



University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Robin Allaby

Article Title: Integrating the processes in the evolutionary system of domestication

Year of publication: 2010

Link to published version:

<http://dx.doi.org/10.1093/jxb/erp382>

Publisher statement: *This is a pre-copy-editing, author-produced PDF of an article accepted for publication in Journal of Experimental Botany following peer review. The definitive publisher-authenticated version Allaby, R. (2010). [Integrating the processes in the evolutionary system of domestication. Journal of Experimental Botany, Vol. 61(4), pp. 935-944 is available online at:*

<http://jxb.oxfordjournals.org/cgi/content/abstract/61/4/935>

Integrating the processes in the evolutionary system of domestication.

Robin Allaby

Warwick HRI, University of Warwick, Wellesbourne, Warwick CV35 9EF.

Abstract

Genetics has long been used as a source of evidence to understand domestication origins.

A recent shift in the emphasis of archaeological evidence from a rapid transition paradigm of hunter-gatherers to agriculturalists to a protracted transition paradigm has highlighted how the scientific framework of interpretation of genetic data was quite dependent on archaeological evidence, resulting in a period of discord in which the two evidence types appeared to support different paradigms. Further examination showed that the discriminatory power of the approaches employed in genetics was low, and framed within the rapid paradigm rather than testing it. In order to interpret genetic data under the new protracted paradigm it must be taken into account how that paradigm changes our expectations of genetic diversity. Preliminary examination suggests that a number of features that constituted key evidences in the rapid paradigm are likely to be interpreted very differently in the protracted paradigm. Specifically, in the protracted transition the mode and mechanisms involved in the evolution of the domestication syndrome have become much more influential in the shape of genetic diversity. The result is that numerous factors interacting over several levels of organization in a domestication system need to be taken into account in order to understand the evolution

of the process. This presents a complex problem of integration of different data types which is difficult to describe formally. One possible way forward is to use Bayesian Approximation approaches that allow complex systems to be measured in a way that does not require such formality.

Introduction

Understanding the process of domestication is fundamental to understanding the transition from hunter-gathering to agriculturist, and is a model example of evolution in action which Darwin himself recognized as a primary source of insight to evolution (Darwin 1859, 1868). Although genetics has been widely studied to try to unravel the evolutionary events of the domestication process in the past few decades, much of the foundation of our understanding is actually drawn from archaeological evidence. This has become most apparent recently as archaeological evidence has caused a paradigm shift in agricultural origins research in the Fertile Crescent from a ‘rapid transition’ model to a ‘protracted transition’ model, which in turn exposed crop origin arguments based on genetic evidence to be dependent on rather than supportive of archaeological evidence. The protracted transition results in a number of corollaries for our interpretation of genetic diversity that are fundamentally different to the patterns of diversity expected under the rapid transition model. Here, consequences for the interpretation of genetic diversity because of this paradigm shift are explored.

The rapid transition paradigm

The process of domestication can be broadly divided into three stages (Harris 1989): 1) wild food procurement (hunting-gathering), 2) cultivation of wild plants (pre-domestication cultivation) and 3) domestication syndrome fixation (emergence of true agriculture). The first two stages are expected to be evident in the archaeological record as wild plant remains found in human settlements at levels higher than expected from any background population. The third stage occurs as phenotypic traits associated with domestication appear in the population, the so-called domestication syndrome traits, and proceed to become fixed. The domestication syndrome trait that has received most attention and is usually taken as the definitive indicator of a domesticated crop is the tough rachis phenotype in which grains are not shed from the plant due to a failure of the rachis to become brittle and shatter as the grain matures. Until recently, it was assumed that the first two stages were brief (Blumler 1992) as there was little evidence for pre-domestication cultivation found in the archaeological record, and the Younger Dryas, which occurred 11,500 years ago had long been considered as an upper limit for cereal growth because conditions would have been too harsh prior to this time (Moore and Hillman 1992, Wright 1976). This coupled with the early occurrence of tough rachis remains 9500 years BP (Hillman 1975) gives apparently only a brief window of opportunity for the pre-domestication cultivation stage.

The sudden appearance of domesticated crops in the archaeological record was further supported by classic field experiments (Fig. 1a) in which Hillman and Davies (1990) demonstrated by using sickle based harvesting proto farmers could have fixed the tough rachis trait within a human generation, 20 years or so. In their model the proportion of mutant phenotypes was enriched in the harvested fraction, where

harvesting took place at a point half-way through the natural grain shedding stage in the cereal's lifecycle, and most of the (wild type) grains that were shed to the ground were naturally predated. The result was the mutant increased in frequency as the next year's stand was of the progeny of the harvest, rather than naturally dispersed wild grain.

Genetic analysis within the rapid transition paradigm

A subtle but fundamentally important point in the development of research into crop origins is that when genetic diversity was used to investigate origins, hypotheses for testing were constructed entirely within the rapid transition paradigm. Classically, workers were interested in whether crops were domesticated once or multiple times. A single origin was associated with inventionism, indicating a single dominant farming group, whereas multiple origins were associated with diffusionism in which no one group would have dominated. It was widely suggested that if the rapid transition was true, then domesticated crops should be monophyletic (Zohary 1999), because a single crop would appear rapidly and there would be no need for further domestication. The implicit alternative is that if the rapid transition is not true, then crops should not be monophyletic but polyphetic, which has consequently been taken as the yardstick for multiple origins and a non-rapid transition. Extensive work has been devoted to establishing monophyly versus polyphyly with mixed results. Some gene phylogenies showed evidence of monophyly, such as in the case of flax (Allaby *et al.*, 2005), while others indicated polyphyly as in the case of rice and barley (Londo *et al.*, 2006, Molina-Cano *et al.*, 2005, Azguvel and Komatsuda 2007). However, the multiregionalism suggested by polyphyly still assumes a rapid process in which gene flow between

nascent crop populations does not occur. Genome-wide based evidence resoundingly supported the notion on single origins for many crops (e.g. Heun *et al.*, 1997, Badr *et al.*, 2000, Matsuoka *et al.*, 2002) by producing phenograms in which domesticated crops were monophyletic. This data was sometimes in conflict with single gene evidence, such as in the case of barley, which became awkward anomalies to what was largely taken as genetic support for the rapid transition paradigm, and a picture of explosive expansion of agriculturalists out of the centres of origin.

The protracted transition paradigm

Just as genetics had come to support the rapid transition paradigm, then new discoveries in archaeology completely changed the intellectual landscape, with all three stages of the domestication process seeing a massive reinterpretation. Evidence for wild plant gathering as early as 23,000 years ago has now been found, some ten millennia earlier than previously thought (Weiss *et al.*, 2004). This is in line with what has been known for a long time about animals; that after 50,000 years ago there was a considerable diversification of the human diet to include many small rodent species during a time of environmental pressure in the last glaciation, termed the broad spectrum revolution (Flannery 1969, Stiner 2001). Furthermore, evidence of the second stage, pre-domestication cultivation, has been established from 12,500 years BP (Weiss *et al.*, 2006, Willcox *et al.*, 2008), and possibly as early as 13,500 years (Willcox *et al.*, 2009), prior to the Younger Dryas. Within the pre-domestication period, there appear to have been numerous beginnings of agriculture, with different species hailing from different localities rather than in a single Neolithic Package (Willcox 2005). Even the last stage of

the domestication process has been shown to have been much more drawn out than previously supposed with the tough rachis mutant taking well over 3000 years to become fixed (Tanno and Willcox 2006). More recently, archaeological evidence from East Asia is also describing an older, more protracted origin of rice (Fuller *et al.*, 2009b) suggesting that the Near East may not be a special case.

