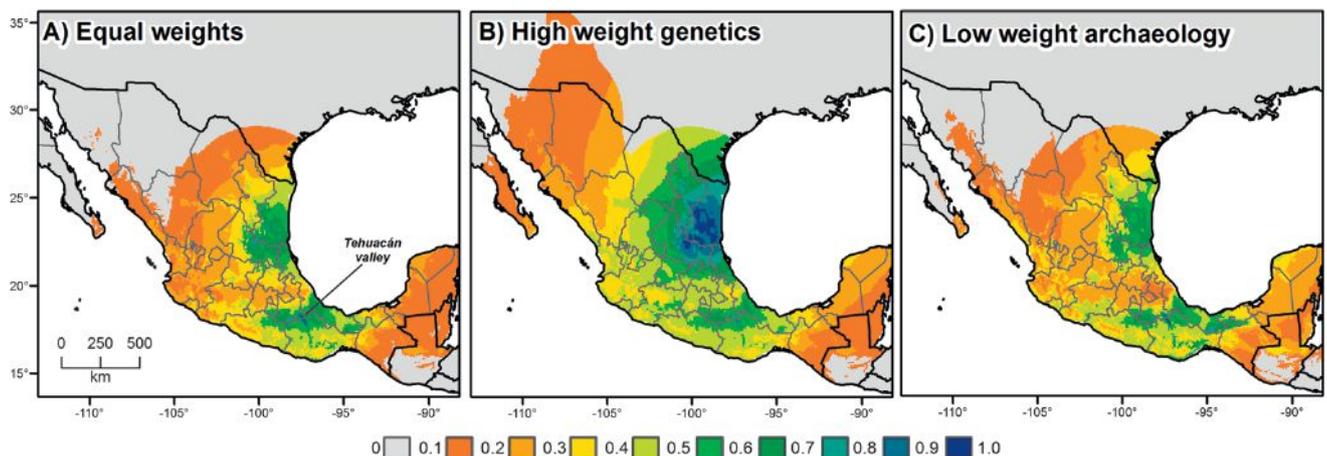


Fig. 2. Consensus models of the likelihood that cultivated chili pepper originated in an area. The models were obtained by combining the four lines of evidence for the origin of domesticated chili pepper (Fig. 1). (A) equal weights; (B) genetics 1/2, all others 1/6 weight; and (C) archaeology 1/10, all others 1/3 weight. After combining, the values were scaled between 0 and 1 and then squared to give more weight to the higher values.

by Harlan and de Wet (37). Confidence in a crop-origin hypothesis is increased when supported by multiple, independent lines of evidence, and improved understanding comes from new evidence in each field and concomitant predictions in other fields (38). Our multidisciplinary approach depends on the independence and strength of evidence from the different fields, each of which has its strengths and weaknesses.

Current archaeobotanical data for chili pepper is mainly based on macroremains from only two sites. In addition to the identification of ancient chili remains at additional sites, our understanding could benefit from the investigation of microfossil data such as starch grains (39) in Mesoamerican sites. Availability of microfossils may provide information on the more ancient distribution and importance of chili peppers and potentially also help distinguish domesticated from nondomesticated remains (as in the case of maize) (40).

The quality of species distribution models depends on having a representative sample of the current distribution of the wild species, the quality of climate data, particularly the modeled past climate data, and the algorithm used. Our sample size was large and the species is widespread, suggesting that the SDM approach should work well (41), as confirmed by a high cAUC score (30). Backcasted climate data for the mid-Holocene is, of course, uncertain; furthermore, we did not consider climate variation during that period. Nevertheless, because we use an ensemble of climate models (Fig. S4), our predictions should be relatively robust (42).

Utilization of linguistic data assumes an understanding of language development, including information relating to language origin, dispersal, and diffusion of traits across languages that is still emerging as new computer approaches are increasingly applied in linguistic analysis (43, 44). PBL provides an assessment of when species acquired substantial salience for prehistoric groups, whether they were merely harvested, cultivated, or eventually domesticated. If a word for a biological species reconstructs for a protolanguage, this is evidence that the species was known to and probably of considerable importance to speakers of the language as shown by Berlin et al. (25) for two closely related Mayan languages, Tzeltal and Tzotzil (Tzeltalan) and by Balée and Moore (26) in a study of plant names in five Eastern Amazonian Tupi-Guaraní languages.

Genetic data are generally based on the analysis of contemporary populations of the wild ancestor of the crop. The wild populations included in this study constitute the largest and most widespread sample used in genetic analyses for this species (29). However, we do not know to what extent the distribution and genetic structure of these populations have changed over the past 6,000 y. Hence, we modeled the past distribution of wild chili peppers based on the assumption that their climatic requirements are the same as today's wild chili pepper population. Correlation between the suitability scores for ancient and current distributions is high (0.92), suggesting that, whereas climate change over the past 6,000 y has likely shifted the species distribution, for the most part, the historical and current ranges of this species overlap. Another potentially confounding factor is gene flow between domesticated and wild chili peppers, which may cause similarities that are not due to ancestor–descendant relationships (45). However, this would not seem very important for chili peppers because they are mostly a self-fertilizing species with minimal outcrossing, which is confirmed by the high levels of homozygosity observed for wild chili pepper populations analyzed here (Dataset S1, *Microsat info*).

The concept of origin of *C. annuum* used in this study encompasses wild plant protection, management, cultivation, and domestication. Within this continuum of increasingly close interaction between humans and plants, distinguishing among these four stages for most crops is difficult. However, with respect to chili pepper, the fact that a Proto-Otomanguean word for the crop was retained in daughter languages attests to its high salience for speakers of the protolanguage. Furthermore, Proto-Otomanguean speakers may have been actively engaged in cultivation, as suggested by the

reconstruction of words for a range of plants, including staple crops such as maize and squash, but also other crops such as avocado and nopal (22). Plausibly, then, the saliency of chili pepper among Proto-Otomanguean speakers reflects cultivation and perhaps incipient domestication and not merely use of a wild plant species.

When analyzed separately, our four lines of evidence do not all suggest the same geographic area as being the most likely place of chili pepper origin. Nevertheless, we identify central-east Mexico as a likely region of initial cultivation or incipient domestication because that interpretation most parsimoniously reconciles all evidence (Fig. 2A). This area extends from southern Puebla and northern Oaxaca to southern Veracruz and encompasses the valley of Tehuacán (Fig. 2A). The Coxcatlán Cave from which preceramic macroremains of chili pepper have been recovered (13) is situated in this valley. Species distribution modeling shows that many parts of the identified area were suited for the wild progenitor of *C. annuum* around the time of first cultivation or domestication in the mid-Holocene and there are currently populations of wild chili pepper that are genetically similar to the domesticated species (Fig. 1D). Near to the valley is the likely center of the Otomanguean homeland. Proto-Otomanguean, spoken in mid-Holocene times some 6,500 y ago, is the oldest ancestral language of the New World for which a term for chili pepper reconstructs. Speakers of contemporary Otomanguean languages live in or close to the region. Otomanguean people, then, may have been the first in the New World to transform wild chilies into the domesticated spice and condiment so widely enjoyed today.

By expressing all data as a distance, whether geographical (archaeological and linguistic data), climatic, or genetic, we have developed a method to bring together different lines of evidence about crop origins into a single framework of analysis. This approach has led to the discovery that the origin of domesticated chili peppers may have been located further south than previously thought (7) and in different regions of Mexico than proposed for common bean (46) or maize (47). Thus, our data do not suggest a single, nuclear area for crop domestication in Mesoamerica, but rather a multiregional model as suggested also for the Southwest Asian (48) and Chinese (49) centers of agricultural origins.

Materials and Methods

Archaeology. We used two locations for which there is evidence of the earliest use of chili: Romero's Cave (near Ocampo, Tamaulipas) and Coxcatlán Cave (Tehuacán Valley, Puebla) (Fig. 1A). We connected these locations by their shortest path, and then computed the distance d (in kilometers) to this path for cells on a raster with 1-km² spatial resolution. We truncated the distances at 1,000 km and used an inverse squared distance decay function, scaled between 0 and 1, $(1 - (d/1,000))$ as a measure of the likelihood that chili was domesticated in a location (grid cell).

Species Distribution Modeling. We used SDM to assess spatial variation in suitability for wild *C. annuum* var. *glabriusculum*, the ancestor of domesticated *C. annuum* (6), during climatic conditions of the mid-Holocene (about 6,000 y ago) (Fig. 1B). Locations where wild *Capsicum* populations currently occur were from collections made in the fall of 2006 and 2007 (Dataset S2, *Coordinates_w_SSR info*) (29) and from additional records obtained from the Global Biodiversity Information Facility (GBIF) (Dataset S3, *Coordinates GBIF*). We used the SDM algorithm MaxEnt (19) to predict suitability during the mid-Holocene according to nine global climate models (*SI Materials and Materials*).

Paleobiolinguistics. The different languages considered are listed in Table 1 and the respective information sources are compiled in *SI Materials and Methods*. Protolanguage dates (Table 1) were calculated through use of ASJP chronology (27). The center of phylogenetic diversity of Otomanguean languages was located from the distribution of places where Otomanguean languages are currently spoken (50), by determining the area where languages spoken in close geographic proximity to one another are found to be affiliated with the largest number of major divisions of the family.

Genetic Distance Analysis. Genetic distance between 139 wild and 49 domesticated pepper accessions (Datasets S2 and S3, *Coordinates GBIF*) were assessed

with data from 17 microsatellite markers (SSRs) developed before this study as described in Dataset S1, *Microsat info*.

Consensus Model. All four data sources were used to create a spatial model on a common raster. All models had values between 0 and 1, with higher scores indicating that a location is more likely to be the area where domestication occurred. We combined these four sources of data into a single consensus model by assigning weights to each indicator. To get a more pronounced differentiation between sites, we squared the values, after first rescaling them between 0 and 1.

ACKNOWLEDGMENTS. Suzanne K. Fish, Eric W. Holman, and Søren Wichmann provided help for this project in its early phase. Holman and Wichmann also

read and commented on the manuscript as did Gene Anderson, Roger Blench, Eric Campbell, Charles Clement, Norman Hammond, Matt Hufford, Sarah Metcalfe, Barbara Pickersgill, Anthony J. Ranere, Brian Stross, and Eric Votava. K.H.K. thanks Horacio Villalón and Sergio Hernández Verdugo for contributions of wild *Capsicum*; Heather Zornetzer for assistance during field collection; and Derek van den Abeelen, Raúl Durán, Tiffany Chan, Jonathan Kong, and James Kami for assistance in the laboratory work for the genetic analyses; and the Fulbright program, the University of California Institute for Mexico and the United States (UC MEXUS), and the Department of Plant Sciences (Graduate Student Research assistantship) for funding. We thank the World Climate Research Programme's Working Group on Coupled Modelling (CMIP5) and the climate modeling groups for making their model output available.

1. Smith B (1995) *The Emergence of Agriculture* (Scientific American Library, New York).
2. Gepts P, et al. (2012) *Biodiversity in Agriculture: Domestication, Evolution, and Sustainability* (Cambridge Univ Press, Cambridge, UK), pp 630.
3. Gepts P (2004) Domestication as a long-term selection experiment. *Plant Breed Rev* 24(Part 2):1–44.
4. Burke JM, Burger JC, Chapman MA (2007) Crop evolution: From genetics to genomics. *Curr Opin Genet Dev* 17(6):525–532.
5. Bosland PW (1994) Chiles: History, cultivation and uses. *Spices, Herbs and Edible Fungi*, ed Charambous G (Elsevier, New York).
6. Pickersgill B (1971) Relationships between weedy and cultivated forms in some species of chili peppers (genus *Capsicum*). *Evolution* 25:683–691.
7. Loaiza-Figueroa F, Ritland K, Cancino JAL, Tanskley SD (1989) Patterns of genetic variation of the genus *Capsicum* (Solanaceae) in Mexico. *Plant Syst Evol* 165:159–188.
8. Aguilar-Meléndez A, Morrell PL, Roose ML, Kim SC (2009) Genetic diversity and structure in semiwild and domesticated chiles (*Capsicum annuum*; Solanaceae) from Mexico. *Am J Bot* 96(6):1190–1202.
9. Votava EJ, Nabhan GP, Bosland PW (2002) Genetic diversity and similarity revealed via molecular analysis among and within an in situ population and ex situ accessions of chiltepin (*Capsicum annuum* var. *glabriusculum*). *Conserv Genet* 3(2):123–129.
10. Smith CE (1967) Plant remains. *The Prehistory of the Tehuacan Valley*, ed Byers DS (Univ of Texas Press, Austin, TX), pp 220–255.
11. Smith CE, Jr. (1987) Current archaeological evidence for the beginning of American agriculture. *Studies in the Neolithic and Urban Revolutions, The V. Gordon Childe Colloquium*, ed Manzanilla L (British Archaeological Reports, Oxford, UK), pp 81–101.
12. Mangelsdorf PC, McNeish RS, Willey GR (1965) Origins of Middle American agriculture. *Natural Environment and Early Cultures*, ed West RC (Univ. of Texas Press, Austin, Texas), pp 427–445.
13. McClung de Tapia A (1992) The origins of agriculture in Mesoamerica and Central America. *The Origins of Agriculture: An International Perspective*, eds Cowan CW, Watson PJ (Smithsonian Institution Press, Washington, DC), pp 143–171.
14. Long A, Benz B, Donahue D, Jull A, Toolin L (1989) First direct AMS dates on early maize from Tehuacán, Mexico. *Radiocarbon* 31:1035–1040.
15. Smith BD (1997) Reconsidering the Ocampo Caves and the era of incipient cultivation in Mesoamerica. *Latin American Antiquity* 8(4):342–383.
16. Perry L, Flannery KV (2007) Precolonial use of chili peppers in the Valley of Oaxaca, Mexico. *Proc Natl Acad Sci USA* 104(29):11905–11909.
17. Ranere AJ, Piperno DR, Holst I, Dickau R, Iriarte J (2009) The cultural and chronological context of early Holocene maize and squash domestication in the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106(13):5014–5018.
18. Elith J, Leathwick JR (2009) Species distribution models: Ecological explanation and prediction across space and time. *Annu Rev Ecol Syst* 40(1):677–697.
19. Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. *Ecol Modell* 190(3–4):231–259.
20. Jarvis A, Lane A, Hijmans RJ (2008) The effect of climate change on crop wild relatives. *Agric Ecosyst Environ* 126(1–2):13–23.
21. Jarvis A, et al. (2005) Use of GIS for optimizing a collecting mission for a rare wild pepper (*Capsicum flexuosum* Sendtn.) in Paraguay. *Genet Resour Crop Evol* 52(6): 671–682.
22. Brown CH (2010) Development of agriculture in prehistoric Mesoamerica: The linguistic evidence. *Pre-Columbian Foodways: Interdisciplinary Approaches to Food, Culture and Markets in Ancient Mesoamerica*, eds Staller JE, Carrasco MD (Springer, Berlin), pp 71–107.
23. Fowler CS (1972) Some ecological clues to Proto-Numic homelands. *Desert Research Institute Publications in the Social Sciences* 8:105–117.
24. Hill JH (2001) Proto-Uto-Aztecan: A community of cultivators in central Mexico? *Am Anthropol* 103:913–934.
25. Berlin B, Breedlove DE, Laughlin RM, Raven PH (1973) Cultural significance and lexical retention in Tzeltal-Tzotzil ethnobotany. *Meaning in Mayan Languages*, ed Edmonson MA (Mouton, The Hague), pp 143–164.
26. Balée WL, Moore D (1991) Similarity and variation in plant names in five Tupi-Guarani languages (Eastern Amazonia). *Bulletin of the Florida Museum of Natural History* 35(4):209–262.
27. Holman EW, et al. (2011) Automated dating of the world's language families based on lexical similarity. *Curr Anthropol* 52(6):841–875.
28. Wichmann S, Muller A, Velupillai V (2010) Homelands of the world's language families: A quantitative approach. *Diachronica* 27(2):247–276.
29. Kraft KH, Luna-Ruiz JD, Gepts P (2013) A new collection of wild populations of *Capsicum* in Mexico and the southern United States. *Genet Resour Crop Evol* 60(1): 225–232.
30. Hijmans RJ (2012) Cross-validation of species distribution models: Removing spatial sorting bias and calibration with a null model. *Ecology* 93(3):679–688.
31. Brown CH, Clement CR, Epps P, Luedeling E, Wichmann S (2013) The paleobiolinguistics of domesticated chili pepper (*Capsicum* spp.). *Ethnobiology Letters* 4:1–11.
32. Kaufman TS (1990) Early Otomanguean homeland and cultures: Some premature hypotheses. (University of Pittsburgh Working Papers in Linguistics 1), pp 91–136.
33. Kraft KH, de Jesús Luna-Ruiz J, Gepts P (2010) Different seed selection and conservation practices for fresh market and dried chili farmers in Aguascalientes, Mexico. *Econ Bot* 64(4):318–328.
34. Lee JM, Nahm SH, Kim YM, Kim BD (2004) Characterization and molecular genetic mapping of microsatellite loci in pepper. *Theor Appl Genet* 108(4):619–627.
35. Minamiyama Y, Tsuru M, Hirai M (2006) An SSR-based linkage map of *Capsicum annuum*. *Mol Breed* 18(2):157–169.
36. de Candolle A (1882) *L'Origine des Plantes Cultivées [The Origin of Cultivated Plants]* (Appleton, New York), pp 468.
37. Harlan JR, de Wet JMJ (1973) On the quality of evidence for origin and dispersal of cultivated plants. *Curr Anthropol* 14:51–62.
38. Ammerman AJ, Cavalli-Sforza LL (1984) *The Neolithic Transition and the Genetics of Populations in Europe* (Princeton Univ. Press, Princeton, NJ), pp 176.
39. Perry L, et al. (2007) Starch fossils and the domestication and dispersal of chili peppers (*Capsicum* spp. L.) in the Americas. *Science* 315(5814):986–988.
40. Holst I, Moreno JE, Piperno DR (2007) Identification of teosinte, maize, and Tripsacum in Mesoamerica by using pollen, starch grains, and phytoliths. *Proc Natl Acad Sci USA* 104(45):17608–17613.
41. Wisz MS, et al. (2008) Effects of sample size on the performance of species distribution models. *Divers Distrib* 14(5):763–773.
42. Araújo MB, New M (2007) Ensemble forecasting of species distributions. *Trends Ecol Evol* 22(1):42–47.
43. Perreault C, Mathew S (2012) Dating the origin of language using phonemic diversity. *PLoS ONE* 7(4):e35289.
44. Bouckaert R, et al. (2012) Mapping the origins and expansion of the Indo-European language family. *Science* 337(6097):957–960.
45. Papa R, Gepts P (2004) Asymmetric gene flow and introgression between wild and domesticated populations. *Introgression from Genetically Modified Plants into Wild Relatives and Its Consequences*, eds Den Nijs D, Bartsch D, Sweet J (CABI, Oxon, UK), pp 125–138.
46. Kwak M, Kami JA, Gepts P (2009) The putative Mesoamerican domestication center of *Phaseolus vulgaris* is located in the Lerma-Santiago basin of Mexico. *Crop Sci* 49(2): 554–563.
47. Matsuoka Y, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99(9):6080–6084.
48. Fuller DQ, Willcox G, Allaby RG (2011) Cultivation and domestication had multiple origins: Arguments against the core area hypothesis for the origins of agriculture in the Near East. *World Archaeol* 43(4):628–652.
49. Cohen DJ (2011) The beginnings of agriculture in China: A multiregional view. *Curr Anthropol* 52(5):S273–S293.
50. Avila-Blomberg A de, Moreno-Díaz NG (2008) Distribución de las lenguas indígenas de México [Distribution of the indigenous languages of Mexico]. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, Mexico, DF.

Particularism and the retreat from theory in the archaeology of agricultural origins

Kristen J. Gremillion^{a,1}, Loukas Barton^b, and Dolores R. Piperno^{c,d}

^aDepartment of Anthropology, Ohio State University, Columbus, OH 43210; ^bDepartment of Anthropology, University of Pittsburgh, Pittsburgh, PA 15260; ^cDepartment of Anthropology, Smithsonian National Museum of Natural History, Washington, DC 20013-7012; and ^dSmithsonian Tropical Research Institute, Apartado Postal 0843-03092, Panama, Republic of Panama

Edited by James O'Connell, University of Utah, Salt Lake City, UT, and approved October 28, 2013 (received for review June 26, 2013)

The introduction of new analytic methods and expansion of research into previously untapped regions have greatly increased the scale and resolution of data relevant to the origins of agriculture (OA). As a result, the recognition of varied historical pathways to agriculture and the continuum of management strategies have complicated the search for general explanations for the transition to food production. In this environment, higher-level theoretical frameworks are sometimes rejected on the grounds that they force conclusions that are incompatible with real-world variability. Some of those who take this position argue instead that OA should be explained in terms of local and historically contingent factors. This retreat from theory in favor of particularism is based on the faulty beliefs that complex phenomena such as agricultural origins demand equally complex explanations and that explanation is possible in the absence of theoretically based assumptions. The same scholars who are suspicious of generalization are reluctant to embrace evolutionary approaches to human behavior on the grounds that they are ahistorical, overly simplistic, and dismissive of agency and intent. We argue that these criticisms are misplaced and explain why a coherent theory of human behavior that acknowledges its evolutionary history is essential to advancing understanding of OA. Continued progress depends on the integration of human behavior and culture into the emerging synthesis of evolutionary developmental biology that informs contemporary research into plant and animal domestication.

evolutionary theory | behavioral ecology

Over the last decade there has been a major expansion of knowledge regarding the timing and socioecological context of plant domestication and emerging agricultural systems. This wealth of data is due in large part to methodological innovations (e.g., in genetics and paleogenomics, in the analysis of plant micro- and macroremains and biological residues, and in the physical and biogeochemical analyses of anthropogenic sediments), re-excavations of some important archaeological sites, and the expansion of archaeological research into regions whose record of agricultural origins has been until recently poorly known [such as New Guinea (1, 2), lowland areas of Mesoamerica and northern South America (3–7), and northern and southern China (8–13)]. These research activities have enriched both the scale and resolution of the data relevant to agricultural origins worldwide. One result of this welcome enhancement of the empirical record is wider acknowledgment of the variability in the historical pathways taken by emerging food production systems across space and time. The dichotomy between foraging and food production has been discarded in favor of a continuum of landscape, plant, and animal management strategies that sometimes resist classification. However, for some scholars (2, 14), the richly detailed records of change seem to have dampened the appeal of general explanations for the transition to agriculture.

This trend, coupled with the critical stance toward hypothetico-deductive science currently influential in archaeological thought (15), has motivated many archaeologists to favor interpretive modes of explanation. In origins of agriculture (OA) research, one result of this changing climate has been a proliferation of historical narratives that highlight local events and processes and downplay general principles. Although we recognize the value of empirical rigor, we argue that the progress of OA research is significantly hampered by the denial of explanations derived from a systematic program of theoretically driven hypothesis testing. We also reject assertions that evolutionary theory is teleological, needlessly reductionist, ahistorical, and therefore irrelevant to understanding human behavior. Without a coherent theory of human behavior, arguments about agricultural origins must rely on assumptions that are often intuitive, unstated, and largely implicit. Coupled with the increasing emphasis on human exceptionalism, the erosion of scientific method is troubling because it creates a rift between researchers in biology and archaeology whose collaboration is essential to advancing understanding of OA. To address these weaknesses, we contend that evolutionary theory (broadly construed to include cultural as well as biological processes) must play a central role in OA research.

In arguing for the systematic use of theory as a tool for the advancement of knowledge, we are not making the claim that evolutionary theory is the only possible realm from which to choose. However, neo-Darwinism has withstood repeated challenges to its core elements, having earned status as an overarching framework for explaining the diversity of life. For this reason, evolutionary theory is central to understanding the root causes of human behavior and indeed culture itself (16). An even stronger case for the evolutionary perspective on human behavior can be made now that the distinctive features of cultural transmission and evolution are becoming more fully incorporated into the synthesis (e.g., refs. 17–20).

We divide our discussion of these issues into three sections. The first elaborates upon some of the weaknesses of particularism in OA research. Next, we sharpen the focus to respond to critiques of evolutionary approaches to human behavior [particularly the research program of human behavioral ecology (HBE) and one of its tools, optimal foraging theory (OFT)]. We argue that recent calls to abandon HBE or OFT are misguided,

Author contributions: K.J.G., L.B., and D.R.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: gremillion.1@osu.edu.

and we explain why the particularist alternatives might provide more explanatory depth and power when situated within a program of behavioral ecology. Finally, we assess the prospects for a synthetic evolutionary research program for the study of domestication and agriculture.

Limitations of Particularism in OA Research and the Need for Theory

The research trend that we critique in this paper shares several characteristics with the approach labeled historical particularism by historians of anthropology (21). Most closely associated with Franz Boas and his students, historical particularism emphasized the collection and compilation of ethnographic data without the distorting effects of theory. Boas and his adherents, weary of unsubstantiated theorizing unsupported by rigorous fieldwork, viewed theory as an obstacle to scientific investigation. Generalizations about culture were suspect because exceptions to them could always be found. Any reliable generalizations would be derived inductively from large bodies of data. Similar recommendations and concerns have been voiced in the OA literature (14, 22–24), with the important difference that current expressions of particularism, rather than disavowing theory, embed it in arguments without explicit acknowledgment or relegate it to a minor role that prioritizes empirically based inference.

We argue that the retreat from theory and the embrace of particularism are not only unwarranted but also counterproductive. Particularistic approaches that reject generalization fail to treat inductive inferences as hypotheses to be tested and substitute appeals to agency for a consistent theory of human behavior. These strategies are problematic for reasons we detail below.

Particularistic Approaches Reject General Principles in the Mistaken Belief That They Imply Universal Causes and Ignore History. In the current climate, it is not surprising that some OA researchers call for a more particularistic approach to understanding why human groups became involved in plant cultivation and animal husbandry. Among the reasons given for rejecting generalizations about causality is that contradictory cases can be answered only by adding qualifiers (ref. 2, p. S380). Some investigators conflate universal explanations (i.e., “prime movers” like environmental change, population pressure, or divine inspiration), which imply that all transitions to agriculture conform to the same causal sequence, with the principles and assumptions of HBE, which do not

(23). The result is a rejection of both universal causation and general principles in favor of “frameworks of explanation that pay close and careful attention to existing relevant archaeological information, that are scaled at the regional level, and that focus on the complex interplay of a range of different environmental and social pre-conditions, prompts, and factors of various kinds” (ref. 23, p. 681).

