Evolution is a continuous trial and error process in which most lineages go extinct without leaving fossil remains. Many of these lineages would be closely related and occasionally hybridized with lineages that gave rise to extant species. Hence, it is likely that one can find genetic signatures of these ancient introgression events in present-day genomes, so-called ghost introgression. The increasing availability of high-quality genome assemblies for non-model organisms and the development of more sophisticated methods for detecting introgression will undoubtedly reveal more cases of ghost introgression, indicating that the Tree of Life is even more reticulated than assumed. The presence of ghost introgression has important consequences for the study of numerous evolutionary processes, including adaptation, speciation, and macroevolutionary patterns. In addition, detailed studies of introgressed regions could provide insights into the morphology of the extinct lineage, providing an unexpected link between genomics and the fossil record. Hence, new methods that take into account ghost introgression will need to be developed.

1. Introduction

Introgression—the exchange of genetic material through hybridization and backcrossing—is a common phenomenon across the Tree of Life.\(^{[1–4]}\) Numerous studies have reported ancient or contemporary gene flow between extant lineages. In some cases, introgression from an extinct lineage could be uncovered by sequencing ancient DNA from fossils. For example, the human genome contains genetic traces from at least two extinct hominins: the Neanderthals\(^{[5]}\) and the Denisovans.\(^{[6]}\) Similarly, modern-day brown bears (\(Ursus arctos\)) show signs of introgression from extinct cave bears.\(^{[7]}\) Given that fossilization is a rare phenomenon and that most fossils do not yield sufficient genetic material, we are probably missing numerous ancient introgression events that left traces of extinct species in present-day genomes (i.e., ghost introgression). Moreover, evolution by natural selection (and genetic drift) is a continuous trial and error process in which most lineages are lost without leaving fossil remains. Many of these lineages would be closely related and occasionally hybridized with lineages that lead to extant species. Hence, it is likely that we can find genetic evidence of these lost lineages in present-day genomes. The availability of high-quality genome assemblies and the development of more sophisticated introgression-detecting methods will allow us to search extant genomes for signatures of introgression from extinct species that are unknown in the fossil record. This exciting opportunity raises several questions: How common is ghost introgression? How often has ghost introgression been adaptive? Can we use the signatures of these ancient introgression events to reconstruct certain characteristics of the extinct species? Moreover, the presence of ghost introgression has important consequences for the study of evolutionary processes, such as adaptation, speciation, and macroevolution.

2. Ghostbusters: Detecting Ghost Introgression

A first clue for the presence of ghost introgression comes from analyses of organelles, such as mitochondrial DNA (mtDNA) and plastid DNA (Figure 1a). Patterns of deep divergence in organelar DNA in combination with shallow nuclear divergence might be due to mitochondrial or plastid capture from an extinct species.\(^{[8–10]}\) However, other demographic and evolutionary processes—including sex-biased gene flow,\(^{[11]}\) androgenesis,\(^{[12]}\) genetic drift\(^{[13]}\) or independent sorting of haplotypes in a large population\(^{[14]}\)—can culminate in similar patterns. Therefore, it is advisable to rule out these alternative explanations and provide additional evidence for introgression from an extinct lineage, such as signatures of ghost introgression in the nuclear genome. This approach is nicely illustrated by Leaf-warblers of the genus \(Phylloscopus.\) Comparison of mtDNA revealed that the Alpine Leaf-warbler (\(P. occisinensis\)) diverged from the Tickell's
Leaf-warbler (*P. affinis*), about four million years ago, whereas
the nuclear genome points to a speciation event roughly 600 000
years ago. The authors suggested that this mito-nuclear discrep-
ancy is due to introgression from an extinct *Phylloscopus*-lineage.
Subsequent analyses of the nuclear genome uncovered several
highly diverged regions that were considered remnants of the
ancient hybridization events that also resulted in the mtDNA
transfer.\[8\]