The archaeological evidence now shows quite clearly that the transition from hunter-gatherer to agriculturist was not rapid at all but a protracted process spread over many millennia. This change in the landscape does rather leave the genetic evidence hanging awkwardly. How can it be that the genetic evidence supports a rapid transition when the archaeology now so clearly contradicts this scenario?

Assessing the information value of genetic diversity

The apparent conflict between archaeology and genetics was recently resolved using a modelling approach to ascertain what pattern of genetic diversity one would expect to see in terms of monophyly versus polyphyly in genome-wide studies given different starting scenarios of multiple or single domestication events (Allaby *et al.*, 2008). This model generated hypothetical plants, with chromosomes and biomarkers that formed breeding populations that could be subject to 'domestication'. The resultant populations could then be subject to phylogenetic analyses to determine if the 'domesticated' plants formed a monophyletic group. The model showed that both domesticate populations of multiple and single origin became monophyletic at a rate that was largely dependent on the population size, with monophyly being reached in around $2N$ generations, where N is the population size. A surprising find was that populations of multiple origins actually

reached monophyly faster than those of a single origin. Earlier coalescent-based studies demonstrated that domesticated crops probably experienced a bottleneck equivalent to a population size of 1500 individuals for 3000 generations (Zhu et al., 2004). An interesting correlation therefore occurs in which temporal dimension of the bottleneck appears to be similar to the fixation time for the tough rachis mutant, which was also enough time for domesticate populations to become monophyletic regardless of the number of origins. The conclusion is that monophyly has low information value for the interpretation of origins, and consequently this evidence does not conflict with the new archaeologically driven stance of a protracted transition. Not only that, the internal conflict within the genetic evidence falls away as genome-wide studies no longer contradict single gene studies either for the same reasoning.

So we can rest easy that genetic and archaeological evidence does not conflict, because the way in which genetic evidence was previously interpreted had low information value. This does leave a rather obvious hole. Where is the information value in the genetic diversity of crops? This is the question we must answer now in order to understand crop origins under a protracted transition. In order to do this we must reconsider what processes are likely to have influenced genetic diversity, beginning with asking the question why did domestication take so long?

Why did domestication take so long?

There are three possibilities that could explain a protracted nature of domestication in terms of selection pressure. Firstly, the original assertion that artificial selection was strong might have been wrong, and in fact selection was weak. A second possibility is

that net selection was weak. A third possibility is that selection or net selection was weak for a long time, and then became strong.

To consider the strength of artificial selection we can return to the classic work of Hillman and Davies (1990) in Fig. 1. There are three obvious points in their model that might be subject to variability resulting in (shown in italics in Fig. 1). The first is the point of harvest, which in the model occurs at the point at which fifty percent of grains have fallen from the plant. An alternative strategy may have been to harvest immature plants (green harvesting) to in order to maximize grain recovery (Fuller 2007). In this case the tough rachis mutant enrichment effect would be greatly reduced. Alternatively, it is argued that for a long time a sickle based harvesting would not have been employed at all, but rather gathering from the ground (Kislev *et al.*, 2004) in which case selection for the tough rachis mutant would have actually been negative because the tough rachis mutation would preclude such grains entering that population. The second point in the Hillman Davies model that might be subject to variability is the predation of fallen grain. The model assumes a rate of 90% predation, which serves to reduce the input of ground fall in which there is negative tough rachis selection relative to the amount sown from the harvest to the plant population of the following year. Before cultivation the background rate of ground fall predation would probably have been dependent on wild plant stand density. Therefore the onset of cultivation would have resulted in an increase in wild plant density, and predation would have lagged behind. If one assumes a sickle pressure process of harvesting, there may not have been significant enrichment of the tough rachis mutant until rodent populations increased as a consequence of cultivation. It is not the intention here to argue that domestication was driven by mice rather than men,

but to point out the effects of different selector agents on the strength and consequent pace of domestication. Beguilingly, however, the work of Willcox *et al.* (2009) does appear to show an increase in rodent droppings over time in Near Eastern sites in the early Holocene, although this is not something the authors themselves comment on. The third point at which the Hillman Davies model might vary is the proportion of crop that is sown from the harvest. Inevitably, the size of the harvest will have varied from year to year. In less productive years one might expect that a lower proportion of the harvest could be spared for sowing, again reducing the strength of tough rachis mutant selection.

An alternative possibility to weak selection pressure, is that the artificial selection pressure was in fact strong, but was not significantly stronger than natural selection (as was previously assumed in the rapid transition paradigm) resulting in a net weak selection. There are a relatively low number of domestication syndrome traits that we are aware of, the principal ones summarized in Fig 2. In each case the artificial selection environment and the natural selection environment are diametrically opposed resulting in an antagonistic struggle between the two (Allaby 2008). The protracted transition means that gene flow between these two environments, which may have been sympatric, could supply each environment with renewed input of alleles that were being selected against, ultimately preventing the rapid fixation of domestication syndrome traits in the artificial environment. It may even be the case that domestication was simply not possible within the wild biogeographic range. A similar scenario is described by Willcox (2005), in which the two environments are in fact allopatric, but gene flow from the wild to the cultivated environment (which occurs outside of the wild range) is facilitated by a 'restocking' behaviour by early cultivators.

The movement of cultivators in and out of the wild biogeographical range leads to the third possibility; selection was weak while cultivators were within the wild biogeographical range, but became strong as they left. In effect it may be the case that ‘domestication events’, if they could still be considered as such, were the result of human expansions. Such a scenario would predict genetically distinct expansions, which intriguingly has been observed in the phylogeographic patterns of emmer wheats (Brown *et al.*, 2006).

In summary, the slow pace of domestication can be explained by incorporating the effects of selectors in the domestication system. A consequence of these selectors is that the protracted transition has a critical difference to a rapid transition in terms of shaping genetic diversity in the extended period of gene flow between the artificial and natural selective environments. Furthermore, the larger timescale associated with the protracted transition leads us to question how these selective forces may have been separated temporally as well as spatially. In the following sections the effects of selectors and gene flow on the rise of the domestication syndrome are considered in a protracted transition.

The chronology of trait fixation

Under the rapid transition paradigm the domestication syndrome traits effectively arose together, which does not distinguish between different selective agents. Recent evidence from archaeology has changed this view (Fuller 2007). It is now apparent that different elements of the domestication syndrome became selected for at different times. The increase in seed size associated with domestication can clearly be seen to occur before the onset of tough rachis fixation in the case of einkorn (Willcox 2004), for instance.

