TITLE

Divergence-based introgression polarization

AUTHORS

Evan S. Forsythe[1,2*], Daniel B. Sloan[1], and Mark A. Beilstein[2]

AFFILIATIONS

[1]Department of Biology, Colorado State University, Fort Collins, CO 80523, USA.

[2]School of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA.

CORRESPONDENCE

Evan. S. Forsythe; 1878 Campus Delivery, Fort Collins, CO 80523; esfors@rams.colostate.edu

ABSTRACT

Introgressive hybridization results in the transfer of genetic material between species, often with fitness implications for the recipient species. The development of statistical methods for detecting the signatures of historical introgression in whole-genome data has been a major area of focus. While existing techniques are able to identify the taxa that exchanged genes during introgression using a four-taxon system, most methods do not explicitly distinguish which taxon served as donor and which as recipient during introgression (i.e. polarization of introgression directionality). Existing methods that do polarize introgression are often only able to do so when there is a fifth taxon available and that taxon is sister to one of the taxa involved in introgression. Here, we present *Divergence-based Introgression Polarization* (*DIP*), a method for polarizing introgression using patterns of sequence divergence across whole genomes, which operates in a four-taxon context. Thus, *DIP* can be applied to infer the directionality of introgression when additional taxa are not available. We use simulations to show that *DIP* can polarize introgression and identify potential sources of bias in the assignment of directionality, and we apply *DIP* to a well-described hominin introgression event.

INTRODUCTION

Hybridization is an influential evolutionary force (Stebbins 1968) that is widespread in natural populations (Yakimowski and Rieseberg 2014; Mallet et al. 2016). Through backcrossing to parental populations, hybrids can serve as bridges for the transfer of alleles and adaptive traits between species or populations, a process known as introgression (Rieseberg and Soltis 1991; Rieseberg et al. 1996; Green et al. 2010; Dasmahapatra et al. 2012; Mallet et al. 2016; Suarez-Gonzalez et al. 2016). Whole genome sequences and advances in phylogenetic methods (Soltis and Soltis 2003) have revealed signatures of historical introgression in scientifically and economically important groups, including well-studied examples in Neanderthals and non-African human populations (Green et al. 2010; Kuhlwilm et al. 2016). Several methods have been developed to identify taxa that exchanged genes during introgression (Huson et al. 2005; Than et al. 2008; Green et al. 2010; Durand et al. 2011; Liu et al. 2014; Martin et al. 2015; Pease and Hahn 2015; Stenz et al. 2015; Rosenzweig et al. 2016). While these methods generally perform well across a variety of biological and experimental scenarios (Zheng and Janke 2018),

theoretical and empirical studies have identified conditions under which each method is susceptible to bias (Eriksson and Manica 2012; Rosenzweig et al. 2016).

One challenging aspect of analyzing introgression is to identify taxa serving as donors *vs.* recipients of genetic material during introgression (i.e. introgression directionality). If hybrids successfully backcross to both parents, alleles will move in both directions, meaning each parent will serve as donor for some introgressed loci and recipient for other loci. However, if backcrosses with one parent but not the other are favored by physiological (Rieseberg and Soltis 1991), selective (Orive and Barton 2002), or biogeographical (Currat et al. 2008) factors, it can lead to asymmetrical (Barton and Hewitt 1985) movement of alleles (directional introgression, denoted hereafter with ⇒). Introgression has been shown to underlie the transfer of adaptive traits to recipient lineages (Whitney et al. 2006; Dasmahapatra et al. 2012; Dannemann et al. 2016; Figueiró et al. 2017), so the ability to infer the directionality of introgression (i.e. polarize introgression) is essential in order to form hypotheses about functional and adaptive consequences.

The majority of tests to detect the occurrence of introgression do not explicitly polarize directionality (Zheng and Janke 2018), and those that can only do so in certain cases. For example, the *D* statistic (Green et al. 2010) is widely-used to infer instances of introgression in a four-taxon system. Introgression polarization is possible under *D* only when data for a fifth taxon are available (Green et al. 2010; Eaton and Ree 2013; Eaton et al. 2015; Pease and Hahn 2015). Moreover, the fifth taxon must be sister to one taxon involved in introgression but cannot itself be involved in introgression. (Pease and Hahn 2015) define this specific configuration of introgressing taxa and sister taxa as “intergroup” introgression and describe how, when these specific five-taxon conditions are met, the branching order of introgressed gene trees indicates directionality. However, the authors also describe how other types of introgression (e.g. “ancestral” introgression) cannot be polarized. Moreover, there are many cases in which a fifth taxon with the required phylogenetic placement is either not sampled or does not exist. In these cases, it is possible to statistically identify introgression using existing methods but not necessarily to polarize introgression. Thus, there is a need for a more widely applicable statistical method to distinguish between bidirectional and unidirectional introgression, while identifying donor and recipient taxa.