We agree that no single cause explains all instances of transition from foraging to food production (25–27); this observation is true but trivial. An explanation that claims universal validity loses its universality when exceptions are found to exist. However, its remaining components—variables, processes, and assumptions—often remain relevant at a more modest level of generality. The fatal flaw in the prime mover arguments is not that they rely on overarching laws or assumptions, but rather that they are applied at an inappropriate scale. Following the same logic, the optimization assumptions used heuristically in many HBE models should not be confused with a claim that economic optimization explains all transitions to agriculture (ref. 23, p. 682).

We agree that multiple broad-scale and local factors are involved in any agricultural transition (25, 26). However, this is a commonsense observation about the world, not a template for explanation (ref. 23, p. 681). In fact, complex phenomena such as OA must be simplified to be understood. A theoretical framework is needed to guide this process to ensure that explanations are not fashioned from a hodgepodge of factors and variables selected at the discretion of the researcher. For this reason, we advocate an approach to understanding OA that examines theoretical assumptions before interpreting data. Such assumptions always exist, acknowledged or not. Brought to the forefront, they serve to anchor observations to bodies of knowledge that have withstood repeated testing. Such high-level bodies of general theory inform middle- and lower-level theories that in turn generate testable hypotheses (16, 28). The explanatory relevance of broadly applicable principles does not conflict with the observation that domestication of any plant or animal is a particular, local, and historically contingent process. If it did, evolutionary theory could not be said to explain the diversity of life without implying that all evolutionary lineages follow the same path.

To be clear, we applaud the attention being paid to accurate reconstructions of the trajectory and context of agricultural transitions at regional and local scales. Detail-oriented analysis of observations and data yield rich

historical accounts of OA (see, for example, refs. 25 and 29–32). However, to the extent that they purport to engage in explanation, such accounts often rely on inductions that are expected to stand on the strength of available evidence alone. Although we recognize that the prioritization of data is often intentional and explicitly stated (e.g., ref. 33), we disagree that it is a good idea to abandon the iterative process of hypothesis testing, revision, and retesting that drives incremental advances in scientific understanding.

Particularistic Explanations Lack a General Theory for the Evolution of Human Behavior Yet Rely on a General Theory of Evolution for All Other Organisms.

There is little disagreement that change in the frequency of attributes in populations of humans, animals and plants under domestication is inherently coevolutionary. Understanding this process requires a general theory that links change in one population to change in another. Certainly, humans possess a unique mode of inheritance and an unusual degree of behavioral flexibility. However, cultural inheritance and the mechanics of human learning that enable it are constrained (at least, in the long run) by processes that constrain the evolution of other organisms. Far from being an obstacle to evolutionary analysis, human uniqueness offers one of the most important justifications for using it.

By not interpreting human behavior as a product of natural selection, particularistic approaches miss the opportunity to explore commonalities in human behavior that enable comparative analysis of agricultural origins in starkly different contexts. Nothing about an evolutionary perspective suggests that natural selection generates a predetermined template for human behavior, nor does it claim to explain everything that humans do. Human social learning has its own distinctive processes that influence patterns of behavior at the population level. However, human abilities (including the capacity for communication, learning, and choice—all components of cultural inheritance) are products of natural selection, and human culture is a function of those abilities (16, 34–36). Because agriculture has been adopted (or inherited) by nearly all human groups and because it was invented separately by many, the process will be understood best by looking at change in the history of human adaptation as a function of attributes shared by all humans, rather than looking at novel attributes unique to a few.

Because they lack a general theory for human behavior, particularistic approaches

rely on diverse and often conflicted principles to account for the decision-making behavior of human agents. In some cases, they retain the rational-actor optimization logic of HBE (e.g., that people will act in their own best interest with regards to survival-related utility); in other cases, they rely on creative or arbitrary combinations of a “bounded rationality” (sensu refs. 37 and 38) drawn opportunistically from ethnographic or contemporary observations; still others seem to gather principles of behavior from social theory (39, 40). Such principles are seldom made explicit; in some cases they are revealed by references to undefined or vaguely conceptualized properties or goals [e.g., “resource-rich,” “intensification of resource manipulation,” “rich array of plant and animal species,” “enrichment,” “stable and sustainable,” “stress,” etc. (23)]. Such terms betray embedded assumptions about the ecological situations that humans prefer and presumably strive for. Consistent theory that illustrates how and why these organizing principles emerge in the first place is essential for explaining the outcomes of diverse motivations in the distant past, particularly when these outcomes are visible today only in the aggregate records of archaeology and population biology.

Researchers who reject general theories of human behavior and cultural evolution as overly deterministic sometimes invoke human agency as an alternative causal mechanism. Agency is a contested term in archaeology, and the OA literature seems to have adopted one sense of it to the exclusion of others. Advocates use the common reading of agency (refs. 15 and 41, p. 67) to illustrate how individual choice and creativity explain cultural change, including the transition to agriculture (ref. 29, p. 44). Although we acknowledge the role of human creativity and innovation in OA, we are not convinced that appeals to agency have much explanatory power, simply because they sidestep the problem of identifying the evolutionary basis of the beliefs and preferences that guide human action (42). In contrast, HBE generates hypotheses about why humans behave the way they do and evaluates whether or not these hypotheses about agency explain the material remains of those behaviors. We agree that human agency is a critical component of evolutionary change, but we do not see it as having an equally important role in the scientific explanation of change.

Another problem with explanation by agency is that it often implies that decision makers have complete information about the long-term consequences of their actions. This feat would require far more involved data

analysis than the simple arithmetic required to find optimal solutions to foraging problems (for example) and is even farther removed from the practical heuristics (e.g., rules of thumb, tacit or explicit learning, rational calculation, etc.) that people actually use to make those decisions. By contrast, simple agency in foraging models is one of individual decision making, but does not imply (and does not require) long-term teleological self-determination.

The Evolutionary Approach Is a Robust, Flexible, and Practical Avenue for an Incrementally Better Understanding of Human Behavior and Cultural Change

Whereas evolutionary biologists seldom have to confront the question of human exceptionalism, the issue is a central one in anthropology and archaeology. The application of evolutionary theory to human behavior has been criticized for excessive adaptationism, reductionism, positivism, and materialism (e.g., refs. 43 and 44). These criticisms have been taken up by many advocates of interpretive archaeology, who question the exclusive claim of science to the production of useful knowledge (many examples appear in ref. 45; see also ref. 24).

Some archaeologists who do embrace evolutionary concepts in their OA research remain uncomfortable with their application to human behavior. Particularly heavy criticism has been leveled at evolutionary ecology (EE), defined as the “study of evolution and adaptive design in ecological context” (ref. 46, p. 3), especially the subdiscipline of HBE and its optimal foraging models. These critiques target either the basic assumptions of EE (many of which are shared by other evolutionary approaches) or elements of HBE or OFT specifically (14, 22, 23). Although the objections are numerous, they commonly revolve around the following themes: (i) explanations based on evolution by natural selection are “covering law” explanations that fail to account for the particular historical circumstances of their social and biological contexts; (ii) simple models are inappropriate for complex phenomena; (iii) the use of multiple simplifying models is compromised when one or more models fail to explain or predict the phenomena at hand; (iv) models do not explain human subsistence (or any other) behavior; (v) explanatory, predictive models deny human agency and are therefore dehumanizing; and (vi) some models do not account for environmental change and are therefore inappropriate for historical or evolutionary analysis. We find these critiques unwarranted and explain our reasoning below.

The Study of Evolution Is a Historical Science. Evolutionary theory is not covering law science (ref. 23, p. 681). It is a scientific approach to history (47) that views change as both cause and consequence of the immediate and long-term fitness implications of inherited and acquired variation. Although focused on the processes of change, nothing about these processes, or the analysis of them, is deterministic, teleological, or scripted. The context and consequences of one case might be completely different from the next (i.e., both historical and particular), but the foundations of evolutionary analysis are unifying. At the same time, it is flexible enough to deal with the hyper-variability of human behavior, which, for a variety of evolutionarily sound reasons, evolves and diversifies rapidly (48). In part, this flexibility comes from the application of a wide range of models that isolate specific aspects of behavioral and evolutionary phenomena.

Simple Models of Complex Phenomena Are Useful Even When They Fail.

Recent reactions to the use of OFT’s simplifying models in hunter-gatherer and OA research (14, 49, 50) take their cue from older critiques in biology and ecology (e.g., refs. 51 and 52) but differ on the issue of culture. For many social science scholars, models that isolate specific aspects of individual behavior in a *ceteris paribus* world (53) are impractical because (i) the *ceteris paribus* world doesn’t exist, and (ii) individuals perform in a complex, indivisible milieu that can be understood only in its holistic totality. For many, the practice of methodological isolation and analytical simplification has “reduced human actors to disemboweled humans who no longer have cultural anchors” (ref. 54, p. 59). They complain that evolutionary models isolate individuals from populations and remove both from the thing that binds them, which is culture.

Reservations about heuristic models are welcome (indeed these reservations and the resulting critiques may lead to better models); outright rejection of them is not. Unfortunately, and in many cases, rejection of simple, heuristic models is based on an idealized (but often unstated and unsupported) view of how the world works and on the complaint that too many people take the modeled predictions as fact, with or without empirical backing (14). Such critiques mistakenly conclude that the optimization assumptions used heuristically in many HBE models constitute a claim that optimal resource use explains all transitions to agriculture (ref. 23, p. 682).

The use of simple models (like those used in OFT) does not imply that the predictions of the models must be true. All things being equal (in the *ceteris paribus* world), the models should explain the behavior of organisms foraging optimally to maximize reproductive success. In the absence of data, the modeled predictions are simply that, and if the assumptions are correct, the predictions simply ought to be true. When data do exist but do not fit the model, either all things are not equal (e.g., the assumptions are incorrect), the model is inappropriately applied (e.g., the scale or resolution of the behaviors does not match those of the resulting data and/or the currency of the model is inappropriate), or the organism simply doesn't conform to expectation. Observations of the final type can also be productive by leading to further hypothesis testing built on related versions of EE developed specifically to explain the proliferation and persistence of suboptimal or maladaptive behavior (17, 55–58).

The criticism that HBE models sometimes fail to yield accurate predictions when tested misses the point that the exploration of the model's vulnerability can itself be a source of insight. Models are productive means for "eliminating problematic answers and identifying and pursuing more promising ones" (ref. 59, p. 171). In other words, failure is an option. For example, when empirical tests of the diet breadth model (DBM) failed to show a good fit of data to model predictions for agricultural transitions in the eastern United States, the findings led to a better grasp of the variables and tradeoffs (such as travel and processing costs) that most influenced subsistence strategies (e.g., refs. 60 and 61). Piperno and Pearsall (62) also discussed how OFT models, whose predictions conformed to Neotropical empirical data on the emergence of food production, were unlikely to address subsequent agricultural intensification satisfactorily, which evoked other explanations. Others (63–65) have investigated the failure of simple foraging models such as the DBM to capture important influences on agricultural decision making, such as yield variability and labor investment. These studies counter claims that applications of OFT demand idealistic purism or seek perfection in model/data fit (14, 50).

Simple, well-used models of OFT like the DBM (66) outline the logic of individual, time-sensitive decisions, and, because individuals make decisions in different contexts, the model anticipates behavioral variation. The extent of this variation in time and space is exactly what we find interesting, and relevant

to OA research. Aggregate patterns of individual behavior are what we see in the archives of the past. Therefore, the study of the past requires evaluation of models against data aggregated over different scales. The DBM is one model that makes this procedure possible. It is not the only model, and there is no reason to expect it to explain all data on all scales of time or space. However, in combination with related models, it is a very powerful analytical tool.

A Challenge to One Model Does Not Negate the Others. One critique of the evolutionary program is that it shifts arbitrarily between models that isolate and address different aspects of phenomena staged at different scales of time, space, or number (e.g., ref. 67). The observation is fair, but the critique is not. Indeed, what makes EE so versatile is that it draws from a "family of models" to address the tradeoffs between generality, realism, and precision inherent to any approach to scientific inquiry (68–70). The important point here is that we draw on models rooted in a unified body of theory, even if some of the models fail to explain the existing data, or if the models explain the data on different scales or levels of analysis.

Models of EE Often Generate Accurate Predictions of Human Subsistence Behavior, Including OA. As previously noted, OFT is one class of models with a reasonable purchase in the explanatory realm, and one that has provided a number of insights for biologists and archaeologists. A review of the past few years of biological literature finds persistent employment of OFT models, with empirical data often showing good fits with their predictions in a variety of organisms (e.g., refs. 71–74). OFT applications in human research also continue apace with valuable insights on varied topics (e.g., refs. 75–78).

A particularly important and well-studied example of subsistence change associated with OA and one that has been a focus of OFT applications is the "broad spectrum revolution" (BSR) (79). The BSR describes the expansion of human diets and the intensification of plant and animal exploitation that preceded agricultural beginnings in many parts of the world. Investigations of the BSR with OFT typically use the diet breadth model (DBM) and have shown a good fit of model predictions with empirical data from archaeological sites in southwest Asia, the Neotropics, and elsewhere (e.g., refs. 6, 62, and 80–87). Research shows that dietary shifts from higher- to lower-ranked plant and animal resources, or a more even dependence

on small-sized dietary items, occurred as a result of climate- and/or human-driven factors, depending on the region involved. Examples include terminal Pleistocene faunal extinctions and vegetational change (Neotropics) and increasing human demographic pressure leading to resource depression (southwest Asia). It is important to reiterate (88) that ethnographic datasets on foraging efficiency from around the world support the ranking of general categories of resources such as small and large game according to their energetic return rates (net energy obtained per unit time spent). Although exceptions occur, most plant foods and small and medium-sized mammals are ranked lower than most large game, and nuts and seeds are often the lowest-ranked of all (80, 88). It would be counterproductive for investigations of cultural developments through time, such as agricultural origins, to ignore those associations.

The collective archaeological data clearly underline how dietary diversification and intensification of plant use by foragers and incipient farmers probably occurred in an environmental milieu of shifting energetic return rates. Arguments that DBM predictions for the BSR are frequently invalidated by the archaeological data (14) are difficult to reconcile with the results of these studies. As a number of scholars have pointed out, the DBM is particularly suited for studying major directional changes in subsistence through time because of its ability to make robust, qualitative predictions of prey choice and dietary diversity. Using the model does not require that every past or present resource ranking set be in perfect order with respect to the general category of resource, package size, and the energetic return rates of individual items.

Ecological Models Accommodate Human Agency. Rejection of models like the DBM in favor of explanation by agency echoes a long-standing debate in evolutionary biology over the levels of causation (89–93). In behavioral science, ultimate explanations view behavior as a product of natural selection (cultural and biological) whereas proximate explanations attempt to understand how individuals identify challenges, make decisions, and initiate action (this being the realm of agency). The former have been caricatured as evolutionary "why" questions whereas the latter have been equated with questions of "how." Many see the conflation of these issues as a general challenge to evolutionary theory (93) whereas others recognize that both are important to understanding the structure

of biological, cultural, and socio-ecological change (42, 92).

HBE models, specifically the contingency models such as the DBM, seek ultimate explanations by identifying optimal solutions to immediate problems. They assume that organisms possess an evolved general capacity to achieve these optimal solutions, at least on average. Although it is true that models like the DBM rest heavily on ultimate rather than proximate explanation, this trait is no reason to reject them. Indeed, users of the model regularly acknowledge that it is silent about proximate causes while recognizing that many such causes exist. Both proximate and ultimate causes shape evolution, and the study of evolution must be able to identify where and when the feedbacks between “hows” and “whys” are reciprocal (92). The HBE modeling format makes this step possible by being agnostic about proximate causes, and testing hypotheses premised on ultimate causation. This agnosticism does not deny human agency but notes that agency, practice, and habitus (for example) are ultimately constrained by natural selection, at least over time. Because the models are not absolute covering-law statements, they can accommodate feedbacks between proximate and ultimate causes when, for example, natural selection (the ultimate cause) shapes the way people make decisions (the proximate cause), and those decisions further shape the environment of selection.

Whether or not the domestication of plants, animals, people, or landscapes entails morphological, genetic, or associational change, by design (i.e., “agency”) or by accident, abruptly or gradually, understanding the process on the human side of things requires an archaeology that tracks change in human behavior. Archaeological data typically say little about individuals but speak volumes about individuals in aggregate. HBE provides the theoretical foundation for understanding aggregate patterns as the product of individual decisions. Indeed, archaeological data are best viewed as the products of traits most common across a population of individuals; change in the spatial and temporal patterns of archaeological data reflects change in the distribution or frequency of traits within and among populations, through space and time.

Ecological Models Accommodate Environmental Change. Some critics of OFT suggest that it assumes an unchanging environment (refs. 14 and 50, p. 257). They argue that OFT models do not account for change in the environment of selection that creates domestic plants and animals and suggest that they should be discarded in favor of

explanations drawn from niche construction theory (NCT) (94). NCT highlights the multiple kinds of feedback that flow to the niche constructor and to other populations whose environments its niche-constructing activity influences (e.g., ref. 95). In doing so, NCT makes it possible to better characterize the ecological context in which optimality must be assessed. That hunter-gatherers did (and do) modify landscapes and resources for calculated gain is inarguable (96–98); however, the extent and scope of this practice are a matter of considerable debate (49). In the case of domestication and agriculture, landscape modification has long been recognized as an important phenomenon (99–101). However, we do not agree with the contention that OFT models cannot accommodate dynamic environments and argue that EE is well-positioned to support and complement NCT.

Although environmental stasis might be necessary to some aspects of model building (e.g., the contingency of the DBM assumes a static prey ranking), the simple models of foraging theory are equally applicable to dynamic environments (102–104). For example, the decision contingency of the DBM (66) is based on the marginal rate of return established by two important inputs: encounter rate (a function of prey abundance) and postencounter return rate (a function of prey value offset by the costs of acquiring and processing it). Data for the latter define the ranking of different prey, which is fixed until either the size (or total nutritional value) of the prey changes, or the energetics of capturing and processing it do. Thus, for each decision contingency there might be a different prey ranking, and, if so, the “optimal” diet in that instance may be different from during the previous instance. The standard logic of the DBM is that lower-ranked prey enter the diet as the abundance of higher-ranked prey declines; increasing abundance of lower-ranked prey has only a minimal effect on the decision to include them in the diet (105, 106). However, environmental modification may affect more than just absolute abundance and rates of encounter: change in the distribution of taxa also affects postencounter return rates that, in turn, affect the ranking of different prey taxa. Harvesting and replanting can raise the density of plants within a patch, which can alter the ranking of the resource (by reducing the costs of collecting the seeds) and boost its dietary importance (which may have been constrained initially by scarcity) (107). In other cases, selective, intensive harvests may reduce the body size of the average individual prey, again affecting the prey ranking and therefore the optimal decision.

EE offers a systematic way to evaluate the motivations for niche-construction activities such as landscape management and its intensification. In fact, applications of behavioral ecological models in archaeology often highlight the active role of humans in shaping their habitats (62, 80, 97, 108). Assertions to the contrary (50) neglect a substantial amount of literature on the subject. Critically, HBE permits an evaluation of the overall return rates associated with resource acquisition, which must include the costs of investing in land management as well as the returns of a managed landscape. EE (including HBE and OFT) can be useful for understanding the evolution of landscape management practices and other forms of niche construction and for evaluating the relationship between land, labor, and resource productivity.

We see a problem with using NCT as a “single unifying approach for integrating consideration of human and non-human modification of ecosystems” (ref. 50, p. 265) because it does not systematically address how or why humans alter their resource-acquisition patterns. There is little doubt that intensive human activity modifies ecosystems in diverse ways that nonetheless follow a limited set of pathways (50). However, a more intriguing question is why altered ecosystems engender or permit novel behavioral strategies in some cases but not others. Indeed, if hunter-gatherer landscape modification (including intentional enhancement of resource abundance) is a common feature of hunter-gatherer adaptive systems, why do so few cases generate the environment of selection that yields domestic plants and animals? Clearly, more is at work than the pervasiveness of niche construction: insights from NCT require the evolutionary logic of human decision making in tandem with a tighter consideration of the interactions between multiple taxa in novel settings to explain the origins of agriculture. Without a theory of organismal behavior, such as HBE, the alternatives imply either (*i*) that humans invented agriculture and its long-term consequences knowingly (as some advocates of human agency would have us believe), or (*ii*) that agriculture emerged almost randomly as a function of dynamic feedbacks between genes, environment, and the interactions between them (according to the strict version of NCT).

Integration of Evolutionary Frameworks

Evolutionary theory as applied to archaeological questions such as OA remains diverse, as expected for a relatively new but maturing field. However, the various approaches share

a commitment to evolutionary explanation even though they emphasize different processes, questions, and types of causation. EE (inclusive of HBE and OFT) (61, 64, 109), niche construction theory (19, 26), and models of cultural transmission and gene-culture coevolution (19, 110–112) all have the potential to link empirical findings to a well-established body of knowledge, but they target different implications of selection. OFT predicts how organisms might mobilize evolved decision-making mechanisms to maximize utility in the face of conflicting needs; niche construction theory asks how environmental engineering sets up pathways of ecological inheritance that affect the fitness of the organism and its descendants; cultural transmission theory asks how patterns of social learning peculiar to our species generate patterns of descent and divergence in cultural lineages.

Debates over the relative utility of these approaches have been helpful in distinguishing between fundamental incommensurabilities and superficial differences, thereby moving us closer to synthesis. We agree with Bettinger and Richerson that “Darwinian theory accommodates a virtually limitless range of interesting, fruitful projects. It is perhaps this diversity that causes us sometimes to forget the basic commonalities that unite the Darwinian enterprise”(ref. 112, p. 222). To declare that any of these theoretical perspectives is fatally flawed (14, 45) closes the door on potentially productive alliances among them.

For example, we believe that NCT can continue to move forward by joining forces with other evolutionary programs. Complementarity between NCT and EE has already been illustrated in case studies of New World hunter-gatherers and agriculturalists (108). Broughton et al. (ref. 108, p. 371) note that EE “focuses on how behavioral adjustments to changing socio-ecological conditions create novel selective pressures that in turn drive other changes in morphology and behavior” and therefore “draws on the same evolutionary logic that underlies niche construction theory.” Others (e.g., ref. 113) illustrate how NCT-based hypotheses can explain technological change across the transition to agriculture. O’Brien and Laland (19) make a strong case that gene-culture coevolution theory and NCT strengthen and support explanations of adaptive complexes related to agriculture, such as adult lactose tolerance and persistence of sickle-cell disease.

Conclusions

We believe strongly that the most significant archaeological contributions to our under-

standing of domestication will be those that acknowledge that humans are subject to the forces of evolution. We find little justification for the argument that, of all species, only *Homo sapiens* is capable of transcending these forces. We also advocate the use of theoretical frameworks that serve to link testable hypotheses to ultimate explanations for human behavior. However, we also wish to make it clear that, in doing so, we are not advocating any approach that ignores the distinctiveness of cultural evolution or the great behavioral plasticity that characterizes human interactions with the natural world. Nor are we in favor of glossing over historical particulars in the interests of advancing general theories of change.

Calls to purge OA research of HBE or OFT fail to distinguish between levels of theory and their respective epistemological roles. In OA research, certain core elements of Darwinian evolutionary theory (as advanced and amended by the neo-Darwinian

synthesis and, more recently, evolutionary and ecological developmental biology) have earned the status of basic assumptions. The reality of evolution by natural selection is such a core idea, arguably the most important one for OA research. It unifies approaches that use different analytic tools and methods and target different kinds of questions.

The research programs that cluster under the umbrella of evolutionary theory have generated, and continue to generate, testable hypotheses that contribute to our understanding of OA. Of these, HBE and OFT have a longstanding record of productivity, and, more recently, models of cultural inheritance have begun to show considerable promise. Although there is no way of predicting which evolutionary perspectives will prove most useful in the long run, one thing is certain: a retreat from theory in favor of particularism will ensure that none of them achieve their full potential.