Seed size increase is likely to be the result of selective pressure provided by sowing (Fuller 2007). Normal wild cereals drop grains on the ground surface where it is advantageous to be small and narrow to avoid predation and settle between soil particles. Once buried in the act of sowing, it becomes more advantageous for a grain to be larger in order to have the energy store required to break through the surface. Experimental evidence has shown that in the case of rice seed size increase can be effected in as little as five generations (Oka and Morishima 1971), supporting the notion that wild plants may have reacted quickly in terms of seed size to pre-domestication cultivation practises. This selection pressure is distinct from the one that caused increases in the frequency of the tough rachis mutation that is associated with harvesting techniques. Different selection pressures gave rise to different traits becoming fixed, and the order in which those selection pressures occurred would have affected the shape of the genetic diversity of the emergent crop. For instance, might traits that are fixed under a single selective pressure be more likely to be co-located in the genomic architecture than traits that were selected by different pressures at different times? Fuller et al. (2009a) explore the possibility of metastable crops of intermediate domestication (e.g. larger grains, no hooks or awns but still brittle rachised) occurring for substantial periods of time. An interesting observation in the comparison between grain size and the tough rachis mutant is that the former is polygenically controlled and the latter monogenically controlled. Intuitively, one might have assumed that the genetically simpler system would have been more easily domesticated. However, quite the reverse might actually be true because complex control of a trait allows for a stepwise increment in that trait which is not

immediately lost in the face of gene flow as in the case of the monogenic trait. Could it be a general rule that monogenic traits are harder to domesticate than multigenic traits?

The fate of parallelisms

An argument that has been powerful for a long time under the rapid transition is that multiple origins for crops should be evident in different mutations for the same trait occurring reflecting those different domestication events (Zohary and Hopf 2000). In the case of the tough rachis mutation, most crops have a single underlying mutation responsible for the trait, with the notable exception of barley that has two (Takahashi 1972, Azguvel and Komatsuda 2007), seemingly supporting the rapid transition paradigm. However, the genetic expectation becomes quite different under the protracted transition paradigm, Fig. 3. If the same trait were fixed by different mutations in two different regions resulting in the same phenotype, when those two cultures came into contact and the resulting crops mixed the initial population would indeed have two mutants. However, those mutants would be selectively neutral with respect to each other and their subsequent fates in terms of allele frequencies would be dictated by neutral processes. Over time genetic drift would remove one of those mutants. The length of time this would take would depend on the lowest mutant frequency. Given that the highest this value could be is 50%, an even mix of populations, then neutral theory tells us that one allele would be removed in approximately $2N$ generations, where N is the population size (Kimura and Ohta 1973). Consequently, one would not necessarily expect multiple mutations for traits after a protracted transition even though more than one may have occurred in the past, Fig 3a. Furthermore, a corollary of this expectation is

that if multiple different traits were fixed in parallel in different regions, then the resulting composite crop might be expected to have single mutations, but with different phylogeographic affiliations, Fig 3b. It may also be the case that when two crops that have had mutants for different domestication syndrome traits selected come into contact may be subject to a force of positive attraction in the artificial selection environment. This could occur if the hybrid progeny that retain both mutants benefit from an additive effect in fitness value in the artificial selection environment relative to the values of each of the progenitors. Such an effect could be termed ‘cultivation magnetism’, and may be reflected in selection signatures.

Recent evidence from rice studies echoes these genetic expectations tantalizingly. There are now several examples of domestication trait mutants that have arisen in parallel in the subpopulations in rice. In almost all cases so far, one of the two mutant types is present at very low frequency, consistent with its removal from the population. In the case of both the waxy (*Wx*) and the red pericarp (*Rc*) genes, mutants have arisen independently in the *indica* and *japonica* subpopulations, but 97% of all types are of the *japonica* type indicating introgression from *japonica* into the *indica* gene pool (Yamanaka et al., 2004, Sweeney et al., 2007). In the case of the fragrance gene (*BADH2*) one mutant allele predominates, again seemingly from the *japonica* subpopulation, but nine other mutants that also give rise to the fragrant trait were also found which are regionally restricted and apparently at low frequencies (Kovach *et al.*, 2009). These data in conjunction with the expectations outlined above suggest that possibly the minor frequency alleles may have been at a higher frequency in the past.

This prediction could be tested using archaeogenetic techniques to test for the presence of alleles in archaeobotanical assemblages.

Evidence for the corollary of the expectation of the fate of parallelisms outlined above, a mosaic genome of trait mutants, has also been seen with rice in the study of Shomura *et al.* (2008). In this case three different domestication trait mutants, waxy (*wx*), shattering (*qsh1*) and seed width (*qsw5*) appear to come from three different geographic regions, giving rise to a mosaic very much like that predicted in Fig 3b. The rice studies also indicate that there are multiple shattering mutants, *sh4* (Li *et al.*, 2006) and *qsh1* (Konishi *et al.*, 2006). In this case *sh4* is present in all tested domesticated rice varieties (Lin *et al.*, 2007), while *qsh1* is restricted to the Japanese area where it presumably arose, and occurred after the *wx* and *qsw5* mutants had arisen in different areas and been combined, possibly in the Thailand area (Shomura *et al.*, 2008).

The protracted transition further complicates the interpretation of parallelisms when spatially explicit population dynamics are considered. It has been generally assumed that mutants such as the tough rachis mutant are rapidly removed from the wild population (Zohary and Hopf 2000). Usually, domestication syndrome traits are associated with loss of function mutations, as in the case of the tough rachis mutant, which are consequently recessive. As a result complete removal from the wild population may take a very long time since heterozygotes display the wild phenotype. In such cases the frequency of a recessive allele can remain quite high in the population, but almost always in the heterozygous state – the classic case being cystic fibrosis in human populations where the frequency of the defective mutant can be as high as 2%, but homozygotes only occur with a frequency of about 0.05%. In the situation in the

protracted transition with artificial and natural selection environments can be adjacent for long periods of time, patch effects may occur in which alleles from one environment are constantly carried into the same local area of the other, Fig 4. Under such circumstances, normal mean field approaches to studying population genetics become inappropriate because they assume panmixis. In this example we have a scenario in which the mean field value of the frequency of a mutant allele derived from the artificial environment occurring in the wild natural selection environment is actually very low, but very high locally making local extinction less likely than the mean field would suggest. The combination of high frequency and re-supply may propagate mutant alleles in the ‘wrong’ environment for long enough for wild population bridges to be formed. In this case the same mutant may be selected for in different regions under culturally independent cultivation regimes. Interestingly, genetic evidence has recently been found for the converse, domesticated populations acting as genetic bridges for wild populations (Ross-Ibarra *et al.*, 2009).