Here, we describe and test a method for inferring directionality of introgression from genome-scale data, which we refer to as *Divergence-based Introgression Polarization* (*DIP*). *DIP* is based on the observation that, when introgression occurs, it alters not only the level of nucleotide sequence divergence between the two species exchanging genes (Rosenzweig et al. 2016) but also divergences with related species that are not directly involved in introgression; these changes occur in systematic and predictable ways according to the directionality of introgression (Fig. 1) (Forsythe et al. in revision; Fontaine et al. 2015; Hibbins and Hahn 2019). *DIP* is calculated from pairwise sequence divergence between taxa involved in introgression and a sister taxon, comparing divergence values obtained from introgressed loci *vs*. non-introgressed loci. It takes as input the same types of data used to infer introgression by existing methods (whole genome/chromosome alignments or single-gene alignments of loci sampled throughout the genome). However, unlike most existing methods, *DIP* is applicable to cases in which only four taxa are sampled, thereby expanding inference of introgression directionality to a broader scope of evolutionary histories.

We present tools to implement the *DIP* method: https://github.com/EvanForsythe/DIP. We also simulate whole genome alignments in which a subset of loci were introgressed either unidirectionally, asymmetrically, or symmetrically. We use these simulated genome alignments to assess how accurately *DIP* polarizes asymmetrical introgression and to investigate the effects of parameters that are known to affect existing introgression inference methods, such as the proportion and timing of introgression (Durand et al. 2011; Martin et al. 2015; Zheng and Janke 2018). We have recently used the principles of *DIP* to document asymmetrical introgression among Brassicaceae species (Forsythe et al. in revision), and here, we also apply *DIP* to empirical data from modern and archaic hominins.

NEW APPROACHES

Introgression alters levels of sequence divergence between taxa, and these changes can differ depending on directionality (Forsythe et al. in revision; Hibbins and Hahn 2019) (Fig. 1). While several statistics focus on the effects of introgression on sequence divergence between species involved in introgression (Feder et al. 2005; Joly et al. 2009; Rosenzweig et al. 2016), here we describe how patterns of sequence divergence in a taxon that is sister to those involved in introgression can be indicative of the directionality of introgression. To define the properties of a

divergence-based introgression test, we use hypothetical species *P1*, *P2*, *P3* and an outgroup, *O*. Species *P1* and *P2* are sister within the species tree, and we model introgression between species *P2* and *P3*. We denote the timing of the three successive speciation events among these taxa as T_{γ}, T_{β}, and T_{α} and the timing of the introgression event between *P2* and *P3* as T_{INT} (Fig. 1A). When introgression has occurred between *P2* and *P3,* some loci will reflect a history of introgression, while other loci will reflect a history of speciation. In applying *DIP*, a gene tree is inferred for each locus, and the resulting topology is used to distinguish introgressed loci from non-introgressed loci. For all loci, we quantify pairwise sequence divergence values between *P2* and *P3* (K_{23}), between *P1* and *P2* (K_{12}), and between *P1* and *P3* (K_{13}) (Fig. 1). The values of K_{23}, K_{12}, and K_{13} on a given gene tree are expected to correspond to T_{INT}, T_{α}, and T_{β} in a way that depends on the introgression history of that gene. Note that K_{23} is the divergence measurement that is most commonly used to indicate the presence of introgression (Feder et al. 2005; Joly et al. 2009; Rosenzweig et al. 2016) because introgression in either direction is expected to reduce K_{23} relative to genes that reflect the species tree, as the divergence time between the sequences of these taxa is reduced from T_{β} to T_{INT} (Fig. 1). In contrast, changes in K_{12} and K_{13} will depend on the direction of introgression. For example, introgression can cause K_{12} to increase corresponding to a change in divergence time from T_{α} to T_{β} but only if introgression occurred from *P3* to *P2* (Fig. 1B). Introgression in the other direction should not affect K_{12}. The effects on K_{13} are also sensitive to the direction of introgression. If it occurs from *P2* to *P3*, introgression should decrease K_{13} based on a change in divergence time from T_{β} to T_{α} (Fig. 1C), but there should be no effect on K_{13} if introgression occurs in the other direction. To quantify these effects, differences are calculated between the mean values of K_{23}, K_{12}, and K_{13} from all loci displaying the species topology (abbreviated SP loci in equations/figures) and the mean values of the same corresponding divergence measurements from all loci displaying the introgression topology (abbreviated INT loci in equations/figures) in the following fashion:

Eq. 1:

$$\Delta K_{23} = \overline{K}_{23}(SP\ loci) - \overline{K}_{23}(INT\ loci)$$

Eq. 2:

$$\Delta K_{12} = \overline{K}_{12}(INT\ loci) - \overline{K}_{12}(SP\ loci)$$

Eq. 3:

$$\Delta K_{13} = \overline{K}_{13}(SP\ loci) - \overline{K}_{13}(INT\ loci)$$

Note that the order of subtraction used in defining these terms is not always the same with respect to species and introgression loci and was chosen such that the effects of relevant introgression are expected to yield positive (rather than negative) *ΔK* in each case. Together, this set of *ΔK* values composes the divergence profile of *DIP*. Below we show the relative magnitudes of these values can be used to differentiate evolutionary histories based on the polarity of introgression. We also use coalescent-based simulations to identify biases that can be introduced by other sources of genealogical discordance such as incomplete lineage sorting (ILS), and we devise additional layers of *DIP* comparisons that can be used to partially alleviate these biases.