Signatures of ghost introgression can be limited to the nu-
clear genome where they might become apparent as highly di-
vergent haplotypes (Figure 1b). For instance, Ai and colleagues\[15\]
uncovered an exceptionally large (14 Mb) region with two diver-
gent haplotypes on the X-chromosome of Chinese wild boars (*Sus
scrofa*). A molecular clock analysis indicated that these haplotypes
diverged $\approx 8.5$ million years ago, much earlier than the known
evolutionary history of this pig species (dating back to about
5 million years ago). Additional analyses ruled out the possibility
of introgression from an extant species, supporting the hypoth-
esis of ghost introgression from an extinct pig species. Similar
to the study on wild boars, Kuhlwilm and co-workers\[16\] found
exceptionally divergent haplotypes in the bonobo (*Pan paniscus*)
genome. To test whether these haplotypes were introduced by in-
trogression with chimpanzees (*P. troglodytes*) or an extinct pri-
mate lineage, they computed the S* statistic, which distinguishes
between ancestral shared haplotypes and introgressed haplotypes
based on their length.\[17\] This approach confirmed previous work
on introgression between bonobos and chimpanzees\[18\] but also
found outlier regions (in terms of the S* statistic) that could not
be traced back to any chimpanzee population, suggesting intro-
gression from an extinct lineage. This finding was confirmed using
another method that uses a hidden Markov model to detect
archaic introgression tracts based on the local density of private
alleles.\[19,20\] Finally, they compared several demographic models
using an Approximate Bayesian Computation (ABC) approach,
revealing that bonobos received between 0.9% and 4.2% genetic
material from an extinct species. This study highlights the power
of combining several methods to detect ghost introgression.

Specifically, model-based approaches (Figure 1c) are becoming
increasingly powerful to confidently detect and quantify ghost in-
trogression, including methods based on the site frequency spec-
trum (SFS)\[21\] and demographic inferences drawn from whole
genome sequences (e.g., G-PhoCS).\[22\] More sophisticated meth-
ods are constantly being developed, as exemplified by recent ap-
proaches, such as reference-free detection of introgression\[23,24\]
and the combination of an ABC approach with deep learning.\[25\]

In summary, deeply divergent haplotypes often provide the
first clue for the presence of introgression from an extinct
lineage.\[8,15\] The hypothesis of ghost introgression can be fur-
ther explored with more sophisticated methods (e.g., S* statistic
and hidden Markov models) that pinpoint putatively introgressed
regions.\[17,19\] Finally, different demographic models can be com-
pared in a statistical framework to confirm or reject the existence
of ghost introgression.\[21–25\]

3. Ghost Introgression Changes Our View on the
Evolutionary Process

We have only recently been able to detect and quantify ghost in-
trogression. Hence, this concept has not been taken into account
when studying the evolutionary history of particular taxa. Ghost
introgression can have important implications in the study of
evolutionary processes, such as adaptation and speciation, and
in depicting phylogenetic relationships (e.g., a network approach
might be more suitable). In the following paragraphs, I discuss
some potential implications of considering ghost introgression
in evolutionary research.

3.1. Ghost Introgression can Speed up Adaptation and Speciation

Introgression can result in the transfer of beneficial alle-
les, i.e., adaptive introgression.\[26,27\] Obviously, the same

Figure 1. Genetic methods for detecting ghost introgression. a) Deep mitochondrial divergence despite shallow nuclear divergence. b) Highly divergent haplotypes, here visualized in a haplotype network. c) Testing different demographic with and without ghost introgression.
principle applies to ghost introgression. Although there are some frameworks to specifically search for signatures of adaptive introgression,[26–30] most studies combine statistical tests for introgression and positive selection to pinpoint cases of adaptive introgression. Indeed, several of the studies mentioned above provided evidence for positive selection on introgressed ghost alleles. For example, archaic introgressed regions in the bonobo genomes harbored several immunity and diet-related genes under positive selection, suggesting that ghost introgression might have contributed to adaptation to diseases and the exploration of novel food sources.[16] Similarly, the large introgressed region on the X-chromosome of wild boars shows signatures of a selective sweep, possibly underlying adaptation to different climates.[15] This genomic region might be one of the many examples of adaptive introgression in the evolutionary history of the wild boar. While expanding out of Asia, this species probably interbred with several other taxa, potentially incorporating locally adapted loci from these taxa into its genome.[31] This would have facilitated expansion and establishment into new areas where the wild boar consequently outcompeted the local species. This hypothetical scenario is reminiscent of the expansion of humans out of Africa. While spreading across the globe, archaic humans interbred with other hominins (e.g., Neanderthals and Denisovans) and human ghost populations[32,33] rapidly acquiring locally adapted loci and outcompeting the local populations. Hence, ghost introgression might help explain some of the most successful colonization histories in the distant past (Figure 2).