Linkage disequilibrium and discerning introgression

In the protracted transition selection of domestication syndrome causing mutants happened over a longer period of time than is assumed in the rapid transition. A consequence of this is that in the face of wild gene flow linkage decay will be greater than has previously been assumed, Fig 5. To date a popular approach has been to establish the existence of linkage blocks around domestication syndrome trait mutant loci relative to the wild type loci (e.g. Kovach *et al.*, 2009). However, we have little understanding of what size of linkage block is expected under protraction. It could be the

case that the earlier traits have very much smaller linkage blocks that may even be hard to detect, whereas bold signals may only be true of the most recent trait selections. This dynamic impacts on the expected requirements of genome architecture. Le Thierry D'Ennequin *et al.* (1999) established through modelling approaches that genomes in which domestication syndrome trait loci are co-located should be associated with a higher probability of successful domestication than those in which such loci are more evenly dispersed. A resulting expectation of this is that one should expect traits in our domesticated species to occur together, and this general observation is indeed being revealed in various crops such as beans, millet rice and tomato (Gepts 2004). The effect of protraction is to accentuate this expectation, we should expect different loci to be even closer than we previously would under the rapid model. Furthermore, if there are interactions between loci both involving mutants then this effect might be expected to be even stronger.

An understanding of the size of linkage blocks that would be expected under a protracted transition may lead to an opportunity to address a long term recalcitrant problem in studies of crop origins, Fig 5; is it possible to distinguish between multiple origins and simple introgression from wild populations? The latter may occur between a crop that became domesticated in one region, and a wild population of a second region to which the crop had been translocated. The result would be a crop that has a mixed phylogeographic signature relating to both locations. Under the rapid transition distinguishing the two possibilities is extremely difficult, because both may be associated with large genomic fragments that could exhibit high linkage disequilibrium because of the lack of time to break down the linkage signature before such fragments have reached

fixation, particularly where extreme population bottlenecks are involved. However, under protraction there is more time for linkage disequilibrium to break down in the absence of selection resulting in two types of phylogeographic mosaic motif; large blocks associated with syndrome trait mutants from different regions, and small blocks associated with introgression. The question remains, how large is the former and how small the latter?

Integrating the domestication system

The consequence of the shift to the protracted transition paradigm is that an understanding of how domestication syndrome traits evolved and became selected has become key. Numerous interacting factors have become influential on the shape of genetic diversity in our crops that were not considered important previously, Fig 6. Moreover, these factors are distributed over several different levels of organization. The genetic control of traits, how genes interact with each other and therefore genome architecture, the spatial distribution of mutants and the biotic and abiotic selection agents all need to be considered. It has become too simplistic to consider one aspect in isolation in order to understand origins as a whole. However, quite different sorts of information are associated with these different levels of organization including genetic, archaeological, ecological and geographical, and the new challenge is to find meaningful ways to integrate them.

It is appropriate to think of the evolution of domestication as a complex system, and it is this system that one should aim to recapitulate in the interpretation of genetic diversity. It is fortunate that due to the rapid progress made in recent years that many

components to the system are understood, such as the molecular mechanisms underlying traits such as vernalization and photoperiod sensitivity (Yan *et al.*, 2003, 2004a, 2004b, 2006, Jones *et al.*, 2008) for instance. However, formal description of such a system is a problem too complex at the moment. One way forward is to use modelling approaches that can describe the interactions of the system, combined with the technique of Bayesian Approximation Computation (Beaumont *et al.*, 2002, Marjoram and Tavaré 2006). The theory behind Bayesian Approximation is that the probability of a model being true, given real summary statistic data, is proportional to the probability that the summary statistics generated by the model are acceptably close to those observed in reality. We can measure this latter probability without knowing the formal parametric distribution that describes it, which means that the approach is immensely powerful in complex situations. In practice, one can construct models that describe the system which have all the components one wishes to include; virtual plants with virtual genes in a virtual landscape subject to selection pressures by virtual farmers. Then various model histories can be explored, and those that produce summary statistics closest to the ones found in reality will be informative about which history is most probable. Furthermore, one can use this approach to explore which summary statistics are likely to be informative – we can answer the question, for instance, about how much linkage there should be around a domestication syndrome trait locus that was subject to selection early in the process of domestication. Or is it even possible for domestication to occur within the biogeographical range of the wild progenitor?

Model systems provide an opportunity to move our understanding of genetics to the next level and identify what patterns we should expect to see in the genetic data for

given histories, and which summary statistics have information value, and which do not as turned out to be the case in the use of the concept of monophyly. In this way genetic evidence should be predictive of archaeology, not the other way round as has been revealed to be the case with the rapid transition paradigm.

Acknowledgements

Thank you to many individuals for useful discussions in the development of this perspective particularly including Terry Brown and Dorian Fuller. Also Moshe Feldman, Wayne Powell, Susan McCouch, Michael Kovach and Roland Von Bothmer for useful discussions and encouragement.

Figure 1. Tough rachis selection model (after Hillman and Davies 1990).

A. Original model of Hillman and Davies in which harvest occurs when half the grains are shed. Sickle harvest enriches the proportion of tough rachis mutants from 0.1% to 0.2%. A higher proportion of grains from this source are then sown the following year (25%) than survive predation on the ground (10%), resulting 0.14% proportion of mutants being sown. B. Three points in the model that may be modified to result in lower or negative selection of the tough rachis mutant; (1) early harvest weakens initial tough rachis enrichment; (2) lower predation rate results in greater survival reducing the proportion of tough rachis sown; (3) higher consumption results in a lower proportion of grains available for sowing. All three examples result in a proportion of mutants lower than the original input leading to negative selection.

Figure 2. Artificial and natural selection of the domestication syndrome.

The traits domestication syndrome that are selected for in the artificial cultivation environment are diametrically opposed to the traits for the same characters in the natural environment. Gene flow between the two environments results in a 'struggle' for each trait.

Figure 3. The fate of parallelisms.

A. Different mutants (shown in red and blue respectively) resulting in the same trait are selected for and fixed in separate regions. Plants are then translocated to a third region in which the region 1 and 2 populations are amalgamated. The initial frequency of each mutant is 50% in the region 3 population but changes with genetic drift because neither mutant is preferentially selected for over the other. In $2N$ generations only a single mutant is expected to be retained. B. Four different loci governing different traits each of which has a different mutant fixed in a different region. Following from A, if these populations are amalgamated, three out of four mutants will be randomly lost through drift for each locus resulting in a random phylogeographic origin for each locus trait.

Figure 4. Fate of a domestication syndrome mutant in a spatially extended system.

A single mutant for a domestication syndrome trait is common in plants (brown squares) from region 1 (red). Gene flow between region 1 and the wild population (grey) leads to a local abundance of plants carrying the mutant, despite selection against it in the wild because of constant replenishment from region 1. The mutant does not spread far into the wild because of negative selection, and so is rare in the wild population as a whole., but

does spread into a second region of artificial selection pressure (blue) where it may be fixed independently to region 1.

Figure 5. Linkage disequilibrium and decay in the rapid and protracted transitions.