RESULTS

DIP: Distinguishing modes of unidirectional and bidirectional introgression

The simplest application of *DIP* involves testing whether ΔK_{23}, ΔK_{12}, and ΔK_{13} are significantly greater than zero and compares these results to the expectations for *ΔK* under different introgression scenarios (Fig. 2). If introgression has occurred in both directions between *P2* and *P3*, then all three *ΔK* values should be positive. However, as noted above, if introgression has occurred exclusively in one direction, the expectation for either ΔK_{12} or ΔK_{13} should remain zero (Fig. 2). To test the performance of *DIP*, we simulated alignments for thousands of loci (5000 bp each) undergoing unidirectional introgression in each direction, as well as symmetric bidirectional introgression (see Methods and Fig. S1). We applied *DIP* to each simulated genome. For the genome simulated under unidirectional *P2⇒P3* introgression, we observed $\Delta K_{23} > 0$, $\Delta K_{12} = 0$, and $\Delta K_{13} > 0$ (Fig. 3A), which is the expected pattern for that direction of introgression (Fig. 1). For the genome simulated under symmetric bidirectional introgression, we observed $\Delta K_{23} > 0$, $\Delta K_{12} > 0$, and $\Delta K_{13} > 0$ (Fig. 3B), which is the expected pattern if some introgression is occurring in both directions. For the genome simulated under unidirectional *P3⇒P2* introgression, we observed $\Delta K_{23} > 0$, $\Delta K_{12} > 0$, and $\Delta K_{13} = 0$ (Fig. 3C), again reflecting

our expected *DIP* profile for that direction. These results indicate that *DIP* can correctly classify all three types of introgression under these simulated conditions.

Next, we explored the performance of *DIP* across a range of different parameter settings, including the proportions of genes in the genome that had been subject to introgression (*pINT*). We also varied the proportions of introgressed loci that moved in one direction or the other [*p(P3⇒P2)*]. We performed a parameter scan (Fig. S1) by generating simulated genomes with different values of *pINT* and *p(P3⇒P2)* and applying *DIP* to each genome (Fig. 3D). We found the expected *P3⇒P2 DIP* profile for the majority of replicated genomes generated with *p(P3⇒P2)*=1 (i.e. unidirectional *P3⇒P2* introgression) (Fig. 3D, red boxes). Further, we found the expected *P2⇒P3 DIP* profile for the majority of replicated genomes generated with *p(P3⇒P2)*=0 (i.e. unidirectional *P2⇒P3* IG) (Fig. 3D, gray boxes). Intermediate *p(P3⇒P2)* values all yielded the expected *DIP* profile for bidirectional introgression for all replicates (Fig. 3D, white boxes). These simulations constitute the basic implementation of *DIP* (hereafter, referred to as Single-*DIP* or *1×DIP*), which can detect the presence of bidirectional introgression (see Fig. 3B profile and Fig. 3D white boxes), but does not report directional asymmetry (i.e. whether either of the two directions predominates) at intermediate values of *p(P3⇒P2)*..

Double-DIP: Detecting asymmetry in cases of bidirectional introgression

Existing introgression polarization methods tend to assume unidirectionality of introgression, but it is also important to consider the possibility of asymmetric bidirectional introgression that falls short of being strictly unidirectional [discussed in (Martin et al. 2015)]. To more directly test for asymmetry in cases of bidirectional introgression, we developed an additional step in the *DIP* analysis, which we refer to as Double-*DIP* or *2×DIP*. The premise of *2×DIP* is that ΔK_{12} for loci introgressed *P3⇒P2* and ΔK_{13} for loci introgressed *P2⇒P3* have the same expected values, as they are both based on a shift in divergence time between T_β and T_α (Fig. 1). Therefore, under symmetric bidirectional (*P3⇔P2*) introgression, we expect genome-wide values of ΔK_{12} and ΔK_{13} to equal each other. Alternatively, if *P3⇒P2* introgression exceeds *P2⇒P3* introgression, we expect genome-wide $\Delta K_{12} > \Delta K_{13}$. *2×DIP* compares the magnitudes of ΔK_{12} and ΔK_{13} by formulating a simple summary statistic, $\Delta\Delta K$, which is defined as follows:

Eq. 4:

$$\Delta\Delta K = \Delta K_{12} - \Delta K_{13}$$

The expectation for the *ΔΔK* summary statistic is zero under symmetric bidirectional introgression, positive under introgression that is biased towards *P2*, and negative under introgression that is biased towards *P3* (Fig. 4).