- 1 Denham TP, et al. (2003) Origins of agriculture at Kuk Swamp in the highlands of New Guinea. *Science* 301(5630):189–193.
- 2 Denham T (2011) Early agriculture and plant domestication in New Guinea and island Southeast Asia. *Curr Anthropol* 52(Suppl 4): S379–S395.
- 3 Piperno DR, Ranere AJ, Holst I, Iriarte J, Dickau R (2009) Starch grain and phytolith evidence for early ninth millennium B.P. maize from the Central Balsas River Valley, Mexico. *Proc Natl Acad Sci USA* 106(13):5019–5024.
- 4 Dillehay TD, Rossen J, Andres TC, Williams DE (2007) Preceramic adoption of peanut, squash, and cotton in northern Peru. *Science* 316(5833):1890–1893.
- 5 Iriarte J (2007) New perspectives on plant domestication and the development of agriculture in the new world. *Rethinking Agriculture: Archaeological and Ethnoarchaeological Perspectives*, eds Denham T, Iriarte J, Vrydaghs L (Left Coast Press, Walnut Creek, CA), pp 167–188.
- 6 Piperno DR (2011) The origins of plant cultivation and domestication in the New World tropics: Patterns, process, and new developments. *Curr Anthropol* 52(Suppl 4):S453–S469.
- 7 Iriarte J (2009) Narrowing the gap exploring the diversity of early food-production economies in the Americas. *Curr Anthropol* 50(5): 677–680.
- 8 Barton L, et al. (2009) Agricultural origins and the isotopic identity of domestication in northern China. *Proc Natl Acad Sci USA* 106(14): 5523–5528.
- 9 Zhao Z (2011) New archaeobotanic data for the study of the origins of agriculture in China. *Curr Anthropol* 52(Suppl 4): S295–S306.
- 10 Fuller DQ, Harvey E, Qin L (2007) Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* 81:316–331.
- 11 Fuller DQ, et al. (2009) The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323(5921):1607–1610.
- 12 Bettinger RL, Barton L, Morgan C (2010) The origins of food production in North China: A different kind of agricultural revolution. *Evol Anthropol* 19(1):9–21.
- 13 Bettinger RL, et al. (2010) The transition to agriculture at Dadiwan, People’s Republic of China. *Curr Anthropol* 51(5):703–714.
- 14 Zeder MA (2012) The Broad Spectrum Revolution at 40: Resource diversity, intensification, and an alternative to optimal foraging explanations. *J Anthropol Archaeol* 31:241–264.
- 15 Hegmon M (2003) Setting theoretical egos aside: Issues and theory in North American archaeology. *Am Antiq* 68:213–243.
- 16 Ellen R (2010) Theories in anthropology and “anthropological theory”. *J R Anthropol Inst* 16:387–404.
- 17 Boyd R, Richerson PJ (1985) *Culture and the Evolutionary Process* (Univ of Chicago Press, Chicago).
- 18 Mesoudi A, Whiten A, Laland KN (2006) Towards a unified science of cultural evolution. *Behav Brain Sci* 29(4):329–347, discussion 347–383.
- 19 O’Brien MJ, Laland KN (2012) Genes, culture, and agriculture: An example of human niche construction. *Curr Anthropol* 53:434–470.
- 20 Richerson PJ, Boyd R, Henrich J (2010) Colloquium paper: Gene-culture coevolution in the age of genomics. *Proc Natl Acad Sci USA* 107(Suppl 2):8985–8992.
- 21 Harris M (1968) *The Rise of Anthropological Theory* (Thomas Crowell, New York).
- 22 Smith BD (2009) Core conceptual flaws in human behavioral ecology. *Commun Integr Biol* 6:533–534.
- 23 Zeder MA, Smith BD (2009) A conversation on agricultural origins: Talking past each other in a crowded room. *Curr Anthropol* 50:681–691.
- 24 Vrydaghs L, Denham T (2007) Rethinking agriculture: Introductory thoughts. *Rethinking Agriculture: Archaeological and Ethnoarchaeological Perspectives*, eds Denham T, Iriarte J, Vrydaghs L (Left Coast Press, Walnut Creek, CA), pp 1–15.
- 25 Zeder MA (2006) Central questions in the domestication of plants and animals. *Evol Anthropol* 15:105–117.
- 26 Smith BD (2007) Niche construction and the behavioral context of plant and animal domestication. *Evol Anthropol* 16:188–199.
- 27 Smith BD (2001) Documenting plant domestication: The consilience of biological and archaeological approaches. *Proc Natl Acad Sci USA* 98(4):1324–1326.
- 28 Ketelaar T (2003) The evaluation of competing approaches within human evolutionary psychology. *Evolutionary Psychology: Alternative Approaches*, eds Scher SJ, Rauscher F (Kluwer, Boston), pp 31–54.
- 29 Zeder MA (2008) The Neolithic macro-(r)evolution: Macroevolutionary theory and the study of culture change. *J Archaeol Res* 17(1):1–63.
- 30 Zeder MA (2008) Domestication and early agriculture in the Mediterranean Basin: Origins, diffusion, and impact. *Proc Natl Acad Sci USA* 105(33):11597–11604.
- 31 Bar-Yosef O (2011) Climatic fluctuations and early farming in West and East Asia. *Curr Anthropol* 52(Suppl 4):S175–S193.
- 32 Belfer-Cohen A, Goring-Morris AN (2011) Becoming farmers: The inside story. *Curr Anthropol* 52(Suppl 4):S209–S220.
- 33 Belfer-Cohen A, Goring-Morris AN (2009) For the first time. *Curr Anthropol* 50:669–671.
- 34 Flinn M (1997) Culture and the evolution of social learning. *Evol Hum Behav* 18:23–67.
- 35 Henrich J, McElreath R (2003) The evolution of cultural evolution. *Evol Anthropol* 12:123–135.
- 36 Nettle D (2009) Beyond nature versus culture: Cultural variation as an evolved characteristic. *J R Anthropol Inst* 15:223–240.
- 37 Jones BD (1999) Bounded rationality. *Annu Rev Polit Sci* 2:297–321.

- 38 Simon HA (1947) *Administrative Behavior* (Macmillan, New York).
- 39 Hodder I (1990) *The Domestication of Europe: Structure and Contingency in Neolithic Societies* (B. Blackwell, Oxford).
- 40 Cauvin J (2000) *The Birth of the Gods and the Origins of Agriculture*, trans Watkins T (Cambridge Univ Press, Cambridge, UK).
- 41 Gardner A (2011) Action and structure in interpretive archaeologies. *Evolutionary and Interpretive Archaeologies: A Dialogue*, eds Cochran EE, Gardner A (Left Coast Press, Walnut Creek, CA), pp 63–82.
- 42 Smith EA (2013) Agency and adaptation. *Annu Rev Anthropol* 42:103–120.
- 43 Ingold T (2000) *The Perception of the Environment: Essays in Livelihood, Dwelling and Skill* (Routledge, London).
- 44 Joseph S (2000) Anthropological evolutionary ecology: A critique. *J Ecol Anthropol* 4:6–30.
- 45 Cochran EE, Gardner A, eds (2011) *Evolutionary and Interpretive Archaeologies: A Dialogue* (Left Coast Press, Walnut Creek, CA).
- 46 Winterhalder B, Smith EA (1992) Evolutionary ecology and the social sciences. *Evolutionary Ecology and Human Behavior*, eds Smith EA, Winterhalder B (Aldine de Gruyter, New York), pp 3–24.
- 47 Nitecki MH, Nitecki DV (1992) *History and Evolution* (State Univ of New York Press, Albany, NY).
- 48 Richerson PJ, Boyd R (1997) Built for speed, not for comfort: Darwinian theory and human culture. *Philosophica* 60:13–43.
- 49 Lightfoot KG, Cuthrell RQ, Striplen CJ, Hylkema MG (2013) Rethinking the study of landscape management practices among hunter-gatherers in North America. *Am Antiq* 78(2):285–301.
- 50 Smith BD (2011) A cultural niche construction theory of initial domestication. *Biol Theory* 6:260–271.
- 51 Gray RD (1987) Faith and foraging: A critique of the “paradigm argument from design”. *Foraging Behavior*, eds Kamil AC, Krebs JR, Pulliam HR (Plenum, New York), pp 69–140.
- 52 Pierce GJ, Ollason JG (1987) Eight reasons why optimal foraging theory is a complete waste of time. *Oikos* 49:111–118.
- 53 Boyer P (1995) *Ceteris paribus* (all else being equal). *How Things Are: A Science Tool-Kit for the Mind*, eds Brockman J, Matson K (William Morrow, New York), pp 169–175.
- 54 Yengoyan AA (2004) *Anthropological History and the Study of Hunters and Gatherers: Cultural and Non-Cultural Hunter-Gatherers in History, Archaeology and Anthropology*, ed Barnard A, Berg, New York), pp 57–66.
- 55 Henrich J (2002) Decision making, cultural transmission, and adaptation in economic anthropology. *Theory in Economic Anthropology*, ed Ensminger J (Alta Mira, Walnut Creek, CA), pp 251–295.
- 56 Henrich J (2004) Demography and cultural evolution: How adaptive cultural processes can produce maladaptive losses—the Tasmanian case. *Am Antiq* 69(2):197–214.
- 57 McElreath R (2004) Social learning and the maintenance of cultural variation: An evolutionary model and data from East Africa. *Am Anthropol* 106(2):308–321.
- 58 Richerson PJ, Boyd R (2005) *Not by Genes Alone: How Culture Transformed Human Evolution* (Univ of Chicago Press, Chicago).
- 59 Bird D, O’Connell JF (2006) Behavioral ecology and archaeology. *J Archaeol Res* 14:143–188.
- 60 Gremillion KJ (2004) Seed processing and the origins of food production in eastern North America. *Am Antiq* 69:215–234.
- 61 Gremillion KJ (1998) Changing roles of wild and cultivated plant resources among early farmers of Eastern Kentucky. *Southeastern Archaeol* 17:140–157.
- 62 Piperno DR, Pearsall DM (1998) *The Origins of Agriculture in the Lowland Neotropics* (Academic, San Diego).
- 63 Barlow KR (2006) A formal model for predicting agriculture among the Fremont. *Behavioral Ecology and the Transition to Agriculture*, eds Kennett DJ, Winterhalder B (Univ of California Press, Berkeley, CA), pp 87–102.
- 64 Marston JM (2011) Archaeological markers of agricultural risk management. *J Anthropol Archaeol* 30:190–205.
- 65 Foster HT II (2003) Dynamic optimization of horticulture among the Muscogee Creek Indians of the southeastern United States. *J Anthropol Archaeol* 22:411–424.
- 66 MacArthur RH, Pianka ER (1966) On optimal use of a patchy environment. *Am Nat* 100:603–609.
- 67 Smith BD (2006) Human behavioral ecology and the transition to food production. *Behavioral Ecology and the Transition to Agriculture*, eds Kennett DJ, Winterhalder B (Univ of California Press, Berkeley, CA), pp 289–303.
- 68 Levins R (1966) The strategy of model-building in population biology. *Am Sci* 54:421–431.
- 69 Richerson PJ, Boyd R (1987) Simple models of complex phenomena: The case of cultural evolution. *The Latest on the Best: Essays on Evolution and Optimality*, ed Dupre J (MIT Press, Cambridge, MA), pp 27–52.
- 70 Winterhalder B (2002) Models. *Darwin and Archaeology: A Handbook of Key Concepts*, eds Hart JP, Terrell JE (Bergin and Garvey, Westport, CT), pp 201–224.
- 71 Catania KC (2012) Evolution of brains and behavior for optimal foraging: A tale of two predators. *Proc Natl Acad Sci USA* 109(Suppl 1):10701–10708.
- 72 Lai Y, Chen J-H, Lee L-L (2011) Prey selection of shell-invading leech as predicted by optimal foraging theory with consumption success incorporated into estimation of prey profitability. *Funct Ecol* 25:147–157.
- 73 Danial-Valcroze T, Lesage V, Giard J, Michaud R (2011) Optimal foraging theory predicts diving and feeding strategies of the largest marine predator. *Behav Ecol* 22:880–888.
- 74 Sayers K, Norconk MA, Conklin-Brittain NL (2010) Optimal foraging on the roof of the world: Himalayan langurs and the classical prey model. *Am J Phys Anthropol* 141(3):337–357.
- 75 Levi T, Lu F, Tu DS, Mangel M (2011) The behaviour and diet breadth of central-place foragers: An application of human hunters and Neotropical game management. *Evol Ecol Res* 13:171–185.
- 76 Lupo KD, Fancher JM, Schmitt DN (2012) The taphonomy of resource intensification: Zooarchaeological implications of resource scarcity among Bofi and Aka forest foragers. *J Archaeol Method Theory* 20:420–447.
- 77 Dwairy M, Dowell AC, Stahl J-C (2011) The application of foraging theory to the information searching behaviour of general practitioners. *BMC Fam Pract* 12:90–98.
- 78 Robinson ME, McKillop HI (2013) Ancient Maya wood selection and forest exploitation: A view from the Paynes Creek salt works, Belize. *J Archaeol Sci* 10:3584–3595.
- 79 Flannery KV (1969) Origins and ecological effects of early domestication in Iran and the Near East. *The Domestication and Exploitation of Plants and Animals*, eds Ucko PJ, Dimbleby GW (Aldine, Chicago), pp 73–100.
- 80 Piperno DR (2006) The origins of plant cultivation and domestication in the Neotropics: A behavioral ecological perspective. *Behavioral Ecology and the Transition to Agriculture*, eds Kennett D, Winterhalder B (Univ of California Press, Berkeley, CA), pp 137–166.
- 81 Stiner MC, Munro ND, Surovell TA (2000) The tortoise and the hare: Small game use, the broad spectrum revolution, and Paleolithic demography. *Curr Anthropol* 41(1):39–79.
- 82 Stiner MC, Munro DN (2002) Approaches to prehistoric diet breadth, demography, and prey ranking systems in time and space. *J Archaeol Method Theory* 9:181–214.
- 83 Jones EL (2009) Climate change, patch choice, and intensification at Pont d’Ambon (Dordogne, France) during the Younger Dryas. *Quat Res* 72:371–376.
- 84 Munro ND (2009) Epipaleolithic subsistence intensification in the southern Levant: The faunal evidence. *Integrating Approaches to the Study of Paleolithic Subsistence*, eds Hublin J-J, Richards MJ (Springer, The Hague), pp 141–155.
- 85 Munro ND (2009) Integrating inter- and intra-site analyses of Epipaleolithic faunal assemblages from Israel. *Before Farming: Archaeol Anthropol Hunter-Gatherers*. 2009(1):4.
- 86 Stutz AJ, Munro ND, Bar-Oz G (2009) Increasing the resolution of the Broad Spectrum Revolution in the Southern Levantine Epipaleolithic (19–12 ka). *J Hum Evol* 56(3):294–306.
- 87 Starkovich BM, Stiner MC (2010) Upper Palaeolithic animal exploitation at Klissoura Cave 1 in southern Greece: Dietary trends and mammal taphonomy. *Eurasian Prehistory* 7:107–132.
- 88 Ugan A, Simms S (2012) On prey mobility, prey rank, and foraging goals. *Am Antiq* 77(1):179–185.
- 89 Mayr E (1961) Cause and effect in biology. *Science* 134(3489):1501–1506.
- 90 Tinbergen N (1963) On aims and methods of ethology. *Z Tierpsychol* 20:410–433.
- 91 West-Eberhard MJ (2003) *Developmental Plasticity and Evolution* (Oxford Univ Press, Oxford).
- 92 Laland KN, Sterelny K, Odling-Smee J, Hoppitt W, Uller T (2011) Cause and effect in biology revisited: Is Mayr’s proximate-ultimate dichotomy still useful? *Science* 334(6062):1512–1516.
- 93 Scott-Phillips TC, Dickens TE, West SA (2011) Evolutionary theory and the ultimate-proximate distinction in the human behavioral sciences. *Perspect Psychol Sci* 6:38–47.
- 94 Odling-Smee J, Laland KN, Feldman W (2003) *Niche Construction* (Princeton Univ Press, Princeton).
- 95 Odling-Smee J, Laland KN (2011) Ecological inheritance and cultural inheritance: What are they and how do they differ? *Biol Theory* 6(3):220–230.
- 96 Anderson MK (2005) *Tending the Wild: Native American knowledge and the management of California’s natural resources* (Univ of California Press, Berkeley, CA).
- 97 Blegie Bird R, Bird DW, Codding BF, Parker CH, Jones JH (2008) The “fire stick farming” hypothesis: Australian Aboriginal foraging strategies, biodiversity, and anthropogenic fire mosaics. *Proc Natl Acad Sci USA* 105(39):14796–14801.
- 98 Smith BD (2007) Behavior: The ultimate ecosystem engineers. *Science* 315(5820):1797–1798.
- 99 Anderson E (1952) *Plants, Man, and Life* (Univ of California Press, Berkeley, CA).
- 100 Anderson E (1956) Man as a maker of new plants and new plant communities. *Man’s Role in Changing the Face of the Earth*, ed Thomas WL (Univ of Chicago Press, Chicago), pp 763–777.
- 101 Lewis HT (1972) The role of fire in the domestication of plants and animals in Southwest Asia: A hypothesis. *Man (Lond)* 7(2):195–222.
- 102 Belovsky GE (1988) An optimal foraging-based model of hunter-gatherer population dynamics. *J Anthropol Archaeol* 7(4):329–372.
- 103 Winterhalder B (1986) Diet choice, risk, and food sharing in a stochastic environment. *J Anthropol Archaeol* 5:369–392.
- 104 Winterhalder B, et al. (1988) The populations ecology of Hunter-Gatherers and their prey. *J Anthropol Archaeol* 7:289–328.
- 105 Pyke GH, Pulliam HR, Charnov EL (1977) Optimal foraging: A selective review of theory and tests. *Q Rev Biol* 52:137–154.
- 106 Winterhalder B (1981) Optimal foraging strategies and hunter-gatherer research in anthropology: Theory and methods. *Hunter-Gatherer Foraging Strategies: Ethnographic and Archeological Analyses*, eds Winterhalder B, Smith EA (Univ of Chicago Press, Chicago), pp 13–35.
- 107 Winterhalder B, Goland C (1997) An evolutionary ecology perspective on diet choice, risk, and plant domestication. *People, Plants, and Landscapes: Studies in Paleoethnobotany*, ed Gremillion KJ (Univ of Alabama Press, Tuscaloosa, AL), pp 123–160.
- 108 Broughton JM, Cannon MD, Bartelink EJ (2010) Evolutionary ecology, resource depression, and niche construction theory: Applications to central California hunter-gatherers and Mimbres-Mogollon agriculturalists. *J Archaeol Method Theory* 17:371–421.
- 109 Bettinger RL (2006) Agriculture, archaeology, and human behavioral ecology. *Behavioral Ecology and the Transition to Agriculture*, eds Kennett DJ, Winterhalder B (Univ of California Press, Berkeley, CA), pp 304–322.
- 110 Richerson PJ, Boyd R, Bettinger RL (2001) Was agriculture impossible during the Pleistocene but mandatory during the Holocene? A climate change hypothesis. *Am Antiq* 66:387–412.
- 111 Coward F, Shennan S, Colledge S, Conolly J, Collard M (2008) The spread of Neolithic plant economies from the Near East to northwest Europe: A phylogenetic analysis. *J Archaeol Sci* 35:42–56.
- 112 Bettinger RL, Richerson PJ (1996) The state of evolutionary archaeology: Evolutionary correctness, or the search for common ground. *Darwinian Archaeologies*, ed Maschner HDG (Plenum, New York), pp 221–231.
- 113 Collard M, Buchanan B, Ruttle A, O’Brien MJ (2011) Niche construction and the toolkits of hunter-gatherers and food producers. *Biol Theory* 6:251–259.

Beyond the single gene: How epistasis and gene-by-environment effects influence crop domestication

Andrew N. Doust^{a,1,2}, Lewis Lukens^{b,1}, Kenneth M. Olsen^{c,1}, Margarita Mauro-Herrera^a, Ann Meyer^b, and Kimberly Rogers^a

^aBotany Department, Oklahoma State University, Stillwater, OK 74078; ^bDepartment of Plant Agriculture, University of Guelph, Guelph, Ontario, Canada N1G2W1; and ^cDepartment of Biology, Washington University, St. Louis, MO 63130

Edited by Greger Larson, Durham University, Durham, United Kingdom, and accepted by the Editorial Board October 11, 2013 (received for review July 16, 2013)

Domestication is a multifaceted evolutionary process, involving changes in individual genes, genetic interactions, and emergent phenotypes. There has been extensive discussion of the phenotypic characteristics of plant domestication, and recent research has started to identify the specific genes and mutational mechanisms that control domestication traits. However, there is an apparent disconnect between the simple genetic architecture described for many crop domestication traits, which should facilitate rapid phenotypic change under selection, and the slow rate of change reported from the archeobotanical record. A possible explanation involves the middle ground between individual genetic changes and their expression during development, where gene-by-gene (epistatic) and gene-by-environment interactions can modify the expression of phenotypes and opportunities for selection. These aspects of genetic architecture have the potential to significantly slow the speed of phenotypic evolution during crop domestication and improvement. Here we examine whether epistatic and gene-by-environment interactions have shaped how domestication traits have evolved. We review available evidence from the literature, and we analyze two domestication-related traits, shattering and flowering time, in a mapping population derived from a cross between domesticated foxtail millet and its wild progenitor. We find that compared with wild progenitor alleles, those favored during domestication often have large phenotypic effects and are relatively insensitive to genetic background and environmental effects. Consistent selection should thus be able to rapidly change traits during domestication. We conclude that if phenotypic evolution was slow during crop domestication, this is more likely due to cultural or historical factors than epistatic or environmental constraints.

QTL | genotype-by-environment interactions | $G \times E$ | *Setaria* | domestication syndrome

Domestication is a process of evolutionary change that leads to increased dependence and associated phenotypic modifications in both domesticator and domesticated. Much is known about the phenotypes that change during plant domestication, including the so-called domestication syndrome described from annual crops, which emphasizes retention of seed on the seed head, reduction in lateral branching, reduction in seed dormancy, and increase in seed size (1). Other phenotypic changes occurring during domestication and improvement, such as shifts in flowering time and grain composition, have also been extensively studied (2–4). Recent molecular research has begun to identify and functionally characterize a growing number of the major-effect domestication and crop improvement genes (5–8). However, there is disagreement over the speed of phenotypic evolution during domestication, with an apparent disconnect between the relatively simple genetic architecture of many domestication-related traits, which should facilitate rapid change, and the protracted length of time seen in the archeobotanical record for some domestication phenotypes to become widespread (9–12).

The genetics of domestication traits generally suggest that phenotypic changes during crop domestication could potentially

occur rapidly. Many domestication alleles segregate as Mendelian loci with large effects (7), and the rate of self-pollination in many crops is high, allowing for the expression of both recessive and dominant domestication alleles even when initially at low frequencies in populations. However, archeobotanical evidence from the major cereal crops rice and wheat (10, 13) and some population genetic simulations (9, 14) suggest that genetic modifications underlying phenotypic change may only slowly be translated into geographically widespread domestication phenotypes. Scenarios to explain this apparent discrepancy have focused on the strength of selection during domestication, mediated by the ability of early farmers to recognize and use favorable mutations, as well as population genetic processes such as introgression from wild populations (15). It is also possible that the vagaries of human history, including wars, epidemics, and other interruptions to food production (16), may have led to protracted domestication rates. There has been much debate as to whether the slow increase in domesticated phenotypes is real or a preservation artifact and what assumptions can be made about the botanical knowledge of early farmers (9–12, 17). Here we explore whether background effects such as gene-by-gene interactions (epistasis) and gene-by-environment ($G \times E$) effects can affect the efficacy of selection on domestication alleles. In theory, this middle ground between molecular genetic change and its expression during development could have significantly slowed the process of phenotypic evolution during crop domestication.

Epistasis refers to genetic interactions between loci, which may be either biallelic or higher order in nature. The term has been defined in a number of different ways [reviewed by Phillips

Significance

Recent archaeological studies of crop domestication have suggested a relatively slow spread and fixation of some key domestication traits, such as the loss of seed shattering. In contrast, genetic studies often indicate that domestication traits have a fairly simple genetic basis, which should facilitate their rapid evolution under selection. Here we examine previously underexplored factors that could account for this apparent disconnect: the roles of gene-by-gene interactions (epistasis) and gene-by-environment effects in shaping the rate of phenotypic evolution during domestication. Analysis of a *Setaria* mapping population, together with a review of evidence from the literature, suggests that these genetic factors, although important, are unlikely to have played a major role in constraining the rate of phenotypic evolution during domestication.

Author contributions: A.N.D., L.L., and K.M.O. designed research; A.N.D., M.M.-H., and K.R. performed research; A.N.D., L.L., K.M.O., M.M.-H., and A.M. analyzed data; and A.N.D., L.L., K.M.O., and M.M.-H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. G.L. is a guest editor invited by the Editorial Board.

¹A.N.D., L.L., and K.M.O. contributed equally to this work.

²To whom correspondence should be addressed. E-mail: andrew.doust@okstate.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1308940110/-DCSupplemental.

(18)]; here we restrict our use to describe the phenomenon whereby the phenotypic effects of a given allele at a gene can vary depending on the allelic composition of the other loci that make up an individual's genotype. This definition has been described as “compositional epistasis” (18) and is synonymous with “physiological epistasis” as used by Cheverud and Routman (19) and “functional epistasis” as used by Carter et al. (20). Epistasis has been shown to contribute to additive genetic variances and covariances, thereby affecting the response to selection; depending on the direction of these effects in morphospace, it may dramatically change the response to selection in only a few generations (20). In the context of domestication genetics, various outcomes are possible for any pair of loci, from complete lack of gene interaction, with only additive effects explaining the phenotype (Fig. 1A), to the effects of wild and domestication alleles differing across genetic backgrounds (Fig. 1B and C).

The geographical distribution and population structure of many crops and their wild relatives may predispose these species to epistatic effects. Both crop species and their wild progenitors are often characterized by substantial population genetic structure, with distinct genetic subpopulations that are geographically separated and evolutionarily diverged (21, 22). This is especially true for inbreeding species, where local adaptation is not easily homogenized by introgression from other populations and where local variants can readily become established and persist through time. If adaptive differentiation among populations has led to evolutionary divergence in gene regulatory networks, a domestication allele that was initially selected upon in one population could, in theory, have very different phenotypic effects when introgressed into a population growing in a different region.

G × E interactions refer to situations in which the relative effects of alleles change across environments. The visible morphological traits selected during domestication, such as branching, inflorescence development, and seed size, are strongly affected by the environment (23), and the phenotypic effects of favorable domestication alleles may have been difficult for early farmers to consistently detect. If we assume that early agriculturalists, like those today, sought to reduce the impacts of both predictable and unpredictable environmental variation on crops, a resulting expectation is that domestication alleles would be less sensitive to environmental changes than predomestication, ancestral alleles.

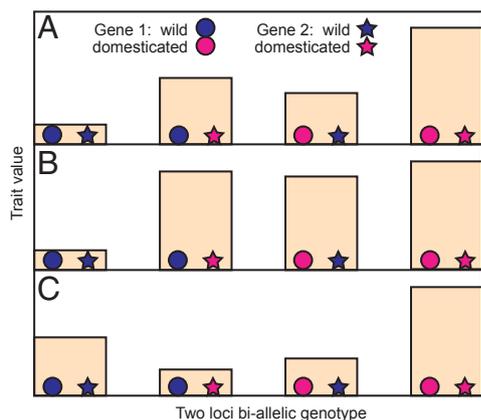


Fig. 1. Three sets of phenotypic outcomes that could result from biallelic epistatic interactions between two loci. Circles and stars represent genotypes homozygous for alleles that affect a domestication trait, where higher trait values are selectively favored during domestication. Pink genotypes have alleles from the domesticated population. Blue genotypes have alleles drawn from a population of close wild relatives. (A) No interaction. Additive effects of alleles explain all of the phenotypic variation. (B) Interaction such that domestication alleles at one or the other locus give trait values close to the domesticate double homozygote. (C) Less than additive interaction such that the wild-domesticate gene pair combinations have genotypic values lower than wild or domesticate double homozygotes.

Such phenotypic stability would facilitate selection, enabling the spread of domesticated crops away from their native environments into other climates and latitudes (24).