A mutant at a domestication syndrome trait locus (blue line) is initially associated with a linkage block (pink) on a chromosome region (green). Over time the linkage block decays as recombination occurs with other genotypes in the population in that region. In the protracted transition the increase in frequency of the trait is slowed by gene flow between the artificial and natural environments resulting in a longer period of linkage decay. After fixation no further decay is visible. A second trait locus (red line) is also selected. Selection strength is additive when the loci are on the same linkage block. In the extended decay of the protracted transition the second locus is broken away for the first losing the additive effect, suggesting that the two loci need to be closer than under the rapid transition.

Figure 6. The domestication evolution system.

The factors affecting the evolution of domestication are distributed over at least four different interacting levels of organization; the genetic mechanisms of control of domestication syndrome traits (green); the arrangement of genes within the genome (pink); the density of similar and different genotypes around individuals (brown); the selection pressures produced from the natural and artificial environments (blue).

References

Allaby RG (2008) The rise of plant domestication: life in the slow lane. *Biologist* 55: 94-99.

Allaby R.G., Fuller D.Q., Brown T.A. (2008) The genetic expectations of a protracted model for the origins of domesticated crops. *Proceedings of the National Academy of Sciences USA* 105:13982-13986.

Allaby RG, Peterson G, Merriwether DA, Fu Y-B (2005) Evidence of the domestication history of flax (*Linum usitatissimum* L.) from genetic diversity of the *sad2* locus. *Theor Appl Genet* 112:58-65.

Azguvel 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.* 100:1009-1015.

Badr A, Müller K, Schäfer-Pregl R, El Rabey H, Effgen S, Ibrahim HH, Pozzi C, Rohde W and Salamini F (2000) On the Origin and Domestication History of Barley. *Mol Biol Evol* 17:499-510

Beaumont MA, Zhang W, Balding DJ (2002) Approximate Bayesian Computation in population genetics. *Genetics* 162:2025-2035.

Blumer MA (1992) Independent inventionism and recent genetic evidence on plant domestication. *Econ. Bot.* 46:98-111.

Brown T.A., Lindsay S., and Allaby R.G. (2006) Using modern landraces of wheat to study the origins of European Agriculture. In *Darwin's Harvest*. Pp. 197-212. Eds. Motley, T.J., Zeregra, N., Cross, H. Columbia University Press, New York.

Darwin C (1859) Variation under domestication. In *Origin of Species*, Murray, London, pp 7-43.

Darwin C (1868) *Variation of Animals and Plants under Domestication*, 2 vols, Murray, London.

Flannery K (1969) Origins and ecological effects of early domestication in Iran and the Near East. P. 73-100 in *The domestication and exploitation of plants and animals*, eds. Ucko P and Dimbleby G. Duckworth, London.

Fuller D (2007) Contrasting Patterns in Crop Domestication and Domestiatiion Rates: Recent Archaeobotanical Insights from the Old World. *Annals of Botany* 100: 903-924.

Fuller DQ, Allaby RG, Stevens C (2009a) Domestication as Innovation: The entanglement of techniques, technology and chance in the domestication of cereal crops. *World Archaeology* in press.

Fuller DQ, Qin L, Zheng Y, Zhao Z, Chen X, Hosoya LA, Sun G-P (2009b) The Domestication process and domestication rate in rice: spikelet bases from the Lower Yangtze. *Science* 323: 1607-1610.

Gepts P (2004) Crop domestication as a long term selection experiment. *Plant Breeding Reviews* 24(2): 1-44.

Harris DR (1986). An evolutionary continuum of people-plant interaction. In *Foraging and Farming: the evolution of plant exploitation*. Eds. Harris DR, Hillman GC. Unwin Hyman, London. Pp 11-26.

Heun M, Schäfer-Pregl R, Klawan D, Castagna R, Accerbi M, Borghi B, Salamini F (1997) Site of einkorn wheat domestication identified by DNA fingerprinting. *Science* 278:1312-1314.

Hillman GC (1975) The plant remains from Tell Abu Hureyra: A preliminary report. *Proc. Prehist. Soc.* 41:70-73.

Hillman GC and Davies MS (1990) Domestication rates in wild-type wheats and barley under primitive cultivation. *Biological Journal of the Linnean Society* 39:39-78.

Jones H, Leigh F, Mackay I, Bower M, Smith L, Charles M, Jones G, Jones M, Brown T, Powell W 2008. Population based re-sequencing reveals that the flowering time adaptation of cultivated barley originated east of the Fertile Crescent. *Molecular Biology and Evolution* 25:2211-2219.

Kimura M, Ohta T (1973) The age of a neutral mutant persisting in a finite population. *Genetics* 75:199-212.

Kislev ME, Weiss E, Hartmann A (2004) Impetus for sowing and the beginning of agriculture: Ground collecting of wild cereals. *Proceedings of the National Academy of Sciences USA* 101:2692-2695.

Konishi S, Izawa T, Lin SY, Ebana K, Fukuta Y, Sasaki T, Yano M (2006) A SNP caused loss of seed shattering during rice domestication. *Science* 312: 1392-1396.

Kovach MJ, Calingacion MN, Fitzgerald MA, McCouch S (2009) The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proceedings of the National Academy of Sciences USA* 106:14444-14449.

Le Thierry D'Ennequin M, Toupance B, Godelle B, Gouyon PH (1999) Plant domestication: a model for studying the selection of linkage. *J Evol Biol* 12:1138-1147.

Li C, Zhou A, Sang T (2006). Rice domestication by reducing shattering. *Science* 311:1936-1939.

Lin Z, Griffith ME, Li X, Zhu Z, Tan L, Fu Y, Zhang W, Wang X, Xie D, Sun C (2007) Origin of seed shattering in rice (*Oryza sativa* L.). *Planta* 226:11-20.

Londo JP, Chiang, Y-C, Hung K-H, Chiang T-Y, Schaal BA (2006) Phylogeography of Asian wild rice, *Oryza rufipogon*, reveals multiple independent domestications of cultivated rice, *Oryza sativa*. *Proc Natl Acad Sci USA* 103:9578-9583.

Marjoram P, Tavaré S (2006) Modern computational approaches for analysing molecular genetic variation data. *Nat. Rev. Genet.* 7:759-770.

Matsuoka Y, Vigouroux, Y, Goodman, MM, Sanchez G, J, Buckler, E & Doebley, J (2002) A single domestication for maize shown by multilocus microsatellite genotyping. *Proc Natl Acad Sci USA* 99:6080-6084.

Molina-Cano JL, Russell JR, Moralejo MA, Escacena JL, Arias G and Powell W (2005) Chloroplast DNA microsatellite analysis supports a polyphyletic origin for barley. *Theor Appl Genet* 110:613-619.

Moore AMT, Hillman GC (1992) The Pleistocene to Holocene transition and human economy in Southwest Asia: The impact of the Younger Dryas. *American Antiquity* 57:482-494.

Oka H-I, Morishima H (1971) The dynamics of plant domestications: cultivation experiments with *Oryza perennis* and its hybrid with *O. sativa*. *Evolution* 25:356-364.