We explored the performance of *2×DIP* by simulating genomes in the same manner as described above for *1×DIP*. For the genome simulated under unidirectional *P2⇒P3* introgression (*p(P3⇒P2)* = 0), we observed a significantly negative *ΔΔK* (Fig. 5A, $p < 0.0002$), consistent with our expectations. For the genome simulated under symmetric bidirectional introgression, *ΔΔK* did not significantly differ from zero (Fig. 5B, $p = 0.914$), also consistent with expectations. For the genome simulated under unidirectional *P3⇒P2* introgression (*p(P3⇒P2)* = 1), we observed significantly positive *ΔΔK* (Fig. 5C, $p < 0.0002$), again reflecting expectations. These results indicate that *2×DIP* correctly classified all three types of simulated introgression events. As above, we also performed a parameter scan to explore *2×DIP*. We found that genomes simulated with *p(P3⇒P2)* = 0.5 (i.e. symmetric bidirectional introgression) returned *ΔΔK* value that did not significantly differ from zero (Fig. 5D, white boxes). We also found significant *ΔΔK* < 0 for nearly all replicated genomes simulated with *p(P3⇒P2)* < 0.5 and significant *ΔΔK* > 0 for nearly all replicated genomes simulated with *p(P3⇒P2)* > 0.5 (Fig. 5D). The only exception to these patterns were found when ten percent or less of loci in the simulated genome (*pINT* ≤ 0.1) underwent nearly symmetrical introgression (*p(P3⇒P2)* = 0.45 and 0.55).

To test the influence of recombination on *DIP* performance, we also applied an alternative simulation approach in which full chromosomes were simulated under different rates of recombination (resulting in varying haplotype block sizes), while applying the same 5000-bp partition size used in our other analyses (see Methods). We found that *2×DIP* correctly inferred unidirectional introgression regardless of recombination rate (Fig. S2*; p(P3⇒P2)* = 0 and 1) and reliably detected slight (*p(P3⇒P2)* = 0.4 and 0.6) directional asymmetries when the size of haplotype blocks were the same or smaller than the size of the sliding window applied during *DIP* (Fig. S2B-C). However, when haplotype blocks were an order of magnitude larger than the window size, we observed increased noise in *DIP* at intermediate *p(P3⇒P2)* values (Fig. S2A), likely due to pseudoreplication caused by many trees reflecting the exact same genealogy (Fig. S2D), ultimately leading to increased sampling variance (see Discussion). Taken together, these

results indicate that *2×DIP* correctly inferred asymmetrical introgression, even in many cases in which there is only slight asymmetry, meaning it is a sensitive method for polarizing asymmetrical introgression that is robust across a variety of parameter values.

Robustness of DIP to population divergence time

The task of accurately classifying loci as introgressed *vs.* non-introgressed (i.e. INT loci *vs.* SP loci, respectively) based on gene tree topology is an integral part of *DIP;* however, this task is confounded when the topology of a gene tree doesn't accurately reflect the history of introgression (or lack thereof) that occurred at that locus. For example, phylogenetic methods rely on diagnostic synapomorphies to infer gene tree topologies; scarcity of synapomorphies or large amounts of homoplasy in an alignment can lead to phylogenetic error and, thus, inaccurate classification. Another important confounding factor is ILS, which can result in gene trees that reflect a history of deep coalescence at a locus, as opposed to the underlying history of speciation and/or introgression at that locus. This process can result in non-introgressed loci displaying the introgressed topology. Alternatively, because ILS and introgression are not mutually exclusive processes, ILS can also lead to introgressed loci displaying the species topology. Importantly, ILS is also expected to yield gene trees displaying an alternative third topology that is neither the species topology or the introgressed topology (Green et al. 2010) (see Triple-*DIP* below).

Both phylogenetic error and ILS are more pronounced during rapid divergence (i.e. short internal branches) (Fontaine et al. 2015). Moreover, it has been shown that, because *P3⇒P2* introgression trees have longer internal branch lengths than *P2⇒P3* introgression trees, the latter are more prone to both phylogenetic error and ILS (Zheng and Janke 2018), ultimately leading them to be more prone to misclassification in *DIP*. This feature introduces the potential for directional bias in *DIP* (see Discussion). Therefore, we explored divergence times, as an additional parameter that may influence performance. We focus our discussion on the process of ILS, but it should be noted that phylogenetic error also has the potential to occur in empirical datasets.

All previous simulations were implemented with constant and large divergence times (see Fig. 1). To explore the branch length parameter, we modified divergence times by multiplying all of the branch lengths by a scaling factor (SF) (see Methods), essentially modifying the height of the entire tree used for simulations. SFs greater than one yield taller trees, while SFs less than

one yield shorter trees. For each SF, we simulated five replicate genomes and calculated *ΔΔK* for each replicate. We first classified introgressed and non-introgressed loci based on the known history used to simulate the data and plotted the resulting *ΔΔK* values (omniscient *2×DIP*). We found that *2×DIP* correctly inferred asymmetry (or lack thereof) at all branch lengths and that the magnitude of *ΔΔK* was proportional to the SF (Fig. 6A, D and G). However, when working with real datasets it is rare to know if individual loci with introgression topologies are the result of bona fide introgression, as opposed to ILS or errors in phylogenetic inference. To explore the impact of the SF on the ability of *2×DIP* to distinguish between a signature of bona fide introgression versus the effects of ILS, we calculated *ΔΔK* using topology-based (non-omniscient) classification. With this approach, we observed an upward bias in *ΔΔK* at low SFs (Fig. 6B, E, and H). This bias favors inference of *P3⇒P2* introgression even when there is asymmetry in the opposite direction (Fig. 6E). As expected, this bias exists at the SFs for which misclassification of gene trees is most pronounced (Fig. S3), suggesting that it results from ILS (see Discussion).