Ghost introgression can thus provide the raw material for rapid adaptation. Instead of waiting for de novo mutations, species could harness the potential of genetic variation introduced by hybridization with other taxa.[34] The introgressed regions could significantly speed up the rate of adaptation and speciation. Indeed, there is evidence that hybridization between divergent ancestral lineages was crucial for the explosive adaptive radiations of cichlid fish.[15–17] On a larger timescale, ghost introgression might also explain patterns of rapid evolution in the fossil record. For instance, cichleans (whales and dolphins) have evolved from land-dwelling into fully aquatic mammals in less than 50 million years,[38] and genomic analyses of extant cetaceans revealed a reticulated evolutionary history with high levels of introgression.[39] Could it be that ghost introgression between extinct lineages has sped up adaptation to an aquatic lifestyle? Most introgression events between cetacean species occurred during the fast radiation ofrorquals around 10.5 to 7.5 million years ago.[39] Whether ghost introgression contributed to more ancient radiations in cetacean evolution remains to be investigated. Similarly, ghost introgression might explain rapid diversification events in other mammalian groups, such as horses,[40] elephants,[41] and cats.[42]

Finally, ghost introgression could explain the rapid evolution of reproductive isolation between certain taxa: introgression between two species might contribute to increased divergence and potentially reproductive isolation between those species and other related species. For example, a study of the *Picea likiangensis* spruce species complex found that 32 genes introgressed between the taxa *complanata* and *likiangensis*. At these genes, this pair of taxa showed significantly increased divergence from another taxon (*linzhiensis*). In addition, 24 genes introgressed between the taxa *rubescens* and *linzhiensis*, culminating in increased divergence from a third taxon (*likiangensis*).[43] If these introgression patterns were not taken into account, divergence between some taxa might seem to have been established extremely rapidly when considered in the context of classic speciation models. Introgression from an extinct lineage can similarly affect estimation of divergence times, leading to wrong conclusions about the evolutionary history of particular taxa. For the moment, it is difficult to assess how important the role of ghost introgression is in rapid adaptation and speciation events. But the combination of genomic data and more sophisticated modelling techniques will allow researchers to test the hypotheses outlined above.

**3.2. Macroevolutionary Studies Should Take Ghost Introgression into Account**

Because the widespread occurrence of introgression cannot be captured in a classic bifurcating tree, several authors have argued for a phylogenetic network approach.[44–46] The observation of ghost introgression suggests that the Tree of Life might be even more reticulated than expected, questioning whether even phylogenetic networks will be able to capture these complex evolutionary patterns.[47] In addition, the reticulated nature of the evolutionary process has important consequences for macroevolutionary studies that generally assume a fixed species tree to reconstruct the emergence and disappearance of particular traits. What if a certain trait does not follow the species tree but has been horizontally transferred from extinct lineages? The likelihood of a trait being transferred by ghost introgression partly depends on the genetic basis of this trait: it is more likely that a trait with a simple genetic basis (e.g., single locus) is exchanged compared to a polygenic trait. In *Heliconius* butterflies, for instance, the genes encoding wing patterns have been transferred between several species,[48] thwarting the estimation of a bifurcating species tree and reconstructing the evolutionary history of wing patterns.[49] Not taking into account the possibility of ancient
introduction events might lead to errors in the reconstruction of trait evolution (Figure 3). Hence, new macroevolutionary and phylogenetic network methods accounting for introgression from extinct lineages will need to be developed.

3.3. Using Ghost Introgression to Reconstruct Fossil Phenotypes

Apart from gaining a more fine-grained and nuanced understanding of the evolutionary history of present-day species, we could use the information from ghost introgression to reconstruct certain characteristics of the extinct lineages. Recently, Gokhman and colleagues used methylation patterns in the Denisovan genome to reconstruct the putative appearance of this illustrious hominin. Potentially, detailed studies of introgressed regions could provide insights into the morphology of the extinct lineage, perhaps even allowing paleontologists to link the ghost lineage to known fossil remains.

4. Conclusions

Several studies have provided evidence that extinct species can live on in the genomes of present-day species. The availability of high-quality genome assemblies and the development of more sophisticated techniques to detect introgression will result in the discovery of more ghost introgression events. This concept will need to be incorporated into the study of adaptation, speciation, and macroevolution, ultimately leading to a better understanding of the reticulated nature of the evolutionary process.

Acknowledgements

The author would like to thank Per Alström for giving him the idea and inspiration to write this paper. The author is also thankful to Editor-in-chief Andrew Moore and two anonymous reviewers whose comments and suggestions improved this manuscript.

Conflict of Interest

The author declares no conflict of interest.

Keywords

adaptation, demographic modelling, hybridization, macroevolution, phylogenetic networks, reproductive isolation, speciation