Shomura A, Izawa, Ebana K, Ebitani T, Kanegae H, Konishi S, Yano M (2008) Deletion in a gene associated with grain size increased yields during rice domestication. *Nature Genetics* 40:1023-1028.

Stiner M (2001) Thirty years on the “Broad Spectrum Revolution” and paleolithic demography. *Proc. Natl. Acad. Sci. USA* 98:6993-6996.

Sweeney MT, Thompson MJ, Cho YG, Park YJ, Williamson SH, Bustamante CD, McCouch S (2007) Global dissemination of a single mutation conferring white pericarp in rice. *Plos Genetics* 3:1418-1424.

Ross-Ibarra J, Tenaillon M, Gaut B (2009) Historical divergence and gene flow in the genus *Zea*. *Genetics* 181:1399-1413.

Takahashi R (1972) Non-brittle rachis 1 and non-brittle rachis 2. *Barley Genetics Newsletter* 2:181-182.

Tanno KI and Willcox G (2006) How Fast Was Wild Wheat Domesticated? *Science* 311: 1886.

Weiss E, Kislev ME and Hartmann A (2006) Autonomous Cultivation Before Domestication. *Science* 312: 1608-1610.

Weiss E, Wetterstrom W, Nadel D and Bar-Yosef O (2004) The broad spectrum revisited: Evidence from plant remains. *Proceedings of the National Academy of Sciences USA* 101: 9551-9555.

Willcox (2004) Measuring grain size and identifying Near Eastern cereal domestication: evidence from the Euphrates valley. *Journal of Archaeological Science* 31:145-150.

Willcox G (2005) The distribution, natural habitats and availability of wild cereals in relation to their domestication in the Near East: multiple events, multiple centres. *Vegetation history and archaeobotany*. 14: 534-541.

Willcox G, Fornite S and Herveux L (2008) Early Holocene cultivation before domestication in northern Syria. *Vegetation history ad archaeobotany* 17: 313-325.

Willcox G, Buxo R, Herveux L (2009) Late Pleistocene and Early Holocene climate and the beginnings of cultivation in northern Syria. *Holocene* 19:151-158.

Wright HE (1976) The Environmental Setting for Plant Domestication in the Near East. *Science* 194:385-389.

Yamanaka S, Nakamura I, Watanabe KN, Sato Y-I (2004) Identification of SNPs in the *waxy* gene among glutinous rice cultivars and their evolutionary significance during the domestication process of rice. *Theoretical and Applied Genetics* 108:1200-1204.

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J 2006. The wheat and barley vernalization gene *VRN3* is an orthologue of *FT*. *Proceedings of the National Academy of Sciences USA* 103:19581-19586.

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J 2003. Positional cloning of the wheat vernalization gene *VRN1*. *Proceedings of the National Academy of Sciences USA* 100:6263-6268.

Yan L, Helguera M, Kato K, Fukuyama S, Sherman J, Dubcovsky J 2004b. Allelic variation at the *VRN-1* promoter region in polyploidy wheat. *Theoretical and Applied Genetics* 109:1677-1686.

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, San Miguel P, Bennetzen JL, Echenique V, Dubcovsky J 2004a. The wheat *VRN2* gene is a flowering repressor down-regulated by vernalization. *Science* 303:1640-1644.

Zhu Q, Zheng X, Luo J, Gaut B and Song G (2007) Multilocus Analysis of Nucleotide Variation of *Oryza sativa* and Its Wild Relatives: Severe Bottleneck during Domestication of Rice. *Molecular Biology and Evolution* 24:875-888.

Zohary D (1999) Monophyletic vs. polyphyletic origin of crops found in the Near East *Genet Res Crop Evol* 46:133-142.

Zohary D and Hopf M (2000) *Domestication of Plants in the Old World. Third Edition.* Oxford University Press, Oxford.

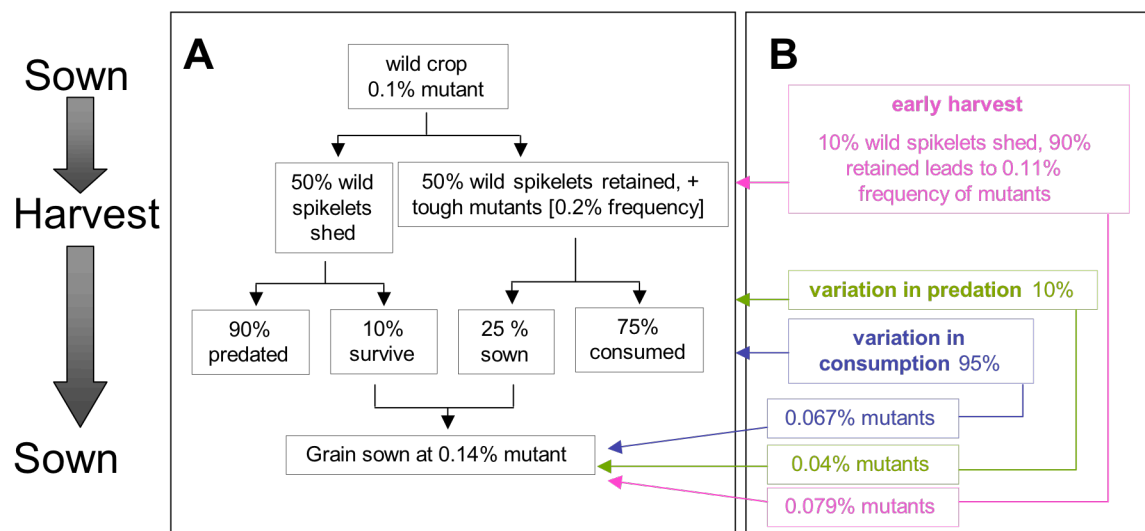


Figure 1. Tough rachis selection model after Hillman and Davies (1990).