We also explored the influence of the timing of introgression relative to speciation nodes. We held the timing of speciation constant while varying only the timing of the introgression event (i.e. relative introgression time). We found that *2×DIP* accurately polarizes asymmetric introgression in all cases under omniscience (Fig. S4A and D). Under non-omniscience, *2×DIP* is accurate when speciation and introgression are separated by a substantial period of time (i.e. relatively recent introgression times) (Fig. S4B). However, we observe a bias in favor of inference of *P3⇒P2* introgression (similar to the bias described above) when introgression occurs immediately following speciation (Fig. S4B) and this effect is compounded when total tree-height is small (i.e. SF=0.1) (Fig. S4E). Below we explore sources of bias and strategies for mitigating its effects.

Triple-DIP: Adjusting for gene tree classification bias

To address the directional bias in *2×DIP* caused by gene tree ILS at short branch lengths, we developed an additional layer that can be applied in *DIP* analysis, which we refer to as Triple-*DIP* or *3×DIP*, so named because it includes an additional *Δ* component (i.e. the "delta of the delta of the delta"). Briefly, in addition to calculating the standard *2×DIP* as above, we also

calculate an alternative *ΔΔK* (*ΔΔK*$_{ALT}$) that substitutes gene trees with the alternative topology, ((*P1*, *P3*), *P2*), for the introgressed loci used in the standard *ΔΔK:*

Eq. 5

$$\Delta\Delta K_{ALT} = \left(\overline{K}_{12}(ALT\ loci) - \overline{K}_{12}(SP\ loci)\right) - \left(\overline{K}_{23}(SP\ loci) - \overline{K}_{23}(ALT\ loci)\right)$$

Note that, K_{23} values are substituted in place of K_{13} values in calculating this version of *ΔΔK* because we are now focusing on a conflicting topology in which *P1* and *P3* are sister to each other. Because *P2* and *P3* are the two taxa subject to introgression, loci with this alternative topology should arise only from ILS and not introgression. Following the logic of standard *D* statistics (Green et al. 2010; Durand et al. 2011), we reasoned that ILS should be equally likely to produce each of the two topologies that conflict with the species tree. Therefore, this alternative *2×DIP* calculation may provide a measure of the amount of bias that is introduced by ILS. In applying *3×DIP,* we weight this value by the counts of loci with the expected (*P3*⇔*P2*) introgression topology (N_{INT} loci) and the alternative topology (N_{ALT} loci). The *ΔΔΔK* summary statistic is calculated as follows (see Methods for derivation):

Eq. 6

$$\Delta\Delta\Delta K = \frac{(\Delta\Delta K\ *\ N_{INT}) - (\Delta\Delta K_{ALT}\ *\ N_{ALT})}{N_{INT} - N_{ALT}}$$

It should be noted that calculation of a *3×DIP* correction is only possible when there is at least some ILS because it relies on the presence of ((*P1*, *P3*), *P2*) loci. As such, when we applied *3×DIP* to genomes simulated with different branch lengths, we were only able to consistently obtain measurements under short-branch conditions (SF < 1.0) where ILS is prevalent (Fig. 6C, F, and I) because these were the only conditions that returned some loci with the relevant topology. Under these short-branch conditions, we found that *3×DIP* reduced but did not eliminate the bias observed in *2×DIP*. While *ΔΔΔK* was still erroneously positive for the lowest branch length values (Fig. 6F and I), the magnitude of *ΔΔΔK* was less than that of *ΔΔK*.

We further explored bias in *2×DIP* and *3×DIP* by simulating short branch trees (with SF of 0.1, 0.2, and 0.3) across a range of *p(P3⇒P2)* values. We first applied omniscient *2×DIP* to give context to the bias introduced. As expected, omniscient *2×DIP* yielded negative *ΔΔK* values for all replicates in which *p(P3⇒P2)* < 0.5 (Fig. 7A). Consistent with the bias observed in Fig. 6, standard (non-omniscient) *2×DIP* yielded erroneously positive *ΔΔK* values, especially for the shortest branch length conditions (Fig. 7B). *3×DIP* reduced the bias, only yielding erroneously positive *ΔΔΔK* values for the highest *p(P3⇒P2)* values and the shortest branch length conditions (Fig. 7C). We also tested the performance of *DIP* in a situation in which ILS has occurred but not introgression (*pINT*=0; SF=0.1) (Fig. S5). Despite the lack of true introgression in these simulations, *1×DIP* produced a profile consistent with *P3⇒P2* introgression (Fig. S5B), although the relative positions of ΔK_{23}, ΔK_{12}, and ΔK_{13} distributions differed from the pattern in Fig. 3C. *2×DIP* also significantly indicated *P3⇒P2* introgression (Fig. S5C), but *3×DIP* produced a *ΔΔΔK* that was not significantly different from zero, again indicating that *3×DIP* is less prone to falsely indicating *P3⇒P2* introgression. However, when we explored bias in the context of relative introgression timing (as opposed to total tree height), we found some situations in which *3×DIP* showed increased directional bias compared to *2×DIP* (Fig. S4). *3×DIP* bias exceeded *2×DIP* bias in situations in which total tree height was large (high SFs) (Fig. S4G) but the opposite was true for low SFs (Fig. S4H). Together, these results indicate that *3×DIP* reduces bias in some (but not all) situations, meaning that information can be gained by applying both *2×DIP* and *3×DIP* when analyzing empirical data.