Here we investigate whether epistasis and/or G × E interactions could help explain the apparent contradiction between the genetic simplicity of many domestication-related traits (both domestication and improvement traits) and the slow rate of phenotypic progress observed for some crops in the archeobotanical record. We review studies of epistasis and show that a number of genes affecting domestication-related traits are epistatic in segregating populations containing wild and domestication alleles. In many cases, domestication alleles have more stable effects in diverse genetic backgrounds than do ancestral alleles, suggesting that selection for these “robust” alleles would be effective in creating rapid phenotypic change. Similarly, the effects of domestication alleles differ across environments but are often less sensitive to environmental variation than wild (predomestication) alleles. As a complement to the literature review, we also examine the effects of epistasis on shattering and the effects of epistasis and G × E on flowering time in a cross between domesticated foxtail millet (*Setaria italica*) and its wild progenitor, green millet (*Setaria viridis*). We find loci of major-effect and epistatic and G × E interactions that significantly explain phenotypic variation. Major-effect loci appear to be less influenced by genetic background than are minor-effect loci. We suggest that for many domestication-related traits, wild populations segregate for alleles that are sensitive to genetic background and environment but that the process of domestication has favored relatively insensitive loci that have large effects on traits and are highly responsive to selection. Thus, the simple genetic architecture observed for many domestication-related traits may be the evolutionary outcome of successful selection for alleles that are relatively insensitive to background effects.

Epistasis Affects Domestication and Crop Improvement Traits

Both domestication traits (i.e., traits that distinguish a crop from its wild ancestor) and crop improvement traits (i.e., agronomic traits that vary among landraces or varieties of a crop) can be affected by epistatic interactions, and epistasis has long been recognized as an important component in the genetic architecture of crop plants. An extreme example is demonstrated by the maize *opaque7* mutant, which produces opaque, soft, and floury kernels. In the inbred dent corn line in which it was discovered (W22), the recessive allele *o7* segregates in a simple Mendelian fashion (25). However, in testcrosses with other maize varieties grown in similar environments, F₂ and backcross populations segregate for the *opaque7* phenotype with greatly reduced frequencies (25). Among 139 genotypes tested, 118 showed significant deviations from the expected 3:1 segregation of normal to opaque kernels. A single F₂ population could have as few as 3.1% opaque kernels. This wide phenotypic variation thus suggests that the visibility of the *opaque7* phenotype for selection during crop breeding would strongly depend on the genetic background in which it arose.

A number of quantitative trait loci (QTL) analyses of domestication traits from biparental populations derived from wild × domesticated plant crosses have reported that the combined effect of two (or more) alleles significantly differs from the sum of the separate allelic effects. In a study of vegetative architecture in an F_{2:3} mapping population derived from a cross between domesticated foxtail millet and its wild ancestor green millet, 10 highly significant epistatic interactions between loci were found, with the proportion of variance explained ranging from 16% to 41% (26). Three interactions were between QTLs that had significant additive effects; two were between a QTL and a marker in another part of the genome, and the remaining five involved loci that showed no significant additive effects (26). Other studies have also identified epistatic interactions between loci that have no significant main effects and between QTL and other loci. A study on epistatic interactions controlling domestication and improvement traits in a backcrossed doubled haploid population of wild

barley (*Hordeum vulgare* ssp. *spontaneum*) and domesticated barley (*H. vulgare* L.) revealed that a significant proportion of phenotypic variation was explained by interactions between loci that did not show significant main effects (27). Such findings are disconcerting given that many studies have failed to consider epistatic effects altogether.

The impact that genetic interactions have on domestication traits can be most clearly observed in lines where wild alleles have been individually introgressed into the genetic background of a modern cultivar, followed by crossing of different introgressed lines. A particularly well characterized example of epistasis is that involving the *teosinte branched1* (*tb1*) locus in maize, which affects apical dominance and lateral inflorescence (ear) development (28, 29). These are key domestication traits that distinguish maize (*Zea mays* L. ssp. *mays*) from its wild ancestor, teosinte (*Z. mays* L. subsp. *parviglumis*), and there has been strong selection at this locus during maize domestication (30). The *tb1* locus, on chromosome 1, interacts with another locus on chromosome 3 to affect the sexuality of the inflorescences (31), and the allelic state of the chromosome 3 gene has a dramatic impact on *tb1* effect size. In teosinte, the lateral branches terminate in male inflorescences (tassels), whereas in maize they terminate in female inflorescences (ears). When maize is introgressed with teosinte alleles at either locus, both teosinte alleles have little effect on inflorescence sex, with either allele only changing ~20% of the spikelets from female to male. However, when both loci are homozygous for the teosinte alleles, the inflorescence produces up to 90% male spikelets (31). The maize *tb1* allele is known to segregate in teosinte populations (30), although these alleles are likely currently under strong negative selection (32). One would expect that rare teosinte plants homozygous for maize alleles would exhibit strongly feminized inflorescences.

The maize–teosinte *tb1* example illustrates a common theme in wild–domesticated systems, where the ancestral allele is sensitive to other modifying ancestral alleles but the domestication allele is insensitive. This pattern is well documented in other crop species such as barley, where there are two genes that control shattering, *Btr1* and *Btr2*. Wild alleles of both genes are required to confer shattering, but domestication alleles at either locus convert the inflorescence to nonshattering, regardless of the allelic state of the other locus (33–35) (Fig. 1B). Similarly, investigation of domestication-related traits in a wild × domesticated barley population found that for several traits controlled by interacting loci (heading date, plant height, and yield), the wild phenotype was only generated with the introgression of both wild alleles into an introgression line (27). Independence of a major domestication locus from its genetic background was also found by Gu et al. (36) for the *qSD12* locus that controls ~50% of the variation in seed dormancy in rice, whereas minor loci showed multiple epistatic interactions. Similar genetic architectures exist for vegetative architecture in maize and foxtail millet (26, 37) and shattering in sorghum, wheat, and rice (35, 38, 39).

When introduced into domesticated germplasm, epistatic interactions may mean that the effects of interacting pairs of wild alleles are less than the sum of the effects of single wild alleles (Fig. 1C). For example, working in tomato, Eshed and Zamir (40) reported frequent less-than-additive epistasis in which the effect of double heterozygotes of wild and domestication alleles was smaller than the sum of the effects of single heterozygotes. They suggest that such effects are indicative of interacting loci in the same genetic pathway. Von Korff et al. (27), in their investigation of heading date, plant height, and yield in a wild × domesticated barley population, also found less-than-additive epistatic effects.

Alleles that are relatively insensitive to genetic backgrounds within cultivated populations can regain sensitivity within wild populations. For example, the rice *sh4* locus accounts for a large proportion of shattering variation in populations derived from crosses between domesticated *Oryza sativa* and its wild progenitors, the annual species *O. nivara* (41) and perennial species *O. rufipogon* (42). However, in some genetic backgrounds, wild rice plants can be homozygous for the *sh4* domestication allele and yet still show

a shattering phenotype (43). Similarly, weedy rice strains in the United States are almost entirely fixed for the *sh4* domestication allele, yet are highly shattering (44). It is possible that novel modifiers of *sh4* have evolved within weedy rice.

Not all studies report significant epistatic interactions for domestication-related genes. In some cases, mapping studies do not explicitly examine epistatic interactions (e.g., ref. 42), or they only compare them between significant QTL (e.g., ref. 45). In QTL mapping approaches, first-order effects are commonly fitted before second- and higher-order (epistatic) effects, making it difficult to statistically detect epistasis (46). Even when epistasis is tested for among all possible pairs of loci (47), statistical and genetic factors can interfere with successful detection. For example, recent simulation studies have suggested that incomplete linkage disequilibrium (LD) between causal variants and observed SNPs will erode estimates of epistatic variance, leading to inflation of the size and relative importance of additive effects (48). This phenomenon may explain the scarcity of epistatic interactions found in association and multiparental mapping studies such as the nested association mapping (NAM) maize population studies on flowering time, inflorescence architecture, plant height, leaf architecture, and disease resistance (2, 49–52). In general, epistatic interactions have been reported more often in inbreeding species such as rice and *Arabidopsis* than in outcrossing species such as maize, consistent with the larger LD exhibited by inbreeding species (46). Another potential reason for the lack of significant epistatic interactions detected in the NAM maize studies may be the very high statistical stringency required to control for false positives, given the large number of pairwise marker comparisons required.

Environmental Interactions with Domestication Genes

As with alleles with stable effects across genetic backgrounds, alleles with consistent effects across environments are easier and more stable selection targets than those whose effects are sensitive to environmental variation. Domestication may thus have been marked by the selection of alleles with a lack of sensitivity to environmental variation. For example, teosinte is strongly affected by crowding, so that when a wild *tb1* allele from teosinte is introgressed into maize, it produces a highly branched phenotype in uncrowded growing conditions but results in repressed lateral branches resembling domesticated maize in high-density growing conditions (28). In contrast, a *tb1* domestication allele in maize produces an unbranched phenotype regardless of whether plants are grown in high- or low-density environments. The effect of a teosinte *tb1* allele on the percentage of staminate flowers also varies between densities, with plants at low densities having ~20% staminate flowers and those at high densities having only ~5% staminate flowers (31). This is in contrast to the complete insensitivity of the maize *tb1* allele, where plants grown under both densities produce no staminate flowers. The environmental sensitivity of the wild allele may have weakened selection by early agriculturalists when plants with either wild or domesticated alleles were grown together as a crop because the two alleles would have produced similar phenotypes when grown at high densities (31).

In maize plants homozygous with teosinte alleles of both *tb1* and the locus on chromosome 3 (see above), the percentage of staminate flowers rises to over 90% at low densities but only to ~50% at high densities (31). Thus, the phenotypic outcome of the epistatic interaction between these two loci is also significantly affected by differences in planting density.

G × E effects are also important for crop improvement traits, especially as cultivation of early-domesticated crops spreads them into new habitats. For example, the genetic control of flowering time must be modified to reconcile changing photoperiod cues with the plant's reproductive cycle as plants are moved north and south from their center of origin (53). The spread of domesticated crops can lead to selection for insensitive alleles at flowering time genes. For example, teosinte is a tropical short-day plant and will not flower in the longer days of higher latitudes. The major gene affecting photoperiod response in *Z. mays*,

ZmCCT, has teosinte alleles highly sensitive to day length but less sensitive or insensitive maize alleles (24). The reduction in sensitivity allows many temperate maize varieties to flower in longer-day photoperiod regimes than can tropical maize or teosinte.

Genetic Control of Shattering and Flowering Time in *Setaria*

Previous work on shattering by Devos and coworkers (38, 54) established that there are two QTL that control shattering in foxtail millet. We have confirmed this using an F_7 recombinant inbred line (RIL) mapping population derived from a cross between domesticated foxtail millet (*S. italica*) and its wild ancestor, green millet (*S. viridis*) (55). We find two significant QTL: QTL1 (closest marker U30) that contributes over 35% of the phenotypic variation and QTL2 (closest marker U372) that contributes ~8%. There are no significant interactions at the Bonferroni-corrected genome-wide $P < 0.05$ level ($P < 10^{-7}$ for each interaction). However, when we used the more liberal significance level recommended by Holland (47) ($P < 0.0014$ for each interaction), there were 23 significant interactions, of which three were between the minor QTL (QTL2) and other loci without main effects. Notably, the other 20 interactions were between loci where neither had a main effect. There was no significant interaction between the major locus QTL1 and any other locus, congruent with the theme that domestication alleles at major-effect QTL are often insensitive to genetic interactions (Fig. 1B). The combined effect of the wild and domestication alleles of the two QTL loci is demonstrated in Fig. 2A, where the degree of shattering in plants with a wild green millet allele of QTL1 is in part a function of the allele at the minor-effect locus, QTL2, whereas shattering in plants with the domesticated foxtail millet allele at QTL1 is unaffected by QTL2. In contrast, Fig. 2B shows the domesticated allele of QTL2 to be more sensitive to the state of an interacting locus without a significant main effect (U283) than is the wild allele.

Flowering time in *Setaria* is strongly affected by environmental variation, including both photoperiod and temperature (56, 57). We searched for QTL and epistatic interactions for flowering time across eight trials that differed in environmental conditions, including photoperiod differences from short day (12 h light and dark) to long day (16 h light and 8 h dark) (58). We had previously reported stable QTL associations across trials, together with evidence of QTL \times environment interactions (58) (Figs. S1 and S2). Of the eight QTL shared across at least two trials, four had significant epistatic interactions with other non-QTL loci.

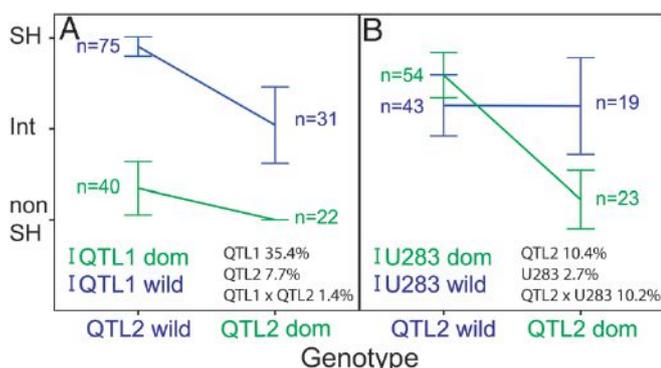


Fig. 2. Epistatic interactions for shattering in a cross between foxtail millet and green millet. The x axis represents the genotypes of homozygous RILs at one locus, and the color of the lines represents the different genotypes of a second locus. The y axis represents values of the shattering trait (SH, shattering; Int, intermediate shattering; NonSH, nonshattering). (A) Interaction between QTL2 (chromosome V) and QTL1 (chromosome IX) ($P < 0.05$ for single test; nonsignificant at the multiple test corrected significance level of $P < 0.0014$). (B) Significant interaction ($P < 0.0014$) between QTL2 and U283, a locus without a significant main effect. Percentages after each locus name and after interaction term reflect percent variance explained.

We used analysis of variance to explore the effects on flowering time of wild and domesticated alleles of QTL and their interacting loci across trials (Table 1 and Tables S1 and S2). For all QTL analyses, variation among trials was significant. QTL main effects were significant in all but one case, that of QTL III-1, where only the interaction between QTL III-1 and another locus was significant. In the other three epistatic interactions, both the QTL main effect and the interaction effect accounted for significant proportions of the flowering time variation.

Significant $G \times E$ effects were observed for three QTL and for one locus interacting with QTL VII-1. Three QTL also showed significant interactions between pairs of epistatic loci and environment. When compared across trials, 15 interactions were found in common across two trials, 5 across three trials, and 2 across four trials (Table S2). Resampling analyses suggest that those interactions found in at least three trials are unlikely to have occurred by chance ($P < 0.001$). Each shared interaction displayed the same direction of effects in the trials in which they occurred (e.g., examples in Fig. S3).

We assessed whether the flowering time data support the theme that domestication alleles of major-effect loci are more robust to genomic or environmental context than wild alleles. Using Levene's test for homogeneity of variance (59), we found that flowering time varied significantly more among genotypes with QTL alleles derived from the wild relative than among genotypes with QTL alleles derived from the domesticate, supporting the view that major-effect loci are robust to genetic background (Table 1). This was true when assessed across genotypes or across both genotypes and trials. However, when assessed across just trials, there was no consistent trend in whether flowering time variances differed between the genotypes with wild or domestication alleles. This is in contrast to the effect of density reported above on maize and teosinte branching and staminate flower production, perhaps because consistent timing of flowering has been under selection not only in the crop, foxtail millet, but also in the wild species green millet, which is now a weed of cultivation.

Conclusions

One of the main findings of our review, supported by our analysis of shattering and flowering time in foxtail millet, is that the phenotypic effects of major domestication alleles are notably less affected by genetic background than are wild alleles, whereas for minor-effect loci, other segregating alleles affect expression. The relative independence of major-effect loci indicates that selection is likely to have favored alleles with large and stable effects. The control of shattering in foxtail millet follows this trend, with a robust major-effect locus and a minor-effect locus that is involved in interactions with several other loci. Similarly, the alleles that control flowering time in domesticated millet often have more stable effects than alleles in the wild relative, across different genetic backgrounds. As with genetic background, the literature review suggests that the effect of the domestication allele is generally more stable than the wild allele across diverse environments, although it is clear from the study of flowering time in *Setaria* that this phenomenon depends on the trait under selection.

Our principal finding, that major domestication alleles are less sensitive to environmental changes and different genetic backgrounds than are wild alleles, suggests that selection on those alleles should be effective and that phenotypic change could therefore be rapid during domestication. This does not accord with archeobotanical observations of a protracted period of phenotypic change for characters such as shattering. How can we explain this apparent paradox? Patterns of crop use, such as the harvesting of grain while still green or the continued harvesting of wild species, may partially explain this paradox (10). For some traits, an additional genetic possibility may involve the limited availability of robust domestication alleles early in the domestication process. Phenotypic change would initially be slow if selection acted on domestication alleles with low phenotypic stability; the rate of evolution would subsequently increase if the early selected alleles were later displaced by more favorable

Table 1. ANOVA of flowering time in a foxtail by green millet cross, using QTL, interacting loci (if any), and trial

QTL	QTL marker (locus 1)	Interacting marker (locus 2)	Trial	Locus 1 × locus 2	Locus 1 × trial	Locus 1 × trial	Locus 1 × locus 2 × trial	QTL allele variance [epistasis]	QTL allele variance [trial]	QTL allele variance [overall]
II-1	66.00***	12.07**	23.05***	71.5***	3.48**	0.80	2.11*	n.s.	n.s.	n.s.
III-1	0.65 n.s.	0.55 n.s.	63.3***	36.69***	0.16	0.71	0.33	<*	n.s.	n.s.
IV-1	297.27***	–	104.35***	–	42.54***	–	–	n.s.	<***	n.s.
V-1	114.52***	–	76.59***	–	1.53	–	–	<***	>***	<***
V-2	62.88***	–	71.33***	–	0.97	–	–	<*	n.s.	<*
V-3	72.53***	–	62.66***	–	0.62	–	–	n.s.	n.s.	n.s.
VII-1	131.89***	104.58***	65.89***	114.50***	6.33***	3.65**	3.50**	<***	>***	<***
VIII-1	73.06***	0.02	81.11***	70.83***	1.28	1.08	2.35*	<***	<*	n.s.

Each QTL is represented by the map marker closest to the peak of that QTL (Table S1). Locus 2 represents the interacting locus as identified from the Epistacy analysis (*Materials and Methods*). Columns 2–8 give ANOVA results for each QTL, interacting locus (if any), trial, and relevant interaction terms. Values in these columns are *F* values, and asterisks denote significance level: **P* < 0.05, ***P* < 0.01, and ****P* < 0.001. The last three columns give the results from Levene's test for homogeneity of variance for each QTL. Terms in square brackets in column headings denote variation included in analysis: [epistasis], variation after removing trial effect; [trial], variation after removing genotypic effect; [overall], variation with both trial and genotypic effect included. Significant results signify that one allele has significantly more or less variation in flowering time than the other. The direction of the angle bracket signifies whether the domestication allele had less (<) or more (>) variation than the wild allele. n.s., not significant.

robust alleles. This reasoning has been followed by Zhang et al. (60), who suggested that the *sh4* locus controlling reduced shattering in rice arose relatively late in the domestication process, with initial selection acting on other loci with less favorable phenotypic effects. This explanation could reconcile the differences between the strong genetic signature for selection on *sh4* and the protracted time for the domesticated phenotype to predominate in the archeobotanical record. However, this explanation needs to be tested, for example, by estimating the relative time of appearance during the domestication process of the various loci that control a domestication trait. Such investigations will require more complete genome information to accurately assess epistatic interactions. As advances in population and quantitative genetics continue to expand our knowledge of major- and minor-effect genes, we will have a greater ability to assess the pervasiveness of epistatic and G × E effects on the process of crop domestication and the history of domestication alleles.

Materials and Methods

Plant Materials and Phenotyping. F₇ RILs were produced by single-seed descent from an F₂ mapping population derived from a cross between foxtail millet (*S. italica* variety B100) and green millet (*S. viridis* A10) (55). A set of 182 RILs were grown for phenotyping in eight field, greenhouse, and growth chamber trials. Two field, two greenhouse, and one growth chamber trial were performed at Oklahoma State University; two field trials took place at the University of Georgia (Athens, GA); and one growth chamber trial was at the Boyce Thompson Institute (Ithaca, NY). Trials varied in photoperiod, seed pretreatment, light intensity, time of year, and experimental design (58). All eight trials were phenotyped for flowering time (58); one field trial (F2_OK, field, OK) was phenotyped for seed shattering. Flowering time was measured as the time from germination to first emergence of the culm inflorescence (58). Shattering was measured on a qualitative scale, with 1 being nonshattering, 2 being intermediate, and 3 being shattering. Shattering was assessed by measuring the ease with which seeds were detached from the rachis by either pulling or bending with tweezers under a dissecting microscope, so that the nonshattering outcomes of breaking of the rachilla above the glumes or along the length of the rachilla could be contrasted with shattering where the grain and glumes separate cleanly from the rachilla along an abscission zone under the glumes. Intermediate shattering referred to those situations where some seeds separated at the abscission zone and some did not. The distinction between shattering and nonshattering closely parallels the method for analyzing shattering reported for the related panicoid grass, sorghum (38).

Twenty seeds were assessed for each of four replicates of each RIL, and measurements were repeated by a second observer.

Map Construction and QTL Analysis. Map construction used 684 SNP, simple sequence repeat, and sequence-tagged site markers (58). QTL analyses were performed in QTLCartographer, using composite interval mapping (CIM) and joint CIM with significance levels established by permutation (58).

Epistasis. Pairwise epistatic interactions were calculated using the program Epistacy (47), initially with a pairwise significance level of *P* < 2.3 × 10⁻⁷ (creating a Bonferroni-adjusted genome-wide significance level of *P* < 0.05) (58) and subsequently with a pairwise significance level of *P* < 0.0014, following the adjustment for multiple testing advocated by Holland (47). Partial *R*² terms for locus main effects and their interaction were calculated in Statistical Package for the Social Sciences (SPSS) version 19 by dividing the partial (type III) sum of squares by the corrected total sum of squares (47).

Statistical Tests and Data Visualization. SPSS version 19 (SPSS Inc.) was used for statistical analysis and to create graphs of phenotypic means and 95% confidence intervals for epistatic interactions. Tests for effect of genetic background on QTL allele variation for flowering time were analyzed using Levene's test for homogeneity of variances (59), after removing trial variation in an initial ANOVA analysis. Tests for effect of environment on QTL allele variation for flowering time were similarly analyzed after removing genotype (RIL) variation in an initial ANOVA analysis. Heat maps of epistatic interactions between all marker pairs were created in R version 2.13.2 (61). To more easily visualize colocalized interactions between trials in the heat maps, markers were grouped into threes along each chromosome, and the most significant *P* value for each group of three was chosen to represent that group. MATLAB version 8.1 (The MathWorks, Inc.) was used to write a resampling program to calculate the probability of trials sharing epistatic interactions between the same pairs of loci, given the number of significant interactions observed in each trial. Each trial was run for 1,000,000 replicates, and the 2.5 and 97.5 percentiles were calculated.

ACKNOWLEDGMENTS. A.N.D. acknowledges Jessica Stromski and the staff of the Cimarron Valley Research Station for their help in field work and funding from Department of Energy (DE-FG02-08ER64636) and Oklahoma Center for the Advancement of Science and Technology (PSB08-007 and PS11-035B). L.L. acknowledges the support of the Natural Sciences and Engineering Research Council and the Ontario Research Fund. K.M.O. acknowledges funding on domestication research by the National Science Foundation Plant Genome Research Program (IOS-1032023).

1. Harlan JR (1973) Comparative evolution of cereals. *Evolution* 27(2):311–325.
2. Buckler ES, et al. (2009) The genetic architecture of maize flowering time. *Science* 325(5941):714–718.
3. Xue WY, et al. (2008) Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat Genet* 40(6):761–767.

4. Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch SR (2009) The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc Natl Acad Sci USA* 106(34):14444–14449.
5. Burger JC, Chapman MA, Burke JM (2008) Molecular insights into the evolution of crop plants. *Am J Bot* 95(2):113–122.