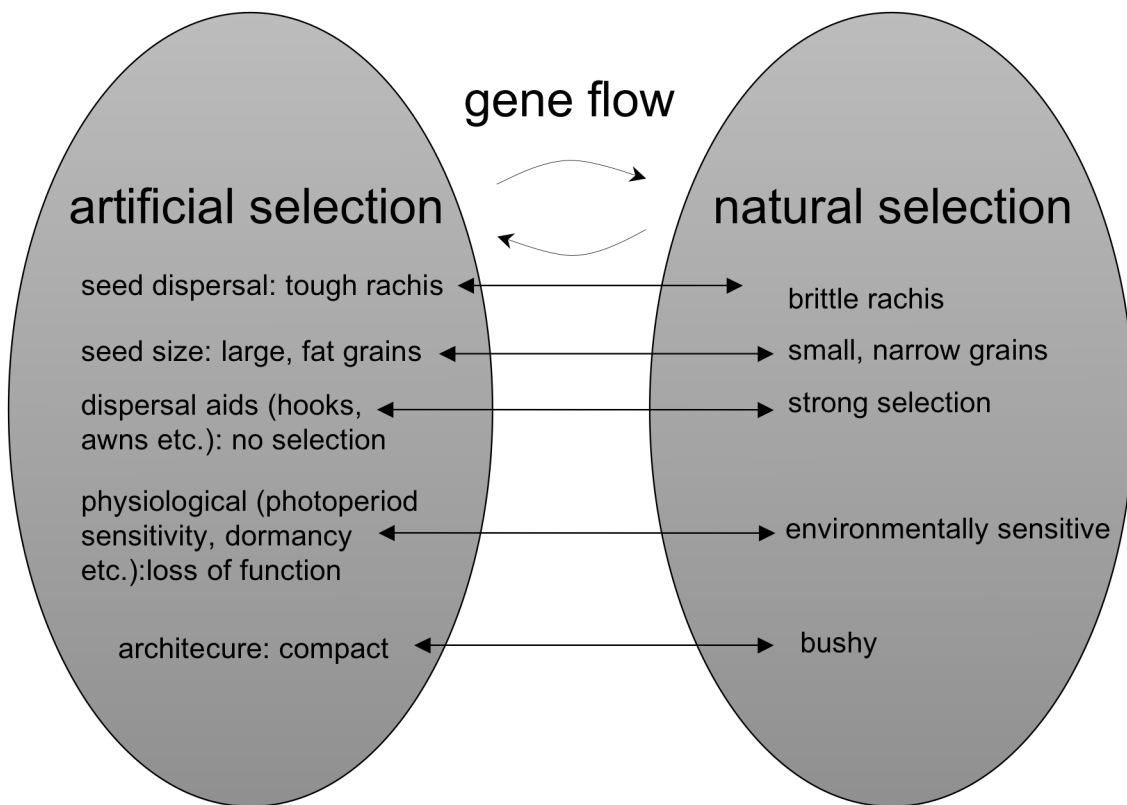


Figure 2.

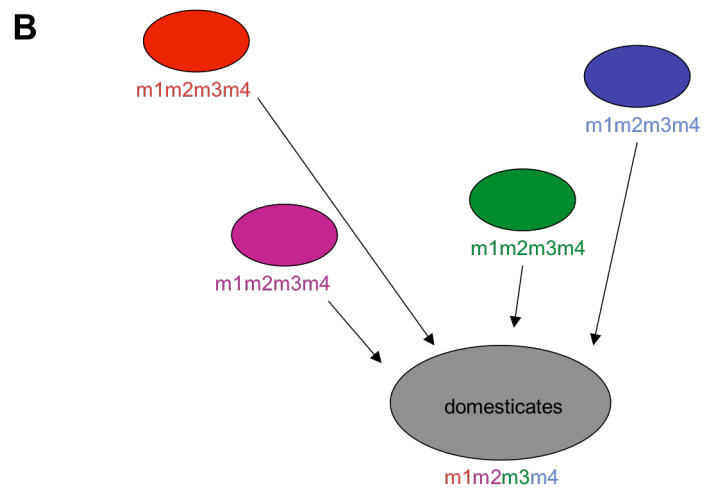
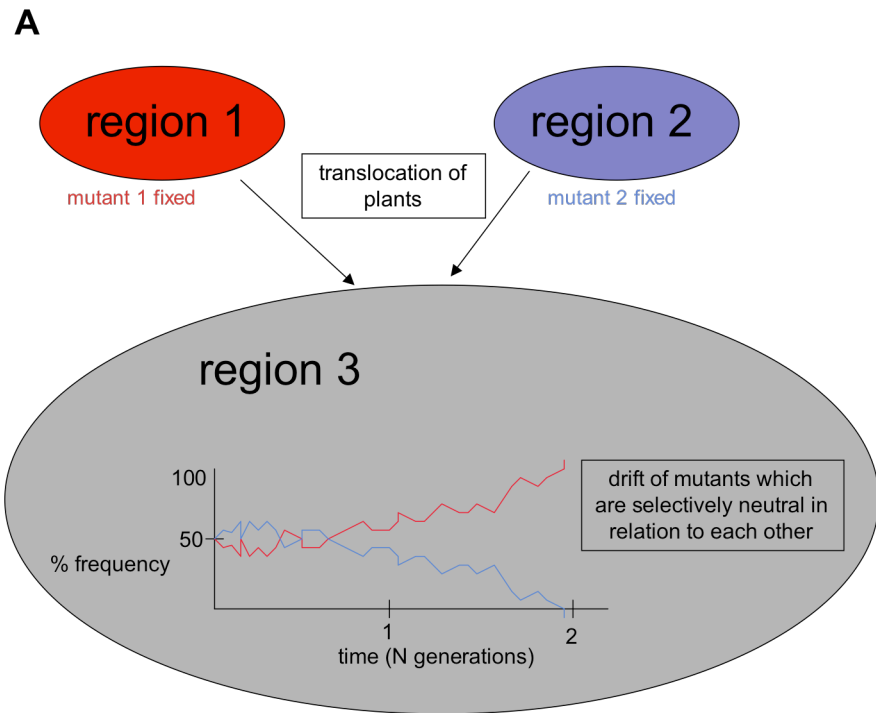


Figure 3.