Analysis of hominin introgression

To understand the performance of *DIP* on empirical data, we applied *DIP* to existing genomic data for introgression that occurred between Neanderthal and a modern human European lineage (Green et al. 2010; Prüfer et al. 2014). Applying a five-taxon version of the *D*-statistic that made use of the phylogenetic position of multiple modern African populations, a previous study (Green et al. 2010) determined that unidirectional introgression occurred from Neanderthal to European lineages. We applied *DIP* to Chromosome 1 from a Neanderthal sample, a Denisovan sample, two modern human (San [African] and French [European]) samples, and the chimpanzee reference genome. The availability of a Denisovan sample allowed us to infer *DIP* in two

different ways using two different taxon sampling schemes (TSS1 and TSS2) (Fig. 8A and F). For both TSSs, there were three gene tree topologies present (Fig. 8B and G), indicating the possibility of misclassification due to phylogenetic error and ILS.

Using TSS1, *1×DIP* yielded a profile indicating the presence of at least some bidirectional introgression (Fig. 8C), a scenario which was not ruled out by (Green et al. 2010). However, it should be noted that, while ΔK_{12} and ΔK_{13} were both significantly positive, the ΔK_{13} was much closer to zero, which would indicate a substantial asymmetry towards Neanderthal⇒French introgression. *2×DIP* and *3×DIP* indicated significantly positive $\Delta\Delta K$ and $\Delta\Delta\Delta K$, respectively (Fig. 8D and E), consistent with asymmetric introgression in the Neanderthal⇒French direction. However, when we applied *DIP* to TSS2, we saw contradictory results. While, *1×DIP* again indicated the presence of bidirectional introgression, although without the near-zero ΔK_{13} (Fig. 8H), *2×DIP* and *3×DIP* yielded positive $\Delta\Delta K$ and $\Delta\Delta\Delta K$, respectively (Fig. 8I and J). *2×DIP* and *3×DIP* from TSS2 thus indicate French⇒Neanderthal introgression. While introgression from modern humans has been inferred in other Neanderthal samples (Kuhlwilm et al. 2016), it is at odds with findings from TSS1 and (Green et al. 2010).

To understand this discrepancy and put our empirical analyses in the context of our simulations, we plotted distributions of divergence estimates (K_{23}, K_{12}, and K_{13}) calculated from two simulated genomes and the TSSs used for the empirical analysis (Fig. S6). The empirical distributions display a wider spread than the simulated distributions, potentially introducing noise into the empirical analysis. Importantly, empirical data also show reduced levels of divergence, even compared to the dataset simulated with the shortest branch lengths (SF = 0.1). This suggests that the biasing factors explored above could be even more at play in the hominin analysis (see Discussion).

DISCUSSION

Intended applications of DIP

Our simulation analyses provide a proof-of-principle that divergence data can be used to polarize introgression in a four-taxon context, narrowing the methodological gap between our ability to identify introgression and our ability to determine the direction of gene transfer. It should be noted that *DIP* is not designed to replace existing methods and act as a frontline test of whether

introgression has occurred. Instead, we recommend cases of introgression first be confidently identified with existing tools (Huson et al. 2005; Than et al. 2008; Green et al. 2010; Durand et al. 2011; Martin et al. 2015; Pease and Hahn 2015; Stenz et al. 2015; Rosenzweig et al. 2016). In these cases, *DIP* can then be used to polarize the direction of introgression, a critical step toward interpreting biological implications. As we have shown above, *DIP* has the potential to distinguish unidirectional and bidirectional introgression and, in cases of bidirectionality, to test for asymmetry between the two directions.

While there are population genetic (Schrider et al. 2018) and five-taxon phylogenetic (Green et al. 2010; Pease and Hahn 2015) methods capable of polarizing introgression, *DIP* offers the ability to detect asymmetric introgression in both directions using a four-taxon context. This will be valuable because very little is known about the extent of reciprocal exchange that occurred during even well-studied introgression events (Green et al. 2010; Kuhlwilm et al. 2016), a deficit that likely stems from an absence of sensitive tools. Another group (Hibbins and Hahn 2019) has recently proposed an approach that overlaps with *DIP*. They introduce a statistic, D_2, which is conceptually similar to ΔK_{13} described here. As such, non-zero values of D_2 indicate the presence of *P2⇒P3* introgression (B⇒C by their nomenclature). *DIP* goes further than this approach because it also uses ΔK_{12} to test for introgression in the opposite direction and $\Delta\Delta K$ to determine the predominant direction of introgression. The primary focus of the recent work by Hibbins and Hahn (2019) is the development of another statistic, D_1, that assesses the timing of introgression relative to speciation events and can be used in assessing possible cases of homoploid hybrid speciation. This is an elegant application of the same type of divergence-based logic that underlies *DIP* to a biological question that cannot currently be addressed with our method. We suggest that further improvements in polarizing introgression can be made by combining the explicit coalescent-based modeling of Hibbins and Hahn (2019) with the more comprehensive summary provided by *1*×, *2*×, and *3*×*DIP*.