6. Doebley JF, Gaut BS, Smith BD (2006) The molecular genetics of crop domestication. *Cell* 127(7):1309–1321.
7. Gross BL, Olsen KM (2010) Genetic perspectives on crop domestication. *Trends Plant Sci* 15(9):529–537.
8. Olsen KM, Wendel JF (2013) A bountiful harvest: Genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol* 64:47–70.
9. Allaby RG, Brown TA, Fuller DQ (2010) A simulation of the effect of inbreeding on crop domestication genetics with comments on the integration of archaeobotany and genetics: A reply to Honne and Heun. *Vegetation History and Archaeobotany* 19(2): 151–158.
10. Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates: Recent archaeobotanical insights from the Old World. *Ann Bot (Lond)* 100(5):903–924.
11. Heun M, Abbo S, Lev-Yadun S, Gopher A (2012) A critical review of the protracted domestication model for Near-Eastern founder crops: Linear regression, long-distance gene flow, archaeological, and archaeobotanical evidence. *J Exp Bot* 63(12): 4333–4341.
12. Purugganan MD, Fuller DQ (2011) Archaeological data reveal slow rates of evolution during plant domestication. *Evolution* 65(1):171–183.
13. Tanno K, Willcox G (2006) How fast was wild wheat domesticated? *Science* 311(5769): 1886.
14. Allaby RG, Fuller DQ, Brown TA (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proc Natl Acad Sci USA* 105(37): 13982–13986.
15. Allaby R (2010) Integrating the processes in the evolutionary system of domestication. *J Exp Bot* 61(4):935–944.
16. Tucker RP (2012) War and the environment. *A Companion to Global Environmental History*, eds McNeill JR, Mauldin, ES (John Wiley, Chichester, UK), pp 319–339.
17. Abbo S, Lev-Yadun S, Gopher A (2012) Plant domestication and crop evolution in the near east: On events and processes. *Crit Rev Plant Sci* 31(3):241–257.
18. Phillips PC (2008) Epistasis—The essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* 9(11):855–867.
19. Cheverud JM, Routman EJ (1995) Epistasis and its contribution to genetic variance components. *Genetics* 139(3):1455–1461.
20. Carter AJR, Hermisson J, Hansen TF (2005) The role of epistatic gene interactions in the response to selection and the evolution of evolvability. *Theor Popul Biol* 68(3): 179–196.
21. Huang P, et al. (2012) Phylogeography of Asian wild rice, *Oryza rufipogon*: A genome-wide view. *Mol Ecol* 21(18):4593–4604.
22. Morrell PL, Clegg MT (2007) Genetic evidence for a second domestication of barley (*Hordeum vulgare*) east of the Fertile Crescent. *Proc Natl Acad Sci USA* 104(9): 3289–3294.
23. Sadras VO, Slafer GA (2012) Environmental modulation of yield components in cereals: Heritabilities reveal a hierarchy of phenotypic plasticities. *Field Crops Res* 127: 215–224.
24. Hung HY, et al. (2012) ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. *Proc Natl Acad Sci USA* 109(28): E1913–E1921.
25. McWhirter KS, Brink RA (1978) Canalization of endosperm development on *opaque-7* maize. *Maize Breeding and Genetics*, ed Walden D (John Wiley, New York), pp 373–385.
26. Doust AN, Kellogg EA (2006) Effect of genotype and environment on branching in weedy green millet (*Setaria viridis*) and domesticated foxtail millet (*Setaria italica*) (Poaceae). *Mol Ecol* 15(5):1335–1349.
27. von Korff M, Léon J, Pillen K (2010) Detection of epistatic interactions between exotic alleles introgressed from wild barley (*H. vulgare* ssp. *spontaneum*). *Theor Appl Genet* 121(8):1455–1464.
28. Doebley J, Stec A, Hubbard L (1997) The evolution of apical dominance in maize. *Nature* 386(6624):485–488.
29. Hubbard L, McSteen P, Doebley J, Hake S (2002) Expression patterns and mutant phenotype of teosinte branched1 correlate with growth suppression in maize and teosinte. *Genetics* 162(4):1927–1935.
30. Studer A, Zhao Q, Ross-Ibarra J, Doebley J (2011) Identification of a functional transposon insertion in the maize domestication gene *tb1*. *Nat Genet* 43(11): 1160–1163.
31. Lukens LN, Doebley J (1999) Epistatic and environmental interactions for quantitative trait loci involved in maize evolution. *Genet Res* 74(3):291–302.
32. Hufford MB, et al. (2013) The genomic signature of crop-wild introgression in maize. *PLoS Genet* 9(5):e1003477.
33. Azhacuve P, Vidya-Saraswathi D, Komatsuda T (2006) High-resolution linkage mapping for the non-brittle rachis locus *btr1* in cultivated x wild barley (*Hordeum vulgare*). *Plant Sci* 170(6):1087–1094.
34. Azhaguvél P, Komatsuda T (2007) A phylogenetic analysis based on nucleotide sequence of a marker linked to the brittle rachis locus indicates a diphyletic origin of barley. *Ann Bot (Lond)* 100(5):1009–1015.
35. Sakuma S, Salomon B, Komatsuda T (2011) The domestication syndrome genes responsible for the major changes in plant form in the Triticeae crops. *Plant Cell Physiol* 52(5):738–749.
36. Gu XY, Kianian SF, Hareland GA, Hoffer BL, Foley ME (2005) Genetic analysis of adaptive syndromes interrelated with seed dormancy in weedy rice (*Oryza sativa*). *Theor Appl Genet* 110(6):1108–1118.
37. Doebley J, Stec A, Gustus C (1995) teosinte branched1 and the origin of maize: Evidence for epistasis and the evolution of dominance. *Genetics* 141(1):333–346.
38. Lin ZW, et al. (2012) Parallel domestication of the Shattering1 genes in cereals. *Nat Genet* 44(6):720–724.
39. Onishi K, Takagi K, Kontani M, Tanaka T, Sano Y (2007) Different patterns of genealogical relationships found in the two major QTLs causing reduction of seed shattering during rice domestication. *Genome* 50(8):757–766.
40. Eshed Y, Zamir D (1996) Less-than-additive epistatic interactions of quantitative trait loci in tomato. *Genetics* 143(4):1807–1817.
41. Li CB, Zhou AL, Sang T (2006) Rice domestication by reducing shattering. *Science* 311(5769):1936–1939.
42. Cai HW, Morishima H (2000) Genomic regions affecting seed shattering and seed dormancy in rice. *Theor Appl Genet* 100(6):840–846.
43. Ishikawa R, et al. (2010) Allelic interaction at seed-shattering loci in the genetic backgrounds of wild and cultivated rice species. *Genes Genet Syst* 85(4):265–271.
44. Thurber CS, et al. (2010) Molecular evolution of shattering loci in U.S. weedy rice. *Mol Ecol* 19(16):3271–3284.
45. Nagai K, et al. (2012) Two novel QTLs regulate internode elongation in deepwater rice during the early vegetative stage. *Breed Sci* 62(2):178–185.
46. Holland J (2001) Epistasis and plant breeding. *Plant Breed Rev* 21:27–92.
47. Holland J (1998) EPISTACY: A SAS program for detecting two-locus epistatic interactions using genetic marker information. *J Hered* 89(4):374–375.
48. Hemani G, Knott S, Haley C (2013) An evolutionary perspective on epistasis and the missing heritability. *PLoS Genet* 9(2):e1003295.
49. Coles ND, McMullen MD, Balint-Kurti PJ, Pratt RC, Holland JB (2010) Genetic control of photoperiod sensitivity in maize revealed by joint multiple population analysis. *Genetics* 184(3):799–812.
50. Coles ND, Zila CT, Holland JB (2011) Allelic effect variation at key photoperiod response quantitative trait loci in maize. *Crop Sci* 51(3):1036–1049.
51. Kump KL, et al. (2011) Genome-wide association study of quantitative resistance to southern leaf blight in the maize nested association mapping population. *Nat Genet* 43(2):163–168.
52. Tian F, et al. (2011) Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet* 43(2):159–162.
53. Olsen KM, Wendel JF (2013) Crop plants as models for understanding plant adaptation and diversification. *Front Plant Sci* 4:290.
54. Devos KM, Gale MD (2000) Genome relationships: The grass model in current research. *Plant Cell* 12(5):637–646.
55. Bennetzen JL, et al. (2012) Reference genome sequence of the model plant *Setaria*. *Nat Biotechnol* 30(6):555–561.
56. Kokubu T, Ishimine Y, Miyaji Y (1977) Variations of growth-period of Italian millet strains, *Setaria italica* BEAUV. and their responses to day-length and temperature: II. Changes of growth-period strains gathered from different districts, both native and foreign, due to the different seeding dates. *Mem Fac Agric Kagoshima Univ* 13(22): 55–75.
57. Schreiber MM, Oliver LR (1971) Two new varieties of *Setaria viridis*. *Weed Sci* 19(4): 424–427.
58. Mauro-Herrera M, et al. (2013) Genetic control and comparative genomic analysis of flowering time in *Setaria* (Poaceae). *G3* 3(2):283–295.
59. Levene H (1960) Robust tests for equality of variances. *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, eds Olkin I, Ghurye SG, Hoëffding W, Madow WG, Mann HB (Stanford University Press, Stanford, CA), pp 278–292.
60. Zhang LB, et al. (2009) Selection on grain shattering genes and rates of rice domestication. *New Phytol* 184(3):708–720.
61. Team RDC (2011) *R: A Language and Environment for Statistical Computing* (Team RDC, Vienna).

Establishing the validity of domestication genes using DNA from ancient chickens

Linus Girdland Flink^{a,b,1,2}, Richard Allen^a, Ross Barnett^a, Helena Malmström^c, Joris Peters^{d,e}, Jonas Eriksson^f, Leif Andersson^{f,g}, Keith Dobney^h, and Greger Larsson^{a,3}

^aDurham Evolution and Ancient DNA, Department of Archaeology, ^bSchool of Biological and Biomedical Sciences, Durham University, Durham DH1 3LE, United Kingdom; ^cDepartment of Evolutionary Biology, Evolutionary Biology Centre, Uppsala University, 75236 Uppsala, Sweden; ^dDepartment of Veterinary Sciences, Institute of Palaeoanatomy, Domestication Research and the History of Veterinary Medicine, Ludwig Maximilian University Munich, 80539 Munich, Germany; ^eBavarian State Collection of Anthropology and Palaeoanatomy, 80333 Munich, Germany; ^fDepartment of Medical Biochemistry and Microbiology, Uppsala University, SE-75123 Uppsala, Sweden; ^gDepartment of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, SE-75007 Uppsala, Sweden; and ^hDepartment of Archaeology, University of Aberdeen, St. Mary's, Aberdeen AB24 3UF, Scotland

Edited by Dolores R. Piperno, Smithsonian National Museum of Natural History and Smithsonian Tropical Research Institute, Fairfax, Washington, DC, and approved October 29, 2013 (received for review August 23, 2013)

Modern domestic plants and animals are subject to human-driven selection for desired phenotypic traits and behavior. Large-scale genetic studies of modern domestic populations and their wild relatives have revealed not only the genetic mechanisms underlying specific phenotypic traits, but also allowed for the identification of candidate domestication genes. Our understanding of the importance of these genes during the initial stages of the domestication process traditionally rests on the assumption that robust inferences about the past can be made on the basis of modern genetic datasets. A growing body of evidence from ancient DNA studies, however, has revealed that ancient and even historic populations often bear little resemblance to their modern counterparts. Here, we test the temporal context of selection on specific genetic loci known to differentiate modern domestic chickens from their extant wild ancestors. We extracted DNA from 80 ancient chickens excavated from 12 European archaeological sites, dated from ~280 B.C. to the 18th century A.D. We targeted three unlinked genetic loci: the mitochondrial control region, a gene associated with yellow skin color (β -carotene dioxygenase 2), and a putative domestication gene thought to be linked to photoperiod and reproduction (thyroid-stimulating hormone receptor, *TSHR*). Our results reveal significant variability in both nuclear genes, suggesting that the commonality of yellow skin in Western breeds and the near fixation of *TSHR* in all modern chickens took place only in the past 500 y. In addition, mitochondrial variation has increased as a result of recent admixture with exotic breeds. We conclude by emphasizing the perils of inferring the past from modern genetic data alone.

selective sweep | breed formation | animal domestication | *Gallus gallus* | cultural history

The resolution afforded by multiple genetic loci and—more recently—complete genomes has led to an increased understanding of the pattern and process of plant and animal domestication (1, 2). More specifically, genetic analyses have uncovered selective sweeps, quantitative trait loci, and even causative mutations underlying a wide range of behavioral and morphological traits, some of which define specific breeds, and others that differentiate domestic plants and animals from their wild ancestors (1, 3, 4).

Because many of these traits are present in either single or relatively few closely related modern breeds, the earliest occurrences of specific phenotypes (and the underlying causative mutations) are presumed to have occurred well after the initial domestication process. These phenotypes are referred to (at least in the plant genetic literature) as “improvement genes” (2). In animals, these traits include hairlessness in Mexican and Peruvian dogs (5), dorsal hair ridges in Vietnamese, Thai, and Rhodesian Ridgebacks (6), excessive skin folds in western Shar-pei (7), double muscling in two cattle breeds (8), and a curly

coat mutation found in Selkirk Rex cats (9), none of which are thought to have been present during early domestication.

Some causative mutations, however, underlie traits found in numerous, distantly related breeds. Alleles that are fixed in domestic variants—and often presumed to have been under selection at the outset of domestication—are referred to in both the plant (2) and animal (3) domestication literature as “domestication loci” (or domestication genes). In some cases, including gray coloring (10) and altered gaits in horse breeds (11), brachycephaly in dogs (12), and muscle growth in pigs (13), no hypotheses have been proposed for the time-frame of first appearance of these traits. In others, however, the commonality of both small size (14, 15) and chondrodysplasia (16) across modern dog breeds and the widespread occurrence of pea-combs in chickens (17), led the authors of these studies to suggest that the genetic mutations underlying these characteristics were selected for during the early stages of the domestication process. More recently, a whole-genome resequencing study that compared variation in 14 unrelated dog breeds and wolves identified 36 regions potentially targeted during early domestication and included 10 genes that allowed dogs to better digest starches (18). Because increased amylase activity was ubiquitous in dogs but

Significance

Recent studies have identified the genetic basis of numerous traits that differentiate modern domestic species from their wild counterparts. In both plants and animals, traits (and the genes underlying them) found ubiquitously in modern breeds are often presumed to have been selected early during the domestication process. Here, by determining genetic variability in ancient European chickens over the past 2,000 years, we show that a mutation thought to be crucial during chicken domestication was not subjected to strong human-mediated selection until much later in time. This result demonstrates that the ubiquity of mutations, which differentiate modern wild and domestic taxa, does not necessarily imply ancient origins.

Author contributions: L.G.F. and G.L. designed research; L.G.F., R.A., R.B., H.M., and J.P. performed research; J.E. contributed new reagents/analytic tools; L.G.F., R.A., R.B., H.M., J.P., J.E., L.A., and K.D. analyzed data; and L.G.F., J.P., L.A., K.D., and G.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Data deposition: The sequences reported in this paper has been deposited in the GenBank database (accession nos. KF753251–KF753289).

¹Present address: Department of Archaeology, University of Aberdeen, St. Mary's, Aberdeen AB24 3UF, Scotland.

²Centre for Genome-Enabled Biology and Medicine, University of Aberdeen, Aberdeen AB24 3RY, Scotland.

³To whom correspondence should be addressed. E-mail: greger.larsson@durham.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1308939110/-DCSupplemental.

absent in wolves, the authors concluded that this change must have occurred when early dogs began adapting to a starch-rich diet provided by early farmers.

Recent genetic and archaeological research has also shed light on domestic chickens and their primary ancestor, the Red Junglefowl (*Gallus gallus*) (19). Based on archaeological bones identified from Neolithic sites in the Yellow River basin, chickens were thought to have been domesticated as early as 6000 B.C. (20). This conclusion has recently been questioned, however, because bones presumed to originate from chickens in the original faunal analysis (21, 22) have since been shown to be pheasants (23, 24). As a result, a reevaluation of all of the early finds is necessary to establish the true chronology and geography of chicken domestication.

Genes that differentiate modern domestic chickens from Red Junglefowl include those that underlie the *yellow skin* phenotype present in the vast majority of Western, commercial chicken breeds, as well as numerous geographically restricted and fancy breeds. Yellow skin is caused by a recessive allele of the *BCDO2* (β -carotene dioxygenase 2) gene (25). *BCDO2* encodes the β -carotene dioxygenase 2 enzyme that cleaves colorful carotenoids into colorless apocarotenoids (26). Although the expression of the dominant allele in skin tissue results in white skin color, the recessive allele possesses one or more *cis*-acting and tissue-specific regulatory mutations that inhibit expression of *BCDO2* in skin tissue. Provided that sufficient carotenoids are available in the diet, the recessive allele reduces carotenoid cleavage and allows them to be deposited in skin tissue, leading to yellow skin (25). This recessive *BCDO2* allele is thought to have been acquired through hybridization with the Gray Junglefowl (*Gallus sonneratii*) in South Asia (25). Red and Gray Junglefowl are known to hybridize in contact zones in the Indian subcontinent (27, 28), and it is possible that domestic poultry engaged in the same behavior after they were introduced from Southeast Asia. Given the ubiquity and genomic signatures of strong human-driven selection of the *yellow skin* trait in modern, Western commercial chickens (29), Eriksson et al. (25) suggested that this trait was favored by humans after chickens acquired the trait in South Asia, but before the first wave of domesticated chickens arrived in Europe between 900 and 700 B.C. (30, 31).

In addition, a recent analysis of pooled wild and domestic chicken samples revealed strong selection signatures across a number of loci, as well as a missense mutation in the thyroid-stimulating hormone receptor (*TSHR*), a locus possibly linked to shifts in seasonal mating (29). Given its ubiquity in domestic breeds (264 of 271 birds representing 36 global populations were homozygous for the sweep allele; the remaining 7 were heterozygous) and the general absence of the derived allele in Red Junglefowl, the authors of that study concluded that the *TSHR* locus may have played a crucial role during chicken domestication (29).

Here, we investigate whether the *TSHR* gene was selected for during the early stages of chicken domestication (29), and if early poultry keepers favored the *BCDO2* gene that underlies yellow skin in chickens soon after it was acquired from the Gray Junglefowl (25, 29). To do so, we genotyped SNPs linked with the sweep alleles in both *TSHR* and *BCDO2* in 80 ancient European chickens dating from ~280 B.C. to the 18th century A.D. (Table S1 and *SI Materials and Methods*). If *TSHR* played a critical role during the domestication process, all of the samples analyzed here should have been fixed for the derived *TSHR* allele, as has been demonstrated in worldwide modern chicken populations (29). Similarly, if *BCDO2* and the *yellow skin* phenotype was favored and maintained soon after its introgression from Gray Junglefowl, a significant proportion of the ancient European individuals should also possess this phenotype. Finally, we assess the hypothesis that the presence of mitochondrial DNA (mtDNA) control region (CR) haplogroups A–D has resulted from the recent introduction of East Asian chickens into the European

gene pool, and that haplogroup E is historically associated with European chickens (32).

Results

For each ancient individual, we attempted to amplify a 58-bp fragment surrounding the candidate missense (Gly > Arg) SNP in the *TSHR* gene (29), a 51-bp fragment surrounding a SNP in the *BCDO2* gene associated with the *yellow skin* allele (SNP B in table 1 of ref. 25), and a 201-bp fragment of the mtDNA CR (33). Overall, 55 of 80 (69%) ancient chicken remains provided reproducible results for at least one of three loci (Fig. 1, Fig. S1, and Table S1). We observed allelic drop out in a number of heterozygous specimens for both *TSHR* and *BCDO2*. However, we estimated the probability of falsely assigning a true heterozygous individual as a homozygote to <0.01. In addition, external replication on 12 samples yielded identical mtDNA CR and *TSHR* sequences (*SI Materials and Methods*).

Among the 44 specimens from whom the *TSHR* locus was successfully genotyped, 8 individuals were homozygous for the derived (domestic) sweep allele, 14 were homozygous for the wild-type allele, and 22 specimens were heterozygous (Fig. 1 and Table S1). The results of a Fisher's exact test revealed that the sweep allele was significantly less frequent in the ancient sample than in modern chickens ($P < 0.0001$). A binomial probability test demonstrated that observing these frequencies among the ancient samples, assuming the frequencies of the modern samples (Fig. S2 and Table S2) (29), is very unlikely ($P < 0.0001$). In addition, a χ^2 test on *TSHR* genotype frequencies derived from the second to third century A.D. Quintana/Künzing Roman population (Fig. 1 and Tables S1 and S3) revealed that the observed frequencies are consistent with Hardy–Weinberg equilibrium ($P > 0.95$, $\chi^2 = 0.004$, $df = 1$), although one group had fewer than the recommended minimum number of expected genotypes/individuals.

Of the 25 ancient specimens successfully genotyped for the *BCDO2* locus, 20 were homozygous for the white skin allele (found in Red Junglefowl) and five individuals were heterozygous (Fig. 1 and Table S2). Because only one functional copy of the *BCDO2* gene is necessary to effectively cleave carotenoids, the *yellow skin* phenotype can only be expressed in chickens that are homozygous for the yellow skin allele and consume sufficient carotenoids in their diet. Of the 25 successfully genotyped chicken samples, none could express the *yellow skin* phenotype. Because genotype/phenotype frequencies reported previously were selected on the basis of their phenotypes (nonrandom sampling) (25), we did not carry out statistical comparisons of allele frequencies between ancient and modern populations.

The targeted mtDNA CR fragment was successfully sequenced in 38 individuals (Table S1). The topology of a maximum-likelihood tree constructed from an alignment of 201-bp haplotypes matched the neighbor-joining tree generated by Liu et al. (19), confirming previous observations that this specific 201-bp fragment is sufficient for recovering the major clades present in the chicken mitochondrial tree (Fig. 2 and Fig. S3) (33, 34). We identified a total of three haplotypes among the ancient specimens, all of which clustered within the E clade on the chicken mitochondrial tree (19, 35) (Fig. 2 and Table S4). The E3 ($n = 1$) and E6 ($n = 2$) haplotypes (19) were present only in Medieval and post-Medieval chickens from England (Tables S1 and S4), whereas the remaining 35 individuals possessed a 201-bp haplotype corresponding to haplotypes E1, E5, E12, E15, or E16 described using a 519-bp fragment (Table S4) (19). We find a significant difference in haplogroup frequencies (pooled into two groups of chickens: those belonging to haplogroup E and those belonging to haplogroup A–D) between the ancient and modern datasets (Fisher's exact test, $P < 0.002$) (Table S5). Assuming the frequency reported for modern European chickens (Table S5) (i.e., ~15% of modern European

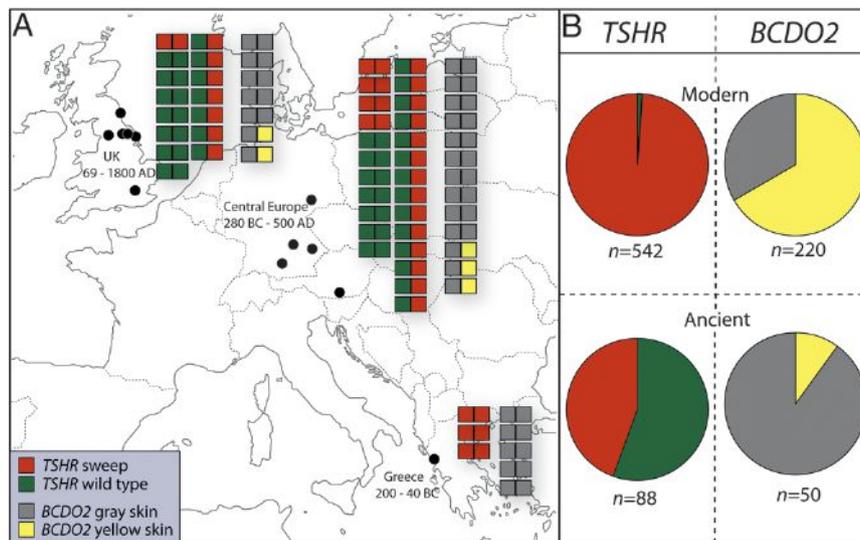


Fig. 1. (A) A map showing the locations and chronology of archaeological sites (black dots) from three different European regions where ancient chicken remains analyzed in this study were excavated (see also *SI Materials and Methods* and Fig. S1). The pairs of colored boxes adjacent to each region show which of two alleles of two nuclear genes (*TSHR* and *BCDO2*) were present in each genotyped individual (Table S1). (B) Pie charts showing the differing allele frequencies of all (pooled) ancient samples (from 290 B.C. to the 18th century A.D.) (Table S2). Allele frequencies found in modern chicken populations at the *TSHR* locus were derived from ref. 29 and at the *BCDO2* locus from ref. 25.

chickens possess haplotypes from clades A–D), a binomial test revealed that the probability of observing only the E haplogroup in 43 ancient specimens (the unique 38 sequences combined with previously published data) (Table S5) is <0.001 .

Discussion

***TSHR* Domestication Locus.** The locus encoding *TSHR* on chromosome 5 in domestic chickens has recently been shown to have undergone a massive selective sweep (29). A nonconservative amino acid substitution (a missense mutation Gly558Arg) was identified as a potential causal and target mutation for the selective sweep. Although the function (and corresponding phenotype) associated with the derived allele remains unknown, it is possible that this gene variant affects photoperiod control and the absence of strict seasonal reproduction, a trait commonly found in domestic animals but rare or absent among their wild relatives (29). Because 264 of 271 modern birds representing 36 globally distributed populations were homozygous for the derived sweep allele (the seven remaining were heterozygous), and

because the Red Junglefowl individuals that also possessed the mutation were thought have acquired it from domestic chickens, Rubin et al. (29) suggested that *TSHR* was a domestication locus. If this selective sweep occurred during the early phase of domestication, all ancient chickens that postdate this event (and are located outside the natural distribution range of Red Junglefowl, thus eliminating the potential for backcrossing with wild birds) should also possess the derived allele.

The results presented here, however, demonstrate that although the derived allele was present in European chickens dated to ~280 B.C. to the 16th–18th century A.D., it was only found on 43% of the typed chromosomes and the wild-type allele persisted at intermediate frequencies until at least the 16th–18th century AD (Fig. 1 and Tables S1 and S2). These results suggest that the *TSHR* mutation was neither a prerequisite, nor critical in the immediate aftermath of chicken domestication. Although the strength of the selection pressure that drove the sweep haplotype in modern populations is not in doubt (29), the data presented here suggest that the fixation of the derived allele in European chickens was likely much closer to the present-day and certainly within the last 500 y, possibly commensurate with the improvement of farmyard animals that began during the industrial revolution (27).

In one possible scenario, the sweep allele was fixed in Western Europe and spread across the world with newly synthesized commercial breeds that emerged during the second half of the 19th century (27). This hypothesis seems overly simplistic, however, given that the *TSHR* allele was found to be fixed (or nearly so) in modern Egyptian Fayoumi chickens and Silkie, Cochin, and Hua-Tung chickens from China (29) that are unlikely to have been strongly influenced by European commercial breeds. It is therefore possible that the *TSHR* sweep allele became fixed in populations originating outside Europe well before the creation of modern breeds, and that these birds in turn replaced the initial populations introduced into Europe. This narrative is supported by the observation that the Classical Greek chickens (Kassope) were the only population analyzed in this study that was fixed for the sweep allele (Fig. 1 and Table S1). Although the earliest Central European chickens that arrived north of the Alps ~600–400 B.C. likely descended from founder populations

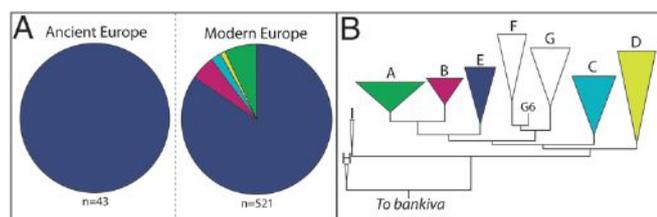


Fig. 2. (A) Two pie charts showing differences in mtDNA haplogroup frequencies between ancient and modern European chickens. The ancient sample consists of specimens sequenced in this study ($n = 38$) and from a previous publication ($n = 5$) (34) and the modern data was compiled from (19, 32, 39) (Table S5). The colors in A correspond to haplogroups depicted in a phylogenetic tree in B. (B) A maximum-likelihood phylogenetic tree constructed using 519-bp haplotypes defined in ref. 19. White (or colorless) clades were not identified in either modern or ancient European chickens. A more detailed maximum-likelihood tree, including detailed description of methods and results (including nodal support values), is described in the *SI Materials and Methods* (see also Fig. S3).

present in eighth to fifth century B.C. Greek settlements on the Mediterranean coast, the first century B.C. chicken population from Kassope could have arrived as part of a secondary introduction of chickens. During the Achaemenid (550–330 B.C.) and Hellenistic (323–31 B.C.) periods, the Aegean region was a commercial hub that drew trade from most of the Near and Middle East, including the Indus Valley (36). Such trade networks could have been linked to the introduction of chicken populations that had undergone a selective sweep at the *TSHR* locus. Crucially, however, none of the Greek individuals possessed the introgressed Gray Junglefowl *BCDO2* allele (Fig. 1), suggesting these two loci were selected for and fixed at different times and places.