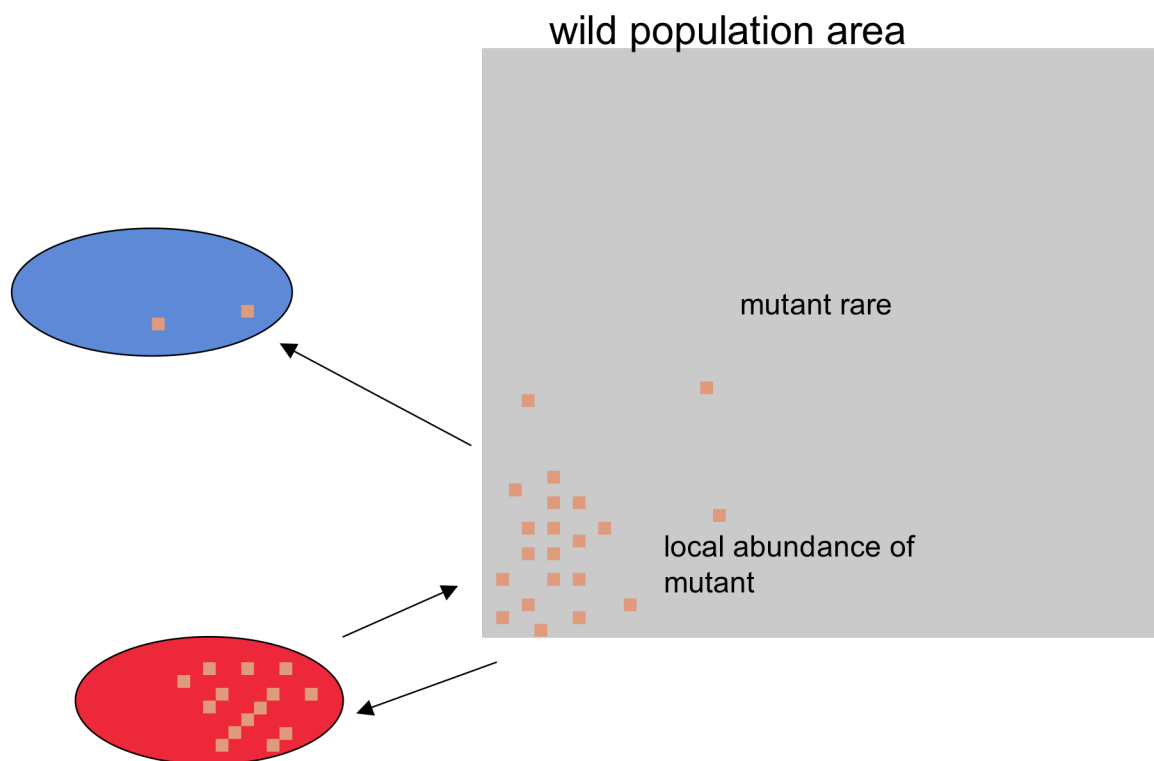


Figure 4

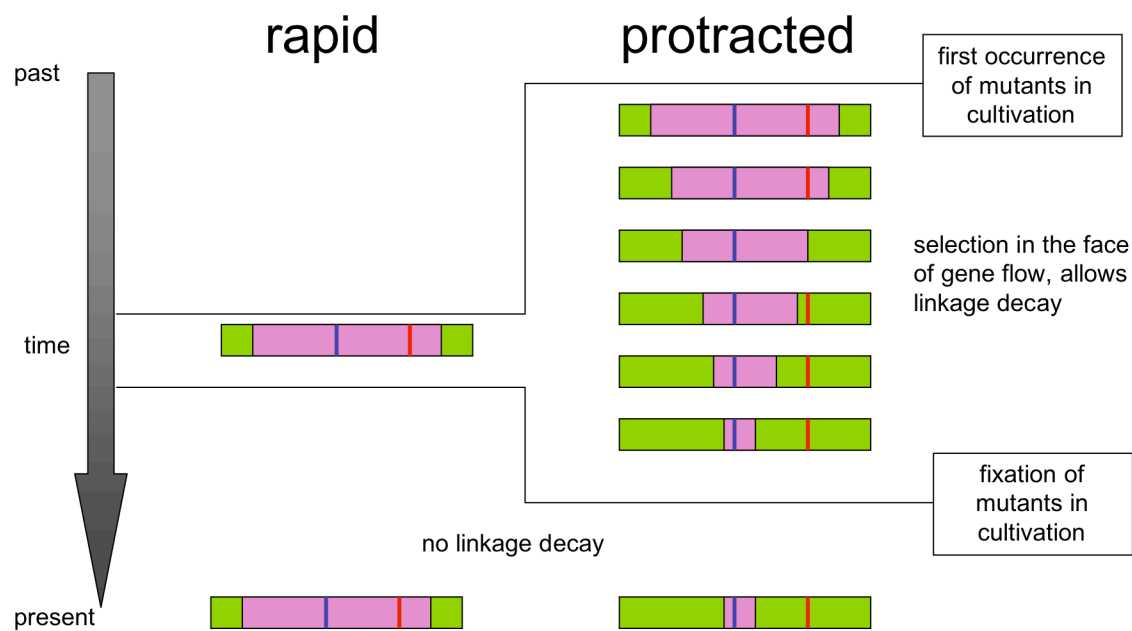


Figure 5.

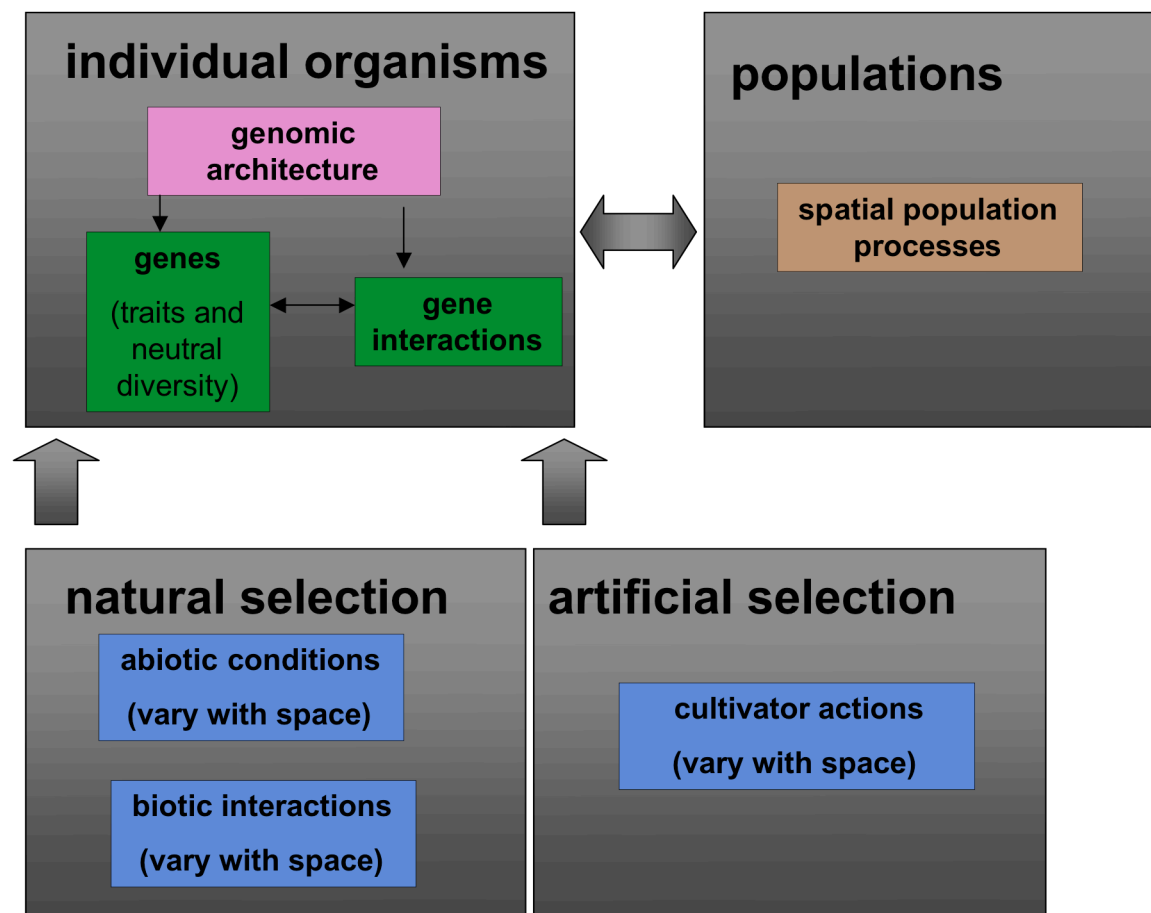


Figure 6