Bias in DIP

It should be noted that the simulation branch length parameters used in Fig. 3 and Fig. 5 resulted in gene trees with relatively deep divergences. These branch lengths were chosen because they emphasize differences in divergence and minimize potential biasing factors, thus providing the clearest view of the general properties of *DIP.* However, it has been shown that timing of

396 population divergence is an extremely influential parameter in introgression analyses (Durand et
397 al. 2011; Martin et al. 2015; Zheng and Janke 2018). This is true, in part, because the length of
398 internal branches is directly related to the extent of ILS that occurs (Maddison and Knowles
399 2006). Short branches lead to increased ILS (Degnan and Rosenberg 2013), which can mimic
400 introgression and introduce noise and bias into introgression analyses. Coalescent simulations,
401 such as those that we performed, capture this phenomenon (Hudson 2002; Degnan and
402 Rosenberg 2009), introducing discordant gene trees at a rate dependent on branch length
403 parameters.
404 Population divergence is additionally important for *DIP* for a more intuitive reason; the
405 magnitude of the *ΔK* measurements, which are the cornerstone of *DIP,* are directly proportional
406 to the length of internal branches, meaning that *DIP* gains power to differentiate between
407 alternative hypotheses as branches are lengthened. Finally, there is a disparity in the accuracy of
408 topology classification for loci introgressed *P3⇒P2 vs.* the opposite direction (Zheng and Janke
409 2018). This disparity stems from the fact that the internal branch on *P2⇒P3* introgression gene
410 trees are shorter than the same branch on *P3⇒P2* introgression gene trees, making for fewer
411 diagnostic synapomorphies by which to infer the introgression topology. This disparity is most
412 pronounced under conditions in which phylogenetically informative synapomorphies are scarce
413 (i.e. short branch lengths). Moreover, the specific disparity between genes introgressed in each
414 direction has an important consequence for simulation analyses, the short internal branch on
415 *P2⇒P3* introgression gene trees results in a higher rate of ILS for these loci compared to other
416 categories of loci, meaning that ILS obscures the introgression history of these loci at a higher
417 rate than loci introgressed in the opposite direction. This disparity is especially problematic for
418 *DIP* because it is likely to introduce a directional bias, favoring inference of *P3⇒P2*
419 introgression.
420 For the above reasons, we performed parameter scans to explore the influence of branch
421 lengths and timing of introgression. We found that *2×DIP* performs as expected when the
422 classification step is bypassed in omniscient mode (Fig 6A, D and G) but bias at short branch
423 lengths arises when introgressed and non-introgressed loci must be classified directly based on
424 the data (Fig. 6B, E, and H). When working with empirical datasets, omniscience about origins
425 and the effects of introgression *vs.* ILS on individual loci is not possible. As such, classification
426 error may be unavoidable, so we sought to develop a strategy to correct for bias that arises from

it, leading to the development of *3×DIP*. A benefit of *3×DIP* is that it is applicable under the conditions in which bias is most pronounced. Following the logic of the *D*-statistic (Green et al. 2010), *3×DIP* is based on the expectation that ILS is equally likely to produce the two topologies that conflict with the species tree: (*P1*(*P2*,*P3*)) and (*P2*(*P1*,*P3*)). Therefore, under the assumption that there has been no introgression between *P3* and *P1*, the number of ALT loci, which are defined by having the (*P2*(*P1*,*P3*)) topology, provides an estimate for the number of identified loci displaying the introgressed topology that were actually the result of ILS. Accordingly, *3×DIP* applies a correction for ILS that is proportional to the frequency of these ALT loci. We found that *3×DIP* reduces directional bias at short branch lengths (Fig. 6C, F, and I; Fig. 6) and does not provide false positive results in the complete absence of introgression (Fig. S5). These results indicate that *3×DIP* is a step toward overcoming directional bias; however, bias persisted for the shortest branch length simulations, meaning that there are biological scenarios in which *3×DIP* is not free from bias. Further, under situations in which introgression occurs immediately following speciation, we observed cases in which *2×DIP* exhibited less bias than *3×DIP* (Fig. S4G).

The basic premise of *3×DIP* is that the number of ALT loci serves as a proxy for the number of loci that have a true history of speciation but display an introgression topology due to ILS. This assumption appears valid in a scenario with ILS but not introgression, as indicated by the ability of *3×DIP* to eliminate bias under these simulated conditions (Fig. S5). However, *3×DIP* does not account for the fact that ILS occurs not only for loci with a speciation history, but also loci with an introgression history. In other words, some of the loci that exhibit the ALT topology will have a true history of introgression, making these loci an imperfect proxy for the number of loci with a speciation history affected by ILS. This can cause undesired behavior of *3×DIP* in situations in which most or all of the ALT topologies stem from loci with a history of *P2⇒P3* introgression. Therefore, we suggest that there is a benefit to applying all three variations of *DIP* to provide the most comprehensive view of introgression directionality.