The *BCDO2* Yellow Skin Locus. The study that discovered the *TSHR* sweep (29) also confirmed a selective sweep encompassing the *BCDO2* locus. Unlike *TSHR*, however, *BCDO2* is only fixed in a limited number of (often commercial and geographically widespread) breeds (25). The *yellow skin* phenotype was previously shown to result from the presence of a *BCDO2* allele that domestic chickens acquired not from their primary ancestor—the Red Junglefowl—but from introgression between domestic chickens and Gray Junglefowl indigenous to South Asia (25).

The data presented here reveal that although the *yellow skin* allele was present on ~10% of chromosomes in ancient European chickens, not a single ancient bird was homozygous for the Gray Junglefowl allele (and therefore capable of expressing the *yellow skin* phenotype) (Tables S1 and S2). This finding is consistent with 17th–19th century records suggesting that a number of widespread, prolific, and economically important breeds raised in western and southern Europe, including the Dorking, Houdan, Sultan, Spanish, or B/W Bantams, clearly had white legs, but the Hamburgh, Polish, Turkish, and Crève Cœur breeds had leg colors ranging between slate blue and dark leaden-blue (27, 37). However, some 17th century European breeds, including the Padua, did possess yellow legs (37) and the trait was explicitly mentioned in relation to heavy, fast-growing types, including Cochin, Brahma, and Malay breeds imported by sea (and therefore named Captain's birds) from the Far East into Europe and the United States during the 1820s–1850s (27). It is therefore possible that the rise in frequency of the *yellow skin* phenotype occurred only after recent introductions of foreign birds to Europe and the breed formation process that consequently led to the creation of modern, widespread, commercial broiler, and egg-laying breeds. Tegetmeier (27), for example, noted the remarkable rapid growth and great size attained by crossbred birds produced by mating Cochins with the large traditional French Crève Cœur, La Flèche, and Houdan breeds. The ubiquity of the *yellow skin* phenotype in commercial, modern, and some rare, geographically restricted breeds can, therefore, be explained by rapid worldwide spread of newly synthesized commercial chickens. This suggestion is further supported by the fact that the *yellow skin* phenotype appears infrequently or is completely absent in rare breeds, such as the Friesian Fowl, Houdan, and Westfälischer Totleger (25).

The Mitochondrial Control Region. The mtDNA control region is a widely used locus in chicken genetic studies and has frequently been used to investigate domestication, admixture, and migration (19, 34, 35, 38). Although previous studies have concluded that modern European chickens (and Western commercial breeds) primarily possess mitochondrial haplotypes belonging to the E clade (19, 35), additional studies (32, 39) have demonstrated that haplogroups A–D are also present in a wide variety of breeds, most likely as a result of the intentional importation and hybridization of East Asian breeds into Europe during breed formation and breed improvement over the past 500 y (32).

Although a recent study of complete mtDNA chicken genomes revealed the existence of 14 mitochondrial clades found worldwide

(35), all 38 ancient chickens sequenced in this study, and five archaeological Spanish chickens typed in a previous study (34), possessed exclusively E-clade haplotypes (which represent 3 out of the 14 major clades in ref. 35) (Tables S1 and S5). These results first demonstrate that chickens initially introduced to Europe possessed a small fraction of the variability present in Southeast Asia, where chickens were originally domesticated. Second, these results demonstrate that all haplotypes belonging to other haplogroups therefore represent recent introductions. Interestingly, because all but three ancient European chickens possessed a single haplotype (probably corresponding to the common E1 haplotype) (19) (Table S4), the presence of haplotypes E3 and E6 in Medieval and post-Medieval contexts from the United Kingdom may imply secondary introductions.

Finally, although the nuclear loci typed in this study show a dramatic reduction in variability between ancient and modern populations, the mitochondrial signatures reveal an initial uniformity followed by an increase in haplogroup diversity in modern birds. These seemingly divergent genetic patterns reflect different human goals over the last two centuries, including not only intensive selection for traits associated with behavior, production (*TSHR*), and skin color (*BCDO2*), but also for phenotypic variability in fancy breeds achieved through the importation of East Asian varieties (32).

Conclusions

A variety of genetic techniques and analytical approaches have led to the identification of selective sweeps and causative mutations that differentiate populations of domestic plants and animals from their modern wild counterparts. The ubiquity of sweep alleles across numerous breeds has often been used as an argument to infer their ancient origins, and as an argument that the SNPs and associated traits were selected for during the early phases of domestication (4, 14, 17, 18, 40). The patterns of allelic diversity presented here, however, challenge the hypothesis that modern variation can be directly mapped onto the past. Instead, the allelic variability within *TSHR* in ancient chickens demonstrates that this locus was not crucial for the early development of domestic chickens, and that the fixation of the derived allele (at least in European chickens) took place only in the past few hundred years. Similarly, although the *yellow skin* allele was present in ancient European chickens, the phenotype was very rare in the past. In addition, the fixation of the *BCDO2* allele in numerous modern chicken breeds was likely also a recent occurrence, possibly coincident with recent breed formation instigated during the industrial revolution (27), and the subsequent development of modern commercial chicken breeds. This pattern demonstrates the ease with which we can underestimate the potential of alleles to become ubiquitous through rapid strong selection leading to fixation, followed by geographic proliferation through human-assisted migration.

This result is consistent with most ancient DNA studies that have revealed the hazards of assuming that modern genetic data can be used to draw accurate conclusions about past population variability. For example, a mutation in the *NAM-B1* gene, associated with increased grain size in wheat, was thought (on the basis of fixation in modern cultivars) to have been selected for during the early phases of domestication (40). A genetic survey of 19th century historical seeds, however, revealed that fixation of the modern, sweep allele occurred only recently during crop improvement (41). A similar study of ancient maize concluded that, although two genes (*tb1* and *pb1*) had been fixed for the domestic variant by ca. 4,000 y ago, one gene (*su1*) still possessed significant variability as long ago as ca. 2,000 y ago (42). Similarly, a number of studies of ancient mtDNA have shown that lineage replacement (often rapid and geographically widespread) was common among both domestic (43, 44) and wild (45, 46) populations.

The processes of recent breed formation, coupled with strong directed selection and admixture with exotic breeds, have rapidly and radically shaped the gene pools of modern domestic plants and animals, often distorting their deeper genetic history (47). As a result, hypotheses regarding past population dynamics drawn solely from modern datasets require verification through direct observation. This need is especially true where zooarchaeological or historical records suggest recent origins for some traits. As mentioned at the outset, conclusions drawn from the genetic variability in modern samples suggest that small size (14, 15) and chondrodysplasia (16) in dogs, and pea-combs in chickens (17), occurred early during domestication. However, small dogs (<30-cm shoulder height) do not appear in the Eurasian archaeological record until after the first millennium B.C. (30), the earliest reported evidence for chondrodysplasia is found in Egyptian tomb art dating to the late third millennium B.C. (48), and the peacomb trait in chickens was still rare in mid-19th century English fowl (27).

These lines of evidence alone do not undermine claims for much earlier appearances of these and other domestication-related traits, but they do suggest that modern genetic data should be considered cautiously and within a broader context that includes an appreciation for the potential of rapid and widespread demographic shifts. Future studies that investigate ancient genetic variability in loci known to underlie behavioral, dietary, and phenotypic differences between wild and domestic plants and animals will help to reveal the first appearance and timing of selection pressures. These data will lead to the establishment of a significantly more robust geographic and temporal interpretative framework to more fully understand the early patterns and processes of domestication.

Materials and Methods

Ancient chicken specimens were selected to represent three geographical locations and two major time bins (first, Late Iron Age, Late Hellenistic and Roman contexts, dating to ~280 B.C. to the fifth century A.D., and second, Medieval and post-Medieval contexts, dating to approximately the 10th to 18th century A.D.) (Table S1). The majority of chicken bones come from closed contexts that have been dated either through stratigraphic (cultural) association or direct radiocarbon dating (*SI Materials and Methods*).

DNA extraction was performed in dedicated ancient DNA laboratories at Durham University and Uppsala University following stringent laboratory procedures according to commonly applied guidelines (49, 50). The work was carried out by laboratory personal wearing protective laboratory coats and overshoes, or coveralls and dedicated laboratory clogs, and double pairs of gloves (with the outer pairs of gloves changed in between every step of the preparation/extraction procedure). All equipment and work surfaces are routinely cleaned before and after each use with a dilute solution of bleach [5–10% (wt/vol) active sodium hypochlorite] followed by rinsing with ddH₂O and ethanol [70–99% (vol/vol)]. A strict one-way system for entering the laboratories is in use to avoid carry-over of post-PCR contaminants.

DNA extractions were performed as previously described (44) (see *SI Materials and Methods* for a detailed account of experimental procedures). To authenticate the results, 12 chickens were independently replicated for the CR and *TSHR* SNP at Uppsala University (Table S1). CR sequences were amplified using previously published PCR primers GG144F-GG387R (33), and PCR and sequencing primers for *TSHR* and *BCDO2* were designed in PSQ Assay Design (Qiagen) (Table S6). CR sequences were Sanger-sequenced in both directions at the DNA sequencing facility at Durham University. *TSHR* and *BCDO2* PCR amplicons were genotyped on the Q24 (Qiagen) pyrosequencer in the Department of Archaeology at Durham University, United Kingdom, or on a PSQ 96MA (Biotage) pyrosequencer at Uppsala University. Full PCR cycling conditions and sequencing methods are presented in detail in the *SI Materials and Methods*.

MtDNA CR sequence data were analyzed and assembled in Geneious v.5.4 (51) and manually edited in Se-AL (<http://tree.bio.ed.ac.uk/software/seal>). Reference sequences were compiled from previous publications (19, 32) and aligned with the unique sequences deposited into GenBank (KF753251-KF753289).

ACKNOWLEDGMENTS. We thank Anders Götherström for assistance with data replication; and the London Archaeological Archive and Research Centre, Museum of London, Alex Croom (Arbeia Roman Fort and Museum), and Deborah Jaques for sample material and discussion. Sample material from Kasso and Central European sites is courtesy of Bavarian State Collection of Anthropology and Palaeoanatomy, Munich; material from Beverley is courtesy of Humber Field Archaeology, Hull; and material from York is courtesy of Deborah Jaques. This work was supported by a Research Councils UK Academic Fellowship (to G.L.); Grant NE/F003382/1 from the Natural Environment Research Council; and Grant F/00 128/AX from the Leverhulme Trust. This manuscript resulted from a catalysis meeting entitled “Domestication as an Evolutionary Phenomenon: Expanding the Synthesis” that was awarded and hosted by the National Evolutionary Synthesis Centre, National Science Foundation #EF-0905606 in 2011.

- Andersson L (2013) Molecular consequences of animal breeding. *Curr Opin Genet Dev* 23(3):295–301.
- Olsen K-M, Wendel J-F (2013) A bountiful harvest: Genomic insights into crop domestication phenotypes. *Annu Rev Plant Biol* 64:47–70.
- Andersson L (2012) How selective sweeps in domestic animals provide new insight into biological mechanisms. *J Intern Med* 271(1):1–14.
- Wayne RK, vonHoldt BM (2012) Evolutionary genomics of dog domestication. *Mamm Genome* 23(1–2):3–18.
- Drögemüller C, et al. (2008) A mutation in hairless dogs implicates FOXI3 in ectodermal development. *Science* 321(5895):1462.
- Salmon Hillbertz NHC, et al. (2007) Duplication of FGF3, FGF4, FGF19 and OROAV1 causes hair ridge and predisposition to dermoid sinus in Ridgeback dogs. *Nat Genet* 39(11):1318–1320.
- Olsson M, et al. (2011) A novel unstable duplication upstream of HAS2 predisposes to a breed-defining skin phenotype and a periodic fever syndrome in Chinese Shar-Pei dogs. *PLoS Genet* 7(3):e1001332.
- McPherron AC, Lee S-J (1997) Double muscling in cattle due to mutations in the myostatin gene. *Proc Natl Acad Sci USA* 94(23):12457–12461.
- Filler S, et al. (2012) Selkirk Rex: Morphological and genetic characterization of a new cat breed. *J Hered* 103(5):727–733.
- Rosengren Pielberg G, et al. (2008) A cis-acting regulatory mutation causes premature hair graying and susceptibility to melanoma in the horse. *Nat Genet* 40(8):1004–1009.
- Andersson LS, et al. (2012) Mutations in DMRT3 affect locomotion in horses and spinal circuit function in mice. *Nature* 488(7413):642–646.
- Bannasch D, et al. (2010) Localization of canine brachycephaly using an across breed mapping approach. *PLoS ONE* 5(3):e9632.
- Van Laere A-S, et al. (2003) A regulatory mutation in IGF2 causes a major QTL effect on muscle growth in the pig. *Nature* 425(6960):832–836.
- Gray MM, Sutter NB, Ostrander EA, Wayne RK (2010) The IGF1 small dog haplotype is derived from Middle Eastern grey wolves. *BMC Biol* 8(1):16.
- Sutter NB, et al. (2007) A single IGF1 allele is a major determinant of small size in dogs. *Science* 316(5821):112–115.
- Parker HG, et al. (2009) An expressed *fgf4* retrogene is associated with breed-defining chondrodysplasia in domestic dogs. *Science* 325(5943):995–998.
- Wright D, et al. (2009) Copy number variation in intron 1 of SOX5 causes the Peacomb phenotype in chickens. *PLoS Genet* 5(6):e1000512.
- Axelsson E, et al. (2013) The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature* 495(7441):360–364.
- Liu Y-P, et al. (2006) Multiple maternal origins of chickens: Out of the Asian jungles. *Mol Phylogenet Evol* 38(1):12–19.
- West B, Zhou B-X (1988) Did chickens go north? New evidence for domestication. *J Archaeol Sci* 15(5):515–533.
- Chow BS (1981) The animal remains discovered at Cishan village, Wu'an, Hebei Province. *Kaogu Xuebao* 62(3):339–348. Chinese.
- Chow BS (1983) The animal remains discovered at Baiyingsite, Tangyin, Henan Province. *Papers on Chinese Archaeology* 3:48–50. Chinese.
- Deng H, Yuan J, Song G-D, Wang C-S, Masaki E (2013) The re-examination of the Domestic Chicken in Ancient China. *Archaeology* 6(6):83–96. Chinese.
- Peters J (1998) *Römische Tierhaltung und Tierzucht. Eine Synthese aus archäozoologischer Untersuchung und schriftlich-bildlicher Überlieferung. Passauer Universitätschriften zur Archäologie* 5. (Marie Leidorf, Rahden, Westfalen, Germany).
- Eriksson J, et al. (2008) Identification of the *yellow skin* gene reveals a hybrid origin of the domestic chicken. *PLoS Genet* 4(2):e1000010.
- Kiefer C, et al. (2001) Identification and characterization of a mammalian enzyme catalyzing the asymmetric oxidative cleavage of provitamin A. *J Biol Chem* 276(17):14110–14116.
- Tegetmeier WB (1867) *The Poultry Book: Comprising the Breeding and Management of Profitable and Ornamental Poultry, Their Qualities and Characteristics; To Which is added "The Standard of Excellence in Exhibition Birds," Authorized by the Poultry* (G. Routledge and Sons, London).
- McGowan PJK (1994) Phasianidae (Pheasants and Partridges). *Handbook of the Birds of the World*, eds del Hoyo J, Elliott A, Sargatal J (Lynx Edicions, Barcelona, Spain), pp 434–552.
- Rubin C-J, et al. (2010) Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature* 464(7288):587–591.

30. Benecke N (1994) *Archäozoologische Studien zur Entwicklung der Haustierhaltung in Mitteleuropa und Südsandinavien von den Anfängen bis zum ausgehenden Mittelalter* (Akademie, Berlin).
31. Carrasquilla Hernández F (1992) Some comments on the introduction of domestic fowl in Iberia. *Archaeofauna* 1:45–93.
32. Dana N, et al. (2010) East Asian contributions to Dutch traditional and western commercial chickens inferred from mtDNA analysis. *Anim Genet* 42(2):125–133.
33. Storey AA, et al. (2007) Radiocarbon and DNA evidence for a pre-Columbian introduction of Polynesian chickens to Chile. *Proc Natl Acad Sci USA* 104(25):10335–10339.
34. Storey AA, et al. (2012) Investigating the global dispersal of chickens in prehistory using ancient mitochondrial DNA signatures. *PLoS ONE* 7(7):e39171.
35. Miao YW, et al. (2013) Chicken domestication: an updated perspective based on mitochondrial genomes. *Heredity (Edinb)* 110(3):277–282.
36. Roaf M (1990) *Cultural Atlas of Mesopotamia and the Ancient Near East* (Equinox, Facts on File, New York).
37. Aldrovandi U (1600) *Ornithologiae Tomus Alter cum Indice Copiosissimo Variarum Linguarum* (Giovanni Battista Bellagamba, Bologna).
38. Gongora J, et al. (2008) Indo-European and Asian origins for Chilean and Pacific chickens revealed by mtDNA. *Proc Natl Acad Sci USA* 105(30):10308–10313.
39. Revay T, Bodzsar N, Mobegi VE, Hanotte O, Hidas A (2010) Origin of Hungarian indigenous chicken breeds inferred from mitochondrial DNA D-loop sequences. *Anim Genet* 41(5):548–550.
40. Uauy C, Distelfeld A, Fahima T, Blechl A, Dubcovsky J (2006) A NAC Gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314(5803):1298–1301.
41. Asplund L, Hagenblad J, Leino MW (2010) Re-evaluating the history of the wheat domestication gene NAM-B1 using historical plant material. *J Archaeol Sci* 37(9):2303–2307.
42. Jaenicke-Després V, et al. (2003) Early allelic selection in maize as revealed by ancient DNA. *Science* 302(5648):1206–1208.
43. Malmström H, et al. (2008) Barking up the wrong tree: Modern northern European dogs fail to explain their origin. *BMC Evol Biol* 8:71.
44. Ottoni C, et al. (2013) Pig domestication and human-mediated dispersal in western Eurasia revealed through ancient DNA and geometric morphometrics. *Mol Biol Evol* 30(4):824–832.
45. Barnes I, Matheus P, Shapiro B, Jensen D, Cooper A (2002) Dynamics of Pleistocene population extinctions in Beringian brown bears. *Science* 295(5563):2267–2270.
46. Shapiro B, et al. (2004) Rise and fall of the Beringian steppe bison. *Science* 306(5701):1561–1565.
47. Larson G, et al. (2012) Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci USA* 109(23):8878–8883.
48. Boessneck J (1988) *Die Tierwelt des Alten Ägypten* (C. H. Beck, Munich).
49. Cooper A, Poinar HN (2000) Ancient DNA: Do it right or not at all. *Science* 289(5482):1139.
50. Gilbert MT, Bandelt HJ, Hofreiter M, Barnes I (2005) Assessing ancient DNA studies. *Trends Ecol Evol* 20(10):541–544.
51. Drummond AJ, et al. (2011) Geneious v5.4, Available from www.geneious.com/. Accessed July 1, 2011.



Archaeological and genetic insights into the origins of domesticated rice

Briana L. Gross^{a,1} and Zhijun Zhao^b

^aBiology Department, University of Minnesota Duluth, Duluth, MN 55812; and ^bInstitute of Archaeology, Chinese Academy of Social Sciences, Beijing 100710, China

Edited by Dolores R. Piperno, Smithsonian National Museum of Natural History and Smithsonian Tropical Research Institute, Washington, DC, and approved November 4, 2013 (received for review June 28, 2013)

Rice (*Oryza sativa*) is one of the most important cereal grains in the world today and serves as a staple food source for more than half of the world's population. Research into when, where, and how rice was brought into cultivation and eventually domesticated, along with its development into a staple food source, is thus essential. These questions have been a point of nearly continuous research in both archaeology and genetics, and new information has continually come to light as theory, data acquisition, and analytical techniques have advanced over time. Here, we review the broad history of our scientific understanding of the rice domestication process from both an archaeological and genetic perspective and examine in detail the information that has come to light in both of these fields in the last 10 y. Current findings from genetics and archaeology are consistent with the domestication of *O. sativa japonica* in the Yangtze River valley of southern China. Interestingly, although it appears rice was cultivated in the area by as early as 8000 BP, the key domestication trait of nonshattering was not fixed for another 1,000 y or perhaps longer. Rice was also cultivated in India as early as 5000 BP, but the domesticated *indica* subspecies currently appears to be a product of the introgression of favorable alleles from *japonica*. These findings are reshaping our understanding of rice domestication and also have implications for understanding the complex evolutionary process of plant domestication.

Oryza rufipogon | *Oryza nivara* | domestication gene

Archaeological Evidence for Rice Domestication and Development of Rice Agriculture

Given the broad importance of domesticated rice as a food source, its origin and development from the wild species *Oryza rufipogon* have driven much of the interest and research in archaeology in East and South Asia during the last century. An early focus was the geographic origin of domesticated rice. Several areas were proposed, including India (1), South China (2), the Yangtze River area in China (3), the so-called “belt region” with its great diversity of *Oryza* species along the southern slope of the Himalayas (4), and coastal swamp habitats in Southeast Asia (5). However, there were comparatively few serious studies on the chronology of rice domestication, which was presumed to have occurred about 10,000 y ago, probably because previous research showed that the origins of agriculture in the other parts of world, such as western Asia, took place ~10,000 y ago (6). During the last 10 y, research into rice origins and dispersal has benefitted, as has domestication research in other regions of the world (papers in this volume), from the generation of a considerable amount of new empirical data from archaeological sites, itself driven by the application of new methodological procedures that can better detect *Oryza* and follow its early history. For example, the recent

widespread use of flotation (7–9) in East and South Asia has resulted in the retrieval of rich macrobotanical remains of rice grains and husks from some important sites (10–15). Phytolith analysis has also proved useful for identifying microscopic remains of plants to the genus level, including in very early (e.g., Pleistocene) deposits where grains and husks are not present. This advance has allowed identification of presumably domesticated, or at least cultivated, rice occurring beyond the areas of wild *Oryza* distribution with enough accuracy to separate the two major subspecies of *Oryza sativa* (*indica* and *japonica*) from each other (16–18).

Many scholars now accept that the Yangtze River area in China is the place where rice was originally domesticated as a consequence of these newer findings (11, 19–23). However, as discussed elsewhere in this paper, whether *indica* and *japonica* had single or multiple origins is a question under active research in the genetic and archaeological arenas, and there is little consensus of opinion with regard to the available genetic evidence (24). The resolution of this question depends to a large extent on archaeological research, which has revealed separate cultivation origins for *indica* and *japonica*. Current arguments in archaeology are also focused on fundamental questions of when rice cultivation began in China and how long the domestication process took. Cultivation

and subsequent domestication are increasingly seen as being considerably more separated in time than once thought, as a horizon of what's termed “predomestication cultivation” sometimes lasting thousands of years is being increasingly documented in the Old World (see Introduction in this volume), and this also appears true for rice. Moreover, recent studies suggest that there is no clear boundary line between hunting-gathering and agriculture and that the transformation between the two is not a revolutionary change but rather a slow process of qualitative and quantitative shifts that may have taken thousands of years (6, 11). These questions, in turn, are related to theories of agricultural origins in China and around the world that are currently of great interest in anthropology and archaeology (25). Accordingly, we now focus on new archaeobotanical data bearing on these issues and the subsequent spread of rice into Korean Peninsula, the Japanese archipelago, and India.

Archaeobotanical Data from China. In recent years, flotation of archaeological sediments for recovery of macrobotanical remains of plants has been carried out on more than

Author contributions: B.L.G. and Z.Z. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹To whom correspondence should be addressed. E-mail: blgross@d.umn.edu.

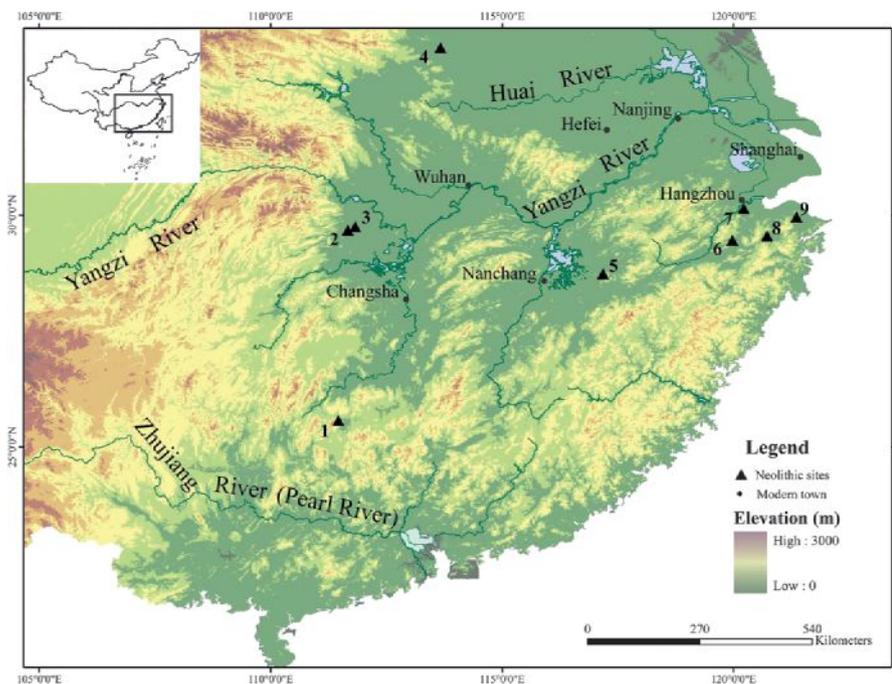


Fig. 1. Locations of the sites with early rice remains in the Yangtze Rivers areas: (1) Yuchanyan; (2) Pengtoushan; (3) Bashidang; (4) Jiahu; (5) Xianrendong/Diaotonghuan; (6) Shangshan; (7) Kuahuqiao; (8) Xiaohuangshan; (9) Hemudu/Tianluoshan.

100 archaeological sites throughout China, and numerous charred plant remains have been retrieved for study (Fig. 1). They are from a variety of crop species including rice (10, 11). The earliest rice remains recovered in China have been reported from three archaeological sites: Xianrendong and Diaotonghuan in Jiangxi Province (26) and Shangshan in Zhejiang Province (27). The cultural remains of these sites are dated to around 10,000 BP (all dates are in calibrated years), although it should be noted that the cultural deposits in the Xianrendong cave site have a very long sequence; the lower layer was recently dated to about 20,000 BP (28). This information suggests that a new date might also be needed for the rice remains found in the upper layers of this site.

Shangshan is an early Neolithic settlement with house and pit features, and artifactual remains that include pottery and stone tools. The cultural assemblage recorded can be divided into two periods: the Shangshan culture dated to ~11,000–9000 BP, and the Kuahuqiao culture dated to 8000–7000 BP (29). Although more than 400 soil samples were floated, only 10 charred rice grains and a few rice spikelet bases were recovered. Most of these belong to the Kuahuqiao cultural horizon. A few rice grains were recovered from the Shangshan culture horizon. Rice remains were also found by other methods. For example, rice husks can be easily iden-

tified in the paste of pottery sherds, and they were commonly found in sherds dating to both of the two periods. Heaps of rice husks were also found in burnt soil blocks from the early period layers of the site. The combined evidence indicates that Shangshan people intensively exploited rice.

The Shangshan rice was believed to be in an early stage of domestication, based on grain size and morphological characteristics (e.g., length-to-width ratios) (27). However, others note that grain size and shape may exhibit considerable variability in wild and domesticated populations, some of which is probably influenced by plant responses to environmental factors and therefore may not be a reliable indicator of early domestication (30, 31). In light of the abundance of rice husks in the pottery and other site contexts, it seems that Shangshan people had a high demand for rice, and the cultivation of rice may have begun at that time (11). Further work is required to unequivocally establish rice cultivation; if it occurred, it could be interpreted as being primarily an attempt to improve the yield of wild rice.

In China, the time around 8,000 y ago appears to have been critical for agricultural origins, not only for rice in the Yangtze River area, but also for millet agriculture in the Yellow River system of northern China. For example, several archaeological sites exhibiting the characteristics of early rice farming have all been dated to around 8000

BP. They are Pengtoushan and Bashidang in Hunan Province (32), Kuahuqiao and Xiaohuangshan in Zhejiang Province (33, 34), and Jiahu in Henan Province (35).

Jiahu was a permanent village dated by dozens of radiocarbon determinations to a period between 9000 and 7800 BP. Flotation work was carried out on a total of 125 soil samples, and a large number of charred plant remains were recovered, including several hundred rice grains (36). Other plant remains include soybean (*Glycine* sp.), water chestnut (*Trapa* sp.), lotus roots (*Nelumbo nucifera*), and acorn (*Quercus* sp.). Zhao's research on the Jiahu rice indicates that it may well be domesticated, as its grain phenotypic characteristics, including size and shape, are much like modern domesticated rice. A discussion has ensued about these characteristics (37, 38). For example, it has been suggested that Jiahu rice might belong to a wild rice species because the grains are remarkably small (31). Alternatively, it has also been argued that the Jiahu rice grains are not small but characterized by a great variation in size, based on measurements of hundreds of rice grains recently recovered from Jiahu (38). Another factor that should be considered in establishing the status of the Jiahu rice is the abundance of weedy grasses, which may represent weeds of cultivation (e.g., *Digitaria* and *Echinochloa* spp.). Furthermore, the location of the Jiahu site is far from the natural distribution of wild rice today. All of these factors indicate that domesticated rice and rice agriculture were established at the site 8,000 y ago.

It should also be noted that the rice remains at Jiahu were accompanied by a large amount of wild food resources, such as lotus and water chestnut, along with fish bones and shells. Quantitative analysis of the plant remains showed that rice was not the dominant plant in the remains (36). It appears that rice did not play a dominant role in the subsistence of the Jiahu people and that the overall subsistence economy was a mixture of plant cultivation, fishing, and other wild resource procurement. This type of mixed subsistence is a pattern coming to light in other regions of the world where early agriculture was established.

The discovery of the Hemudu site in the 1970s resounded throughout the world. Because of the waterlogged condition of the site, organic materials were well preserved (39). A huge number of plant remains have been recovered, among which the most noticeable were rice. Some scholars accordingly suggested that the Hemudu people might have had a very mature agricultural economy based on rice (40). However, little was known about the

phenotypic characteristics of the rice that could reveal if they were morphologically still wild or domesticated. Another unresolved problem was whether the rice was indeed the main food resource at Hemudu, especially considering that a significant number of other edible wild plant remains were found during the excavation.

The Tianluoshan site, discovered in 2004, afforded a chance to answer these questions. Its location is only 7 km from Hemudu, and the cultural deposits of the two sites are almost identical (41). A systematic sampling strategy was applied during the excavation to recover plant remains, including water screening and flotation. More than 200 soil samples have been processed thus far, and a tremendous number of plant remains have been found, including rice, water chestnut, acorns, bottle-gourd, *Euryale ferox*, *Ziziphus jujube*, *Diospyros* sp., and various weed seeds. The most important study of the Tianluoshan data are a systemic analysis of the rice spikelet bases carried out by Fuller et al. (42). The results show that the Tianluoshan rice consists of a high proportion of shattering, WT spikelets, which suggests that the process of rice domestication was not yet complete in the Hemudu period, i.e., sometime about 6500 BP. However, the quantitative analysis of the plant remains suggested that rice was one of the most important food resources at Tianluoshan and that the Tianluoshan people probably engaged in rice farming activities. Nonetheless, rice farming did not replace hunting-gathering as the dominant economy of the Tianluoshan residents or even probably the Hemudu culture. Wild resources such as acorns were still important foods at that time.

Liu et al. (38) published additional research that analyzed the morphological characteristics of both grains and spikelet bases from early occupations at Shangshan, Kuahuqiao, and Hemudu. They also presented new measurements of rice grains recovered from Jiahu. They argued based on all of the data that the process of rice domestication began during the early Holocene at about 9000 BP with the “management of phenotypically wild plants.” Fuller et al. (37) counter argued that the new metric data of Liu et al. were still insufficient to determine the wild or domesticated status for rice remains from the early sites. Analyses of spikelet bases unearthed from the early sites in the lower Yangtze River region, such as Kuahuqiao, Luojiajiao (Majiabang period), and Tianluoshan, show few signs of domestication.

Loss of shattering is a key characteristic of domesticated cereal crops that distinguishes

them from their wild ancestors (31). Furthermore, domestication is a long and slow process in which the proportion of shattering forms should be high in the beginning and then decrease gradually until nonshattering types dominate populations. Before or during this transition, wild or cultivated rice that shattered at maturity may have been harvested at an immature stage to prevent resource loss. Therefore, the study and detection of “immature rice” is a strong conceptual advance and will become a key for understanding the process of rice domestication, as emphasized by Fuller. However, the accurate identification of immature rice is still at issue (43). Broadly speaking, the distinctive characteristics of immaturity, shattering, and nonshattering states based on spikelet bases has the potential to be reliable, whereas characters relating to grain morphometrics and their diagnostic power for domestication need further study.

As can be seen, major questions relating to when rice cultivation ensued and how it can be identified, together with how domestication should be determined, currently surround issues of rice cultivation in China. On the basis of current data, two significant patterns in agricultural origins and development can be identified. First, as shown by the percentages of nonshattering grains through time, the fixation of nonshattering grains under cultivation was a slow process that occurred over a few thousand years. The comprehensive study of the early sites (e.g., Shangshan) suggests that a predomestication cultivation horizon may have existed for rice before domestication and that cultivation probably began 9,000–8,000 y ago. Second, the transition to rice agriculture from hunting and gathering was not a clear-cut revolutionary change but a slower evolutionary process. During this long-term process, hunting and gathering gradually waned, whereas rice agriculture slowly achieved a dominant position and finally became the major subsistence practice. This slow process is a pattern emerging in many other areas of the world.

Archaeobotanical Data from Korea, Japan, and India.

Korea. *Oryza* is not native to the Korean peninsula; thus, rice research focuses on diffusion from China and possible routes by which the plant may have spread. It was conventionally thought that rice farming spread from China to the Korean peninsula during the Early Mumun period (3400–2800 BP). Some studies suggested that rice may have arrived in Korea as early as the Chulmun period (7500–4000 BP) (13). Recent archaeobotanical studies have clarified this issue.

Macrobotanical evidence now indicates that Chulmun subsistence, especially the Middle-Late Chulmun, was clearly based on an agricultural economy. However, Chulmun agriculture was a dry-land farming system based not on rice, but apparently on millets, including both foxtail (*Setaria italica*) and broomcorn millet (*Panicum miliaceum*), legumes such as soybean (*Glycine max*), adzuki bean (*Vigna angularis*), and other crops (14). No Chulmun rice remains have been found through macrobotanical analysis. Nevertheless, rice presence is indicated by phytoliths found in pottery shards dating to the Chulmun period. Future work will address further aspects of rice at this site.

Analysis of plant remains confirmed that rice agriculture was indeed a major part of the subsistence economy of the Mumun period and also suggested that paddy rice farming occurred. This fact suggests that rice agriculture appeared in the Korean peninsula as an already developed rice farming system (13, 14), indicating that rice agriculture diffused into the Korean peninsula. The route of spread is still under active research and debate.

Japan. As with the Korean peninsula, wild rice does not occur in Japan today and probably never did. It is therefore generally agreed that rice agriculture diffused to Japan, and it is thought this occurred during the Yayoi period, which began ~2800 BP (44). Very similar to the situation in the Korean peninsula, early rice farming in Japan appears to have involved paddy rice farming. Although a highly productive rice agriculture probably began in the Yayoi period, it is increasingly likely that domesticated rice was introduced into the Japanese archipelago earlier, during the late Jomon period, some time around 4000 BP. Rice seed impressions on Jomon period pottery documented through scanning electronic microscopic study (45) appear to establish that domesticated rice was introduced into Japan before the Yayoi period, although it is not clear how large a part of the subsistence base rice was at that time. This rice appears to have been a part of a dry-land agriculture system, as other crops of dry-land farming have been found in sites dating to the late and middle Jomon, including various millets such as barnyard millet (*Echinochloa utilis*) and legumes (soybean and adzuki beans).

It is interesting to see the similarity between the Korean peninsula and Japan regarding the early development of agriculture. First, the beginning of agriculture in these two areas was characterized by a dry-land farming system with the major crops being millets and beans. Second, rice agriculture diffused

to these two areas at a relatively late time, but with a paddy rice farming system. This new farming system quickly replaced the indigenous dry-land farming system. It can be suggested this change was one of the reasons for the cultural transitions that took place; i.e., from Mumun to Chulmunin the Korean peninsula and Jomon to Yayoi in Japan. Third, sporadic rice remains were also found in the periods before the occurrence of paddy rice farming system in these two areas. This early arriving rice presumably had little impact on the indigenous dry-land farming system, but more evidence is needed to study its significance.

India. The prehistory of *indica* and *japonica* in India presents one of the more interesting stories of domestication, long distance spread, and subsequent interactions of cultivars within a single genus of plants. Both *O. rufipogon* and another close wild relative, *Oryza nivara*, are native to India and well distributed there today, and probably were present since the Pleistocene (46). The country has a number of long archaeological sequences with good plant records including those in the Ganges River valley in the north where rice, likely wild *O. rufipogon* and *O. nivara*, is documented by 9000 BP (46, 47). It is now recognized that the Indian subcontinent was probably an independent center of agricultural origins with important regions in the Ganges plain and to the south on the Deccan Plateau. Native plants that were cultivated or domesticated before crops were introduced from elsewhere include mung bean and small-seeded grasses, among others (47). The question of an origin of *indica* rice in India has been under active discussion, and recent research has done much to clarify and resolve the issue. It now appears that an independent origin of cultivation of ancestral *indica* or proto-*indica* rice took place in the Ganges plains, but that the plant was completely domesticated only when domesticated *japonica* arrived from China and hybridized with it about 4,000 y ago (47). *Indica* consumption began early, by 8400 BP, and the plant was cultivated and appears to have been a staple food by 5000 BP (47).

Summary. A subject that is as important as the origins and spread of domesticated *Oryza* and that is increasingly informed by multiple empirical databases from genetics and archaeology has naturally given rise to a number of controversies. Points of disagreement include how and when cultivation began and domesticated varieties emerged, how this can be documented in rice's major center of origin in China, and when archaeological rice

remains can be associated with an economy partly or fully dependent on rice as a staple food. These issues have done much to inspire and advance methodological aspects of rice domestication research and make it ever clearer that establishing accurate and feasible criteria for distinguishing domesticated and wild species is of prime importance in agricultural origins research.

Methodological techniques will continue to evolve. For example, recent work by Zhao and Gu (48) demonstrates that identification of spikelet bases has its limitations, despite the fact that pertinent qualitative features have a clear tendency of polarization between wild and domesticated rice. Wild rice spikelet bases have a shallow and round abscission scar and a small distinct vascular pore, whereas domesticated spikelet bases display irregular-shaped and deeper scars. Although this would seem to make them easy to distinguish, the irregular morphology of domesticated rice spikelet bases and their small size can make microscopic measurements very difficult. Applying qualitative criteria alone may introduce perceptual differences and variability in how different scholars evaluate the same features and make repeatability and comparisons between sites difficult. Ongoing examinations of other traits such as rice embryo characteristics (48) may provide additional data on domestication that will independently, or in combination with other criteria, provide more precise identifications of wild and domesticated rice. Productive discussions of theoretical issues are also dependent on such advances.

Genetic Evidence for Rice Domestication

Just as physical remains provide evidence of the presence and transformation of wild rice into domesticated rice in the archaeological record, the genomes of wild and domesticated rice preserve a record of the evolutionary forces they have been subject to over time. These records include information ranging from the possible geographical origin(s) of extant domesticated rice as a whole to the origin and assembly of individual alleles that combine to create the domestication phenotype (e.g., loss of shattering, increase in grain size and number, change in grain color and plant growth stature). Combined with complimentary information from archaeology and ethnobotany, this information can be essential to understanding the process of rice domestication. What has genetic information from wild and domesticated rice yielded to date? Lately, the answer seems to be that genetics has told us a constantly changing story. Although this is not necessarily surprising, given the rapid advances in

genetic technology over the last 20 y, it has had the effect of both expanding our understanding of the origin and evolution of domesticated rice and stirring strong feelings about the history of one of the world's most important food crops.

The two major *Oryza sativa* subspecies are differentiated by a number of morphological and physiological characters, along with a substantial (although incomplete) sterility barrier (49, 50). All genetic analyses have confirmed the distinctiveness of the subspecies and further confirmed the existence of recognized subgroups within these groups (i.e., *temperate* and *tropical japonica* along with *aus*) (51). The consistent genetic and phenotypic distinctiveness has long been considered an indicator that *indica* and *japonica* might have distinct origins. However, genetic distinctiveness alone is not enough to establish independent origins. Multiple origins of a single domesticated species is most convincingly demonstrated when varieties or subspecies within the species show genetic affiliations to existing wild populations with distinct geographic and genetic provenance. What, then, prevents us from quickly identifying the genetic and geographic source populations that gave rise to domesticated rice, given its obvious socioeconomic importance and the extensive genetic resources for the *Oryza* system? After all, the likely geographic origin of the extant maize lineage was pinpointed more than 10 y ago (52, 53), whereas arguments for and against multiple origins of rice have appeared in the literature on a nearly annual basis.

There are three main complicating factors that have made the history of domesticated rice difficult to read based on the patterns in its genome. These factors include (i) a paucity of genetic markers; (ii) a paucity of samples from the wild relative of domesticated rice; and (iii) difficulty resolving the relationship between gene genealogies for domestication genes and gene genealogies for neutral genes. Below, we will describe the progress of rice domestication genetics with these three factors in mind and discuss how the field might be advanced in the future.

Neutral Markers. Many authors have documented evidence consistent with independent origins of the *japonica* and *indica* subspecies based on molecular markers ranging from allozymes to retrotransposons (54–57). However, to a large extent, the recent exchange of papers identifying a single vs. multiple origins of domesticated Asian rice began in 2006, with the publication by Londo et al. showing that phylogeographic evidence was consistent with separate origins

of *indica* and *japonica* (58). This paper surveyed only three sequenced loci, but its strength lay in the relatively broad sampling of both domesticated and wild Asian rice (203 *O. sativa* and 129 *O. rufipogon*), with a strong effort to sample across the range of the wild species. In particular, this paper included many field-collected samples of *O. rufipogon* from China; these samples were particularly important because they include one of the potential geographic origins of domesticated rice. The phylogenetic patterns detected in this paper suggested two likely origins of domesticated rice, with *indica* originating in eastern India and *japonica* originating in southern China.

The possibility that the *O. sativa* subspecies had unique origins was subsequently bolstered by the analysis of Caicedo et al. (59). Although this paper did not include a phylogeographical component (it was focused on the effects of selection on the patterns of genetic variation in domesticated rice), the large number of sequenced regions (111 gene fragments) provided a strongly supported phylogenetic tree indicating that *indica* and *japonica* were not just genetically distinct but also that each subspecies was more closely related to a separate set of *O. rufipogon* than to the other subspecies. In contrast to the Londo et al. analysis, however, this analysis included a relatively small number of *O. rufipogon* samples (21 total), the majority of which were from China or Nepal. Indeed, the phylogenetic tree presented in the paper would seem to indicate that all subspecies of domesticated Asian rice arose from wild populations similar to those found in modern-day China. Overall, the lack of samples from other countries makes it impossible to draw conclusions about the geographic origin of rice (and, indeed, the paper's authors do not try to). A complementary study published in the same year, using 22 sequence-based markers, also showed patterns consistent with independent origins of the two subspecies (60). The level of sampling, however, was still too low (at 30 samples) to pinpoint a possible geographical origin of either subspecies.

Shortly thereafter, an analysis using 60 microsatellites conducted by Gao and Innan suggested "nonindependent" origins of the two subspecies (61). The authors of this paper took the interesting approach of evaluating evidence for bottlenecks at corresponding loci in the genomes of the two subspecies. The logic was that the stochastic nature of a domestication bottleneck would reduce diversity at some neutral loci to a greater extent than other loci, but that levels of diversity at a given locus would not be

correlated across subspecies if they were domesticated independently (assuming an absence of parallel selective pressures acting on these loci and limited gene flow between independent origins). The authors found a significant, positive correlation when they compared the subspecies, consistent with a single origin or extensive recent gene flow. The samples used in this study included 92 individuals, 35 of which were *O. rufipogon*, all sampled from within China (61, 62). Whether this wild reference population was the most appropriate comparison for *indica* varieties that may or may not have been domesticated in that geographical area is not addressed in the paper.

Domestication Genes. This succession of papers coincided with the first set of domestication genes being cloned and characterized in Asian domesticated rice, both of which controlled shattering (*sh4*, *qSH1*) (63, 64). Although the *qSH1* domestication allele was confined to a subset of *japonica* varieties, analyses of *sh4* revealed that the mutations associated with the nonshattering phenotype had a single origin and that a single allele was now distributed across *japonica* and *indica* (despite the sterility barrier between them) (65, 66). The single origin of a gene underlying this major domestication trait prompted researchers to consider possibilities of interactive domestication scenarios such as the "snowballing model" and the "combination model" proposed by Sang and Ge (65, 67). These models attempt to reconcile the divergence at neutral loci with the similarity at domestication loci through various scenarios involving gene flow, although multiple origins were still favored based on the deep divergence between the subspecies (57, 68). Following the shattering genes, the *Rc* gene, underlying a change in pericarp color from red in wild rice to white in domesticated rice, was cloned and its origin was characterized in a diverse collection of more than 400 rice cultivars (69, 70). Much like *sh4*, the *rc* allele (causing white pericarps) was found to be common across both *indica* and *japonica*. The survey also clearly indicated that the domestication allele originated in *japonica* and spread to *indica*.

All three of the genes mentioned thus far can be classified as domestication genes, in that they either control a trait that is critical to a domesticated condition (loss of shattering) or they are found in a large majority of domesticated varieties (white pericarps). The *Waxy* (*Wx*) gene, which was characterized at the molecular level over a number of years (71–74), is best described as an improvement or diversification gene: it is selected in

response to cultural preferences in some areas of the world but not uniformly favored. The majority of *indica* varieties have the fully functional *Wx^a* allele, whereas *japonica* varieties generally carry the *Wx^b* allele (resulting in a stickier grain), or its derivative, the *waxy* allele, which results in fully glutinous rice. The *waxy* allele has spread out of *japonica* and into some *indica* varieties when the glutinous phenotype was favored (71). This example conforms again to an emerging pattern at this point in rice domestication genetic research: domestication and improvement alleles are either restricted to the *japonica* subspecies (e.g., *qSH1*, *Wx^b*), or they originated there and subsequently spread to *indica* (e.g., *rc*, *waxy*) (75). Again, the common origin of important domestication genes might seem to indicate that *O. sativa* was domesticated only once, but this scenario was inconsistent with the deep genetic divergence between the subspecies based on neutral loci. The combination model, suggesting that the domestication process was initiated multiple times and that this was followed by extensive introgression of strongly selected domestication alleles, was considered to be most consistent with the data at this point (75).

These four loci represented an early view into the future of rice domestication genetics based on functional genes: following the identification of these domestication and improvement genes, new genes were identified at a rapid rate (76, 77). In general, when surveys included a broad sampling of rice varieties, the allele associated with a domesticated state showed one of three patterns: (i) the allele was unique to *japonica* (78, 79), (ii) the same allele was found in a subset of both *japonica* and *indica* (80, 81), or (iii) the same allele was found in the majority of *japonica* and *indica* varieties (82, 83). These results revealed the genetic complexity of rice domestication; alleles underlying domestication phenotypes can have different origins and different distributions depending on their desirability and dispersal across the range of domesticated rice. However, in every case where the survey included an evolutionary component, the origin of the domestication allele was found to be the *japonica* subspecies (80, 81). As a result, the phylogenies generated from domestication genes showed *japonica* and *indica* as a monophyletic group (69, 84), consistent with a single origin for allele controlling a domestication trait. In contrast, neutral loci generally recovered a polyphyletic relationship (51, 59), and analyses that included *O. rufipogon* showed that *japonica* and *indica* were more closely related to different populations of the wild species

than to each other, consistent with multiple origins. It should be noted that the discordance in phylogenetic trees is expected under a scenario of introgression at loci controlling domestication traits (76) and illustrates the difference between gene trees and species trees (85), but the question of how to interpret the origins of *indica* and *japonica* in light of these conflicts was still an open one.

Genomic Patterns. These common patterns for domestication genes vs. neutral loci were challenged in a publication by Molina et al. (86), where 630 gene fragments were sequenced in 20 *O. rufipogon*, 20 *indica*, and 16 *japonica* sampled from throughout the native range of wild and domesticated rice. When methods of demographic inference (using *∂a∂i*) were applied to this dataset (87), the results indicated that a single origin for domesticated rice was significantly more likely than multiple origins. The authors found that a single origin was more likely even when they excluded regions that showed evidence of selective sweeps; i.e., gene fragments that looked the most like domestication genes were not influencing the analysis. The paper also included a reanalysis of previously published sequence data using *BEAST (88), a Bayesian approach that generates a species tree based on heterogeneous gene trees: this removes the need for concatenation and potentially provides a more accurate phylogeny. Although the smaller datasets indicated a lack of monophyly, larger datasets (>5,000 bp, five or more loci) were strongly supportive of a single origin for domesticated rice. Overall, the paper is consistent with a single origin of *O. sativa* in the Yangtze Valley of China, and this corresponds to the majority opinion based on archaeological research. Although it has been pointed out that the modeling component of the analysis was influenced by the assumption of no structure in *O. rufipogon* (89), a situation that appears to be unlikely based on subsequent studies (90, 91), this paper opened many questions about genetic research into the origins of domesticated rice (were previous phylogenetic analyses wholly inaccurate? how should the contrasting results from the same starting data be interpreted?). In addition, the study served as a harbinger of a return to a focus on data from neutral loci in addition to domestication genes.

At nearly the same time, an analysis using an almost identical sampling strategy (22 samples each from *O. rufipogon*, *indica*, and *japonica* from throughout the native range of wild rice) applied whole-genome resequencing to evaluate the evolutionary history of domesticated rice (92). In contrast

to the methods of Molina et al. (86), He et al. (92) used a coalescent-based approach that showed strong support for separate origins of *japonica* and *indica* based on the majority of the genome, but also found that regions of low diversity (possibly indicative of a selective sweep) displayed a pattern that was most consistent with a single origin. Overall, the data were consistent with unique origins of *japonica* and *indica* accompanied by extensive gene flow between the subspecies at domestication loci. Clearly, given the similarities of these two studies, it would be optimal to see each analysis (*∂a∂i* vs. coalescent modeling) applied to each dataset. As it currently stands, it is difficult to say whether the different results are due to the differences in the type of data (gene fragments vs. whole genomes), the analytical methods, or the identity of the samples chosen in each case. The unique findings of Molina et al. have not yet been recovered in other studies, despite their potential importance for our understanding of rice domestication.

The large genetic datasets but relatively small sample sizes in these studies were countered in a study focusing firmly on the neutral genetic variation in *O. rufipogon* (90). The study used 42 gene fragments and 180 accessions (108 of which were *O. rufipogon*); this represented the largest combined genetic/sample size at the time (93). Analysis indicated that *O. rufipogon* showed strong population structure and fell into two groups; this corresponds well to previous studies. However, in a surprising finding with no precedent in the literature, the results also showed that *indica* was most closely related to wild rice from southern China, a location

that is usually associated with the origin of *japonica* rice. Even more surprisingly, *japonica* samples showed no close affinity to either of the *O. rufipogon* groups. These results open many questions about the origin of domesticated Asian rice and its relationship with its closest wild relative. The results have yet to be replicated, but indicate that a greater emphasis on sampling *O. rufipogon* may reveal unexpected patterns.

The use of resequenced rice genomes as a tool to understand the evolutionary origins of domesticated *Oryza* has expanded most recently with the publication and analysis of 446 resequenced *O. rufipogon* genomes and 1,083 *japonica* and *indica* genomes (91). This analysis provides one of the most extensive datasets exploring the genetic structure of *O. rufipogon*, which is critical for understanding the evolution of domesticated rice. The analyses based on genomewide variation indicated a close relationship between *japonica* and *O. rufipogon* from far southern China (in the Pearl River valley rather than the Yangtze River valley), and a similarly close relationship between *indica* and *O. rufipogon* from eastern India. When phylogenetic analyses were conducted based on variation at 55 regions that showed evidence of selective sweeps under domestication, both *indica* and *japonica* were most closely related to populations in the Pearl River valley. The authors therefore suggest a single origin of domesticated rice in far southern China, followed by dispersal to Southeast Asia and hybridization with local *O. rufipogon*. This study corresponds well to archaeological evidence relating to the origin of *indica*, but differs in that archaeological research has not indicated the

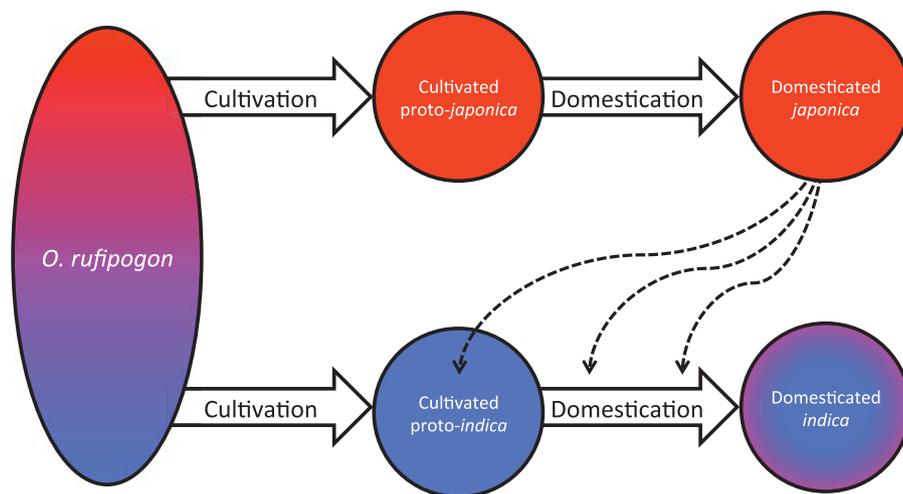


Fig. 2. Possible scenarios for the origin of *japonica* and *indica* from a genetically diverse *O. rufipogon* population, with genetic differentiation represented by different colors. Possible timing for the movement of domestication alleles from *japonica* into *indica* via hybridization is shown with dashed lines. Colors in the final domesticated *indica* circle represent the contributions from the original *O. rufipogon* populations (blue) and introgression from *japonica* (red).

Pearl River as a location for the domestication of *japonica*. One item to note about this paper is that, although the dataset is very impressive, the authors did not take advantage of powerful coalescent or other model-based techniques that could yield more precise of the population history of the domesticated and wild species (e.g., population size and timing of domestication).

Summary. Considering these recent studies, especially in light of the factors that might complicate our understanding of rice domestication, we can see real progress in some areas. As is the case for many other biological systems, the number of genetic markers in a given study has increased by orders of magnitude, so that limited sampling of genetic variation is no longer a concern. Sample size and diversity has not increased as steadily; indeed, there has been a trend toward decreasing sample sizes with an increasing number of molecular markers. Although necessary in some circumstances, this trend can hopefully be avoided in future studies of rice domestication, because sampling sparsely from a large geographical range is bound to introduce error in the form of missing populations and diversity. Finally, the integration of patterns seen at domestication genes and neutral loci is still being resolved. Although the possibility of introgression between *indica* and *japonica* is clearly well accepted, along with the genealogical discordance this can cause across different loci, it is still unclear how to relate this directly to the origin of *indica* and *japonica* (Fig. 2). In particular, should we conclude that there were two origins of domesticated rice and that phylogenetic incongruity is caused by introgression of domestication genes (92)? Or should we conclude that there was a single origin, with *indica* being brought to a domesticated state through a series of hybridization events between *japonica* and *O. rufipogon* populations (91)? The former scenario would suggest that some domestication alleles might have arisen in *indica*, but that these alleles were replaced by more desirable alleles from *japonica*. The later scenario would suggest that all domestication alleles arose exclusively in *japonica*. Which scenario is more likely currently remains unresolved, but the answer seems much closer now than it has in the past.

Conclusions

Domestication is not a single event, but a continuum, within which there are different degrees of codependence between humans and plants (94). Rice has moved far along the continuum (becoming a major food source for humans) and has been extensively modified

in this process, so that it now differs from its closest wild relative for a suite of traits that encompass life history, breeding system, morphology, and physiology. Given the exciting accumulation of new information from both archaeology and genetics about when and where *Oryza sativa* started along this continuum and how domestication traits eventually arose across the range of the species, how can the two fields better inform each other to answer key questions about the origin of *indica* and *japonica*?

Both fields offer unique approaches to identifying the geographical origin of rice domestication, and the bulk of studies indicate that *japonica* originated in the Yangtze River valley, whereas *indica* originated in the Ganges plains (although there are counter examples with interesting alternatives). Archaeological research indicates that appearance of one of the most iconic domestication traits (loss of shattering) was later than originally thought and that it was potentially not complete even by 6500 BP, although other traits (more difficult to detect from physical remains) may well have followed a different trajectory. Genetic approaches have shown that domestication traits are controlled by the same alleles in *japonica* and *indica* and that these alleles often originated in *japonica*, but that the subspecies are divergent at neutral loci. These findings from archaeology and genetics have been combined to suggest the independent origins of rice cultivation in China and India, followed by the introgression of domestication traits from *japonica* into proto-*indica* cultivated

plants to result in the establishment of the domesticated *indica* subspecies (47). This scenario, a type of “domestication by hybridization” for which there is increasing evidence from other plant and animal domesticates (95), is consistent with both the deep divergence between the subspecies at neutral loci and allele sharing at loci controlling key domestication genes.

One question about the domestication process with strong synergistic potential for the two fields deals with the process of domestication and the preferential spread of some domestication alleles over to others. For example, while the most common allele resulting in white rice grains (*rc*) originated in *japonica* and spread to *indica*, there are independent mutations resulting in an identical phenotype found in both the *aus* subspecies of *O. sativa* (closely related to *indica*) and the African domesticate *O. glaberrima* (69, 96). This pattern may indicate the incipient development of domestication traits in some populations, with further spread or development stopped by the arrival of a variety with more favorable characteristics. Although genetic surveys of extant lineages can provide some indication of the frequency of this type of phenomenon, archaeological evidence of the presence of domestication traits (e.g., shattering and grain shape) are required to complete our understanding of the history of rice domestication.

ACKNOWLEDGMENTS. This work is based on the National Evolutionary Synthesis Center (NESCent) catalysis meeting entitled “Domestication as an Evolutionary Phenomenon: Expanding the Synthesis.”

- 1 Vavilov NI (1926) Studies on the origin of cultivated plants. *Bull Appl Biol* 16:139–248.
- 2 Ding Y (1957) The origin and evolution of Chinese cultivated rice. *J Agriculture* 8(3):243–260.
- 3 Yan W (1989) Re-thinking the origin of rice agriculture in China. *Nongye Kaogu* 89(2):74–83.
- 4 Chang T (1976) Rice. *Evolution of Crop Plants*, ed Simmonds NW (Longman, London), pp 98–104.
- 5 Higham C (1995) The transition to rice cultivation in Southeast Asia. *Last Hunters—First Farmers*, eds Price T, Gebauer A (School of American Research Press, Santa Fe, NM), pp 127–155.
- 6 Zeder MA (2011) The origins of agriculture in the Near East. *Curr Anthropol* 52(4):S221–S236.
- 7 Watson P (1976) In pursuit of prehistoric subsistence: A comparative account of some contemporary flotation techniques. *Mid-Continental J Archaeol* 1:77–100.
- 8 Crawford G (1983) *Paleoethnobotany of the Kameda Peninsula Japon* (Museum of Anthropology, Ann Arbor, MI).
- 9 Pearsall D (1989) *Paleoethnobotany: A Handbook of Procedures* (Academic Press, San Diego).
- 10 Zhao Z (2005) Paleoethnobotany and its recent advances in China. *Kaogu (Archaeology)* 7:42–49.
- 11 Zhao Z (2011) New Archaeobotanical data for the study of the origins of agriculture in China. *Curr Anthropol* 52(4):S295–S306.
- 12 Fuller DQ (2006) Agricultural origins and frontiers in South Asia: A working synthesis. *J World Prehist* 20:1–86.
- 13 Ahn S (2010) The emergence of rice agriculture in Korea: Archaeobotanical perspectives. *Archaeol Anthropol Sci* 2(2):89–98.
- 14 Lee G (2011) The transition from foraging to farming in prehistoric Korea. *Curr Anthropol* 52(4):S307–S330.
- 15 Crawford G (2011) Advances in understanding early agriculture in Japan. *Curr Anthropol* 52(4):S331–S346.
- 16 Lu H, et al. (2002) Rice domestication and climatic change: Phytolith evidence from East China. *Boreas* 31(4):378–385.
- 17 Zheng Y, Matsui A, Fujiwara H (2003) Phytoliths of rice detected in the Neolithic sites in the valley of the Taihu Lake in China. *Environ Archaeol* 8(2):177–183.
- 18 Gu Y, Zhao Z, Pearsall D (2012) Phytolith morphology research on wild and domesticated rice species in East Asia. *Quat Int* 287:141–148.
- 19 Bellwood P (2005) *First Farmers: The Origins of Agricultural Societies* (Blackwell, Malden, MA).
- 20 Barker G (2006) *The Agricultural Revolution in Prehistory* (Oxford Univ Press, Oxford).
- 21 Crawford G (1992) Prehistoric plant domestication in East Asia. *The Origins of Agriculture: An International Perspective*, eds Cowan C, Watson P (Smithsonian Institution Press, London).
- 22 Nakamura S (2010) The origin of rice cultivation in the Lower Yangtze Region, China. *Archaeol Anthropol Sci* 2(2):107–113.
- 23 Cohen D (2011) The beginnings of agriculture in China: A multi-regional view. *Curr Anthropol* 52(4):S273–S293.
- 24 Fuller DQ, Sato YI (2008) *Japonica* rice carried to, not from, Southeast Asia. *Nat Genet* 40(11):1264–1265.
- 25 Price T, Bar-Yosef O (2011) The origins of agriculture, new data, new ideas: An introduction to supplement 4. *Curr Anthropol* 52(4):S163–S174.
- 26 Zhao Z (1998) The middle Yangtze region in China is one place where rice was domesticated: Phytolith evidence from the Diaotanghuan cave, northern Jiangxi. *Antiquity* 72(278):885–897.
- 27 Jiang L, Liu L (2006) New evidence for the origins of sedentism and rice domestication in the Lower Yangzi River, China. *Antiquity* 80(308):355–361.

- 28 Wu X, et al. (2012) Early pottery at 20,000 years ago in Xianrendong Cave, China. *Science* 336(6089):1696–1700.
- 29 Jiang L (2005) The Shangshan Neolithic site in Pujiang County, Zhejiang: New evidence of rice civilization in the Lower Yangtze River region (in Chinese). *Gudai Wenming Yanjiu Zhongxin Tongxun* 7:51–55.
- 30 Fuller DQ (2007) Contrasting patterns in crop domestication and domestication rates: Recent archaeobotanical insights from the Old World. *Ann Bot (Lond)* 100(5):903–924.
- 31 Fuller DQ, Harvey E, Qin L (2007) Presumed domestication? Evidence for wild rice cultivation and domestication in the fifth millennium BC of the Lower Yangtze region. *Antiquity* 81(312):316–331.
- 32 Hunan Institute of Archaeology (2006) *Pengtoushan and Bashidang* (Kexue Chubanshe, Beijing), (in Chinese).
- 33 Zhejiang Institute of Archaeology and Xiaoshan Museum (2004) *Kuahuqiao* (Kexue Chubanshe, Beijing), (in Chinese).
- 34 Zhang H, Wang H, Yang W (2005) *Early Neolithic Relics Found in the Xiaohuangshan Site in Shengzhou, Zhejiang* (Zhongguo Wenwu Bao) Chinese.
- 35 Henan Institute of Archaeology (1999) *Wuyang Jiahu* (Kexue Chubanshe, Beijing) Chinese.
- 36 Zhao Z, Zhang J (2009) The report of flotation work at the Jiahu site. *Kaogu (Archaeology)* 8:84–93.
- 37 Fuller DQ, Qin L, Harvey E (2008) Rice archaeobotany revisited: Comments on Liu et al. *Antiquity* 82(315):Project Gallery, online (www.antiquity.ac.uk/ProjGall/fuller1/).
- 38 Liu L, et al. (2008) Evidence for the early beginning (c. 9000 cal. BP) of rice domestication in China: A response. *Holocene* 17(8):1059–1068.
- 39 Liu J (2006) *The Hemudu Culture* (Cultural Relics Press, Beijing) Chinese.
- 40 Ellis JR, Pashley CH, Burke JM, McCauley DE (2006) High genetic diversity in a rare and endangered sunflower as compared to a common congener. *Mol Ecol* 15(9):2345–2355.
- 41 Research Center of Chinese Archaeology (2011) *Comprehensive Study of Natural Remains at the Tianluoshan Site* (Cultural Relics Press, Beijing) Chinese.
- 42 Fuller DQ, et al. (2009) The domestication process and domestication rate in rice: Spikelet bases from the Lower Yangtze. *Science* 323(5921):1607–1610.
- 43 Pan Y (2008) Immature wild rice harvesting at Kuahuqiao, China? *Antiquity* 82(316):Project Gallery, online (www.antiquity.ac.uk/projGall/pan/).
- 44 Imamura I (2007) Chronology of the Yayoi period. *News Letter No.7* (National Museum of Japanese History, Sakura, Japan), pp 6–7 Japanese.
- 45 Obata H (2008) Origins of domesticated plants in Jomon period, Japan. *Prehistoric and Ancient Cultigens in the Far East*, ed Obata H (Univ of Kumamoto, Kumamoto, Japan), Vol 3, pp 43–94. Japanese.
- 46 Fuller DQ, Allaby RG, Stevens C (2010) Domestication as innovation: The entanglement of techniques, technology and chance in the domestication of cereal crops. *World Archaeol* 42(1):13–28.
- 47 Fuller DQ (2011) Finding plant domestication in the Indian subcontinent. *Curr Anthropol* 52(S4):S347–S362.
- 48 Zhao Z, Gu H (2011) Study of identification methods of archaeological rice remains. *Hunan Kaogu Jikan* 8:5–68.
- 49 Oka H-I (1953) Phylogenetic differentiation of the cultivated rice plant. 1. Variations in respective characteristics and their combinations in rice cultivars. *Jap J Breeding* 3:33–43.
- 50 Oka H-I (1958) Intervarietal variation and classification of cultivated rice. *Indian J Genet Plant Breed* 18:79–89.
- 51 Garris AJ, Tai TH, Coburn J, Kresovich S, McCouch S (2005) Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169(3):1631–1638.
- 52 Matsuoka Y, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99(9):6080–6084.
- 53 van Heerwaarden J, et al. (2011) Genetic signals of origin, spread, and introgression in a large sample of maize landraces. *Proc Natl Acad Sci USA* 108(3):1088–1092.
- 54 Cheng C, et al. (2003) Polyphyletic origin of cultivated rice: Based on the interspersed pattern of SINEs. *Mol Biol Evol* 20(1):67–75.
- 55 Second G (1982) Origin of the genetic diversity of cultivated rice (*Oryza spp.*): Study of the polymorphism scored at 40 isozyme loci. *Jpn J Genet* 57:25–57.
- 56 Wang ZY, Second G, Tanksley SD (1992) Polymorphism and phylogenetic relationships among species in the genus *Oryza* as determined by analysis of nuclear RFLPs. *Theor Appl Genet* 83(5):565–581.
- 57 Zhu Q, Ge S (2005) Phylogenetic relationships among A-genome species of the genus *Oryza* revealed by intron sequences of four nuclear genes. *New Phytol* 167(1):249–265.
- 58 Londo JP, Chiang YC, Hung KH, Chiang TY, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc Natl Acad Sci USA* 103(25):9578–9583.
- 59 Caicedo AL, et al. (2007) Genome-wide patterns of nucleotide polymorphism in domesticated rice. *PLoS Genet* 3(9):1745–1756.
- 60 Rakshit S, et al. (2007) Large-scale DNA polymorphism study of *Oryza sativa* and *O. rufipogon* reveals the origin and divergence of Asian rice. *Theor Appl Genet* 114(4):731–743.
- 61 Gao LZ, Innan H (2008) Nonindependent domestication of the two rice subspecies, *Oryza sativa* ssp. *indica* and ssp. *japonica*, demonstrated by multilocus microsatellites. *Genetics* 179(2):965–976.
- 62 Gao LZ, Zhang CH, Jia JZ (2005) Cross-species transferability of rice microsatellites in its wild relatives and the potential for conservation genetic studies. *Genet Resour Crop Evol* 52(7):931–940.
- 63 Li C, Zhou A, Sang T (2006) Rice domestication by reducing shattering. *Science* 311(5769):1936–1939.
- 64 Konishi S, et al. (2006) An SNP caused loss of seed shattering during rice domestication. *Science* 312(5778):1392–1396.
- 65 Sang T, Ge S (2007) Genetics and phylogenetics of rice domestication. *Curr Opin Genet Dev* 17(6):533–538.
- 66 Zhang L-B, et al. (2009) Selection on grain shattering genes and rates of rice domestication. *New Phytol* 184(3):708–720.
- 67 Sang T, Ge S (2007) The puzzle of rice domestication. *J Integr Plant Biol* 49(6):760–768.
- 68 Ma J, Bennetzen JL (2004) Rapid recent growth and divergence of rice nuclear genomes. *Proc Natl Acad Sci USA* 101(34):12404–12410.
- 69 Sweeney MT, et al. (2007) Global dissemination of a single mutation conferring white pericarp in rice. *PLoS Genet* 3(8):e133.
- 70 Sweeney MT, Thomson MJ, Pfeil BE, McCouch S (2006) Caught red-handed: *Rc* encodes a basic helix-loop-helix protein conditioning red pericarp in rice. *Plant Cell* 18(2):283–294.
- 71 Olsen KM, Purugganan MD (2002) Molecular evidence on the origin and evolution of glutinous rice. *Genetics* 162(2):941–950.
- 72 Olsen KM, et al. (2006) Selection under domestication: Evidence for a sweep in the rice *waxy* genomic region. *Genetics* 173(2):975–983.
- 73 Wang Z-Y, et al. (1995) The amylose content in rice endosperm is related to the post-transcriptional regulation of the *waxy* gene. *Plant J* 7(4):613–622.
- 74 Hirano HY, Eiguchi M, Sano Y (1998) A single base change altered the regulation of the *Waxy* gene at the posttranscriptional level during the domestication of rice. *Mol Biol Evol* 15(8):978–987.
- 75 Kovach MJ, Sweeney MT, McCouch SR (2007) New insights into the history of rice domestication. *Trends Genet* 23(11):578–587.
- 76 Gross BL, Olsen KM (2010) Genetic perspectives on crop domestication. *Trends Plant Sci* 15(9):529–537.
- 77 Olsen KM, Wendel JF (2013) Crop plants as models for understanding plant adaptation and diversification. *Front Plant Sci*, 10.3389/fpls.2013.00290.
- 78 Asano K, et al. (2011) Artificial selection for a green revolution gene during *japonica* rice domestication. *Proc Natl Acad Sci USA* 108(27):11034–11039.
- 79 Shomura A, et al. (2008) Deletion in a gene associated with grain size increased yields during rice domestication. *Nat Genet* 40(8):1023–1028.
- 80 Takano-Kai N, et al. (2009) Evolutionary history of *GSS3*, a gene conferring grain length in rice. *Genetics* 182(4):1323–1334.
- 81 Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch SR (2009) The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc Natl Acad Sci USA* 106(34):14444–14449.
- 82 Tan L, et al. (2008) Control of a key transition from prostrate to erect growth in rice domestication. *Nat Genet* 40(11):1360–1364.
- 83 Jin J, et al. (2008) Genetic control of rice plant architecture under domestication. *Nat Genet* 40(11):1365–1369.
- 84 Gross BL, et al. (2010) Seeing red: The origin of grain pigmentation in US weedy rice. *Mol Ecol* 19(16):3380–3393.
- 85 Nichols R (2001) Gene trees and species trees are not the same. *Trends Ecol Evol* 16(7):358–364.
- 86 Molina J, et al. (2011) Molecular evidence for a single evolutionary origin of domesticated rice. *Proc Natl Acad Sci USA* 108(20):8351–8356.
- 87 Gutenkunst RN, Hernandez RD, Williamson SH, Bustamante CD (2009) Inferring the joint demographic history of multiple populations from multidimensional SNP frequency data. *PLoS Genet* 5(10):e1000695.
- 88 Drummond AJ, et al. (2012) Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol* 29(8):1969–1973.
- 89 Ge S, Sang T (2011) Inappropriate model rejects independent domestications of *indica* and *japonica* rice. *Proc Natl Acad Sci USA*, 108(39):E755, author reply E756.
- 90 Huang P, et al. (2012) Phylogeography of Asian wild rice, *Oryza rufipogon*: A genome-wide view. *Mol Ecol* 21(18):4593–4604.
- 91 Huang X, et al. (2012) A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490(7421):497–501.
- 92 He Z, et al. (2011) Two evolutionary histories in the genome of rice: The roles of domestication genes. *PLoS Genet* 7(6):e1002100.
- 93 Gross BL (2012) Rice domestication: Histories and mysteries. *Mol Ecol* 21(18):4412–4413.
- 94 Zeder MA (2006) Central questions in the domestication of plants and animals. *Evol Anthropol* 15(3):105–117.
- 95 Larson G, Burger J (2013) A population genetics view of animal domestication. *Trends Genet* 29(4):197–205.
- 96 Gross BL, Steffen FT, Olsen KM (2010) The molecular basis of white pericarps in African domesticated rice: Novel mutations at the *Rc* gene. *J Evol Biol* 23(12):2747–2753.

PNAS

Open Access

Providing immediate access to current scientific research.



Papers accepted for publication in PNAS can be made immediately available online without cost to the reader when the authors choose the open access option; all PNAS papers are available online for free 6 months after publication.

PNAS
www.pnas.org

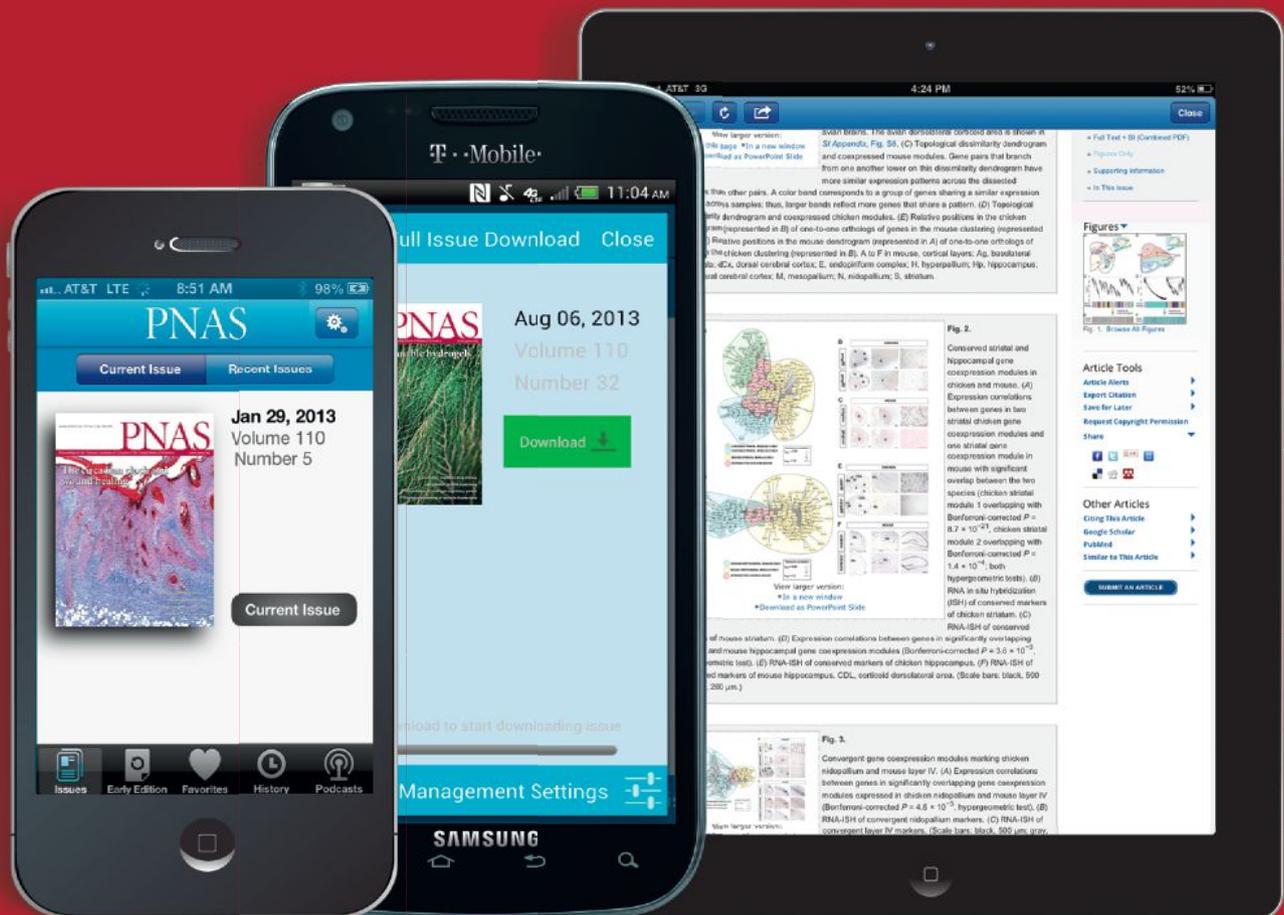
For more information visit www.pnas.org/site/aboutpnas/



The sun logo designates freely available content through the PNAS open access option.

PNAS goes Mobile

Version 1.02



Download the free PNAS iPhone, Android, or iPad app to easily browse and read the latest research published in PNAS every business day by accessing the PNAS tables of contents and full-text articles directly from your mobile device. New features include an in-app Web browser, various table and figure views, emailing options, and a favorites tab. PNAS content is available on your mobile device through either the improved full-text apps or on the mobile site at m.pnas.org.



For Apple



For Android



PNAS

www.pnas.org