Fully overcoming bias introduced into introgression analyses by classification error represents a future goal for the field. With current implementations of *DIP*, inferences of introgression in the *P3⇒P2* direction should be viewed with caution, especially in taxa with very recent divergence times or when introgression occurred very shortly after a speciation event. On

the other hand, it can be viewed as a conservative test for *P2⇒P3* introgression, so identification of introgression in that direction can be interpreted as a much more confident prediction. As suggested above, further progress in this area may come through more complex models that explicitly include ILS that occurs at introgressed loci (Hibbins and Hahn 2019), rather than solely at non-introgressed loci.

A related challenge to *DIP* analyses is associated with the question of how to partition the genome. Arbitrarily breaking chromosomes into loci of a fixed size may be problematic because the resulting 'loci' may either be composed of multiple haplotype blocks with different genealogies due to intralocus recombination or, conversely, an individual haplotype block may contain multiple partitioned 'loci', resulting in pseudoreplication as it will be sampled numerous times by *DIP*. Our simulations of introgression and recombination revealed that these issues do not introduce a directional bias but do dramatically increase the variance of *DIP* when the size of true haplotype blocks is much larger than the window size used by *DIP*. One potential strategy for mitigating this challenge would be to incorporate methods that explicitly infer recombination breakpoints (e.g. the four-game test (Hudson and Kaplan 1985)) into the window-definition phase of *DIP*.

There are also unexplored factors that should be considered when implementing *DIP* because our simulations were run under simplifying assumptions such as random mating, constant population size, and a single bout of instantaneous introgression solely between *P3* and *P2*. Violation of these assumptions in natural populations (Eriksson and Manica 2012; Prüfer et al. 2014; Kuhlwilm et al. 2016; Slon et al. 2018) may introduce additional sources of bias, Our simulation strategies also do not fully capture rate heterogeneity across the genome, branch-specific variation in effective population size/mutation rate, technical biases caused by read-mapping, and introgression from unsampled taxa (i.e. "ghost lineages"). These factors should be investigated in future studies with more complex simulation scenarios.

DIP performance on empirical data

We chose hominin introgression as a test case because it is one of the most famous and best-studied examples of introgression. An additional benefit is that the sampling in the group is dense; several modern human samples as well as samples from ancient Neanderthal and Denisovan tissues are available. A benefit of this dense taxon sampling is that previous studies

have been able to apply five-taxon statistics to polarize introgression, leading to the conclusion that "all or almost all of the gene flow detected was from Neandertals into modern humans" (Green et al. 2010). However, more recent analyses of additional archaic samples from different parts of the hominin geographical range also indicated introgression in the opposite direction (Kuhlwilm et al. 2016) as well as mating between Neanderthals and Denisovans (Slon et al. 2018).

An additional benefit of dense hominin taxon-sampling is that the phylogenetic placement of samples allows us to analyze the same introgression event with four-taxon statistics from two different angles. We devised a TSS in which Neanderthal and a modern human acted as *P3* and *P2*, respectively (TSS1, Fig. 8A) as well as one in which the roles were reversed (TSS2, Fig. 8F). Importantly, these TSSs allowed us to evaluate whether the directional bias described above was strong enough to outweigh the true signature from introgression. *DIP* returned contradictory results for TSS1 and TSS2. In both cases, *2×DIP* and *3×DIP* favored *P3⇒P2* introgression, despite the identity of *P3* and *P2* being reversed in the two cases. The fact that both analyses sided with the directional bias we documented above, suggests that bias may be outweighing the introgression signature. This is consistent with the observation that hominin divergence is both lower and more heterogenous than our simulated branch lengths (Fig. S6), suggesting that biasing factors are strong enough to bias even *3×DIP*. It is worth noting, however, that the magnitude of $\Delta\Delta K$ from TSS1 is higher than that from TSS2 and the variance of $\Delta\Delta\Delta K$ is much larger for TSS2 than for TSS1, meaning the signal favoring Neanderthal⇒French introgression (the expected direction) is stronger and less noisy than the signal in the opposite direction.

Our general takeaway from analysis of hominin data is that, like all introgression analysis tools, there are limits to the conditions under which *DIP* can be reliably applied. Although *3×DIP* may represent a step in the right direction, in the case of hominin introgression, the level of ILS swamps out the signal of introgression. We suggest that incorporating an alternative means of identifying introgressed loci, such as f_d (Durand et al. 2011; Martin et al. 2015), may yield more reliable results when ILS is prevalent, representing an area of future work. For the time being, *DIP* will be most reliable in cases of introgression that occurred at more ancient time scales (Forsythe et al. in revision; Dasmahapatra et al. 2012; Fontaine et al. 2015).

METHODS

Resource availability

URLs for downloading previously published data are provided in place in the following sections. Scripts for reproducing the analyses in this study are available at: https://github.com/EvanForsythe/DIP. Also included are *R* scripts for performing *DIP* on genomic data. All scripts are callable from the command line. Users have the choice of inputting either whole chromosome alignments, which will be divided into single-window (i.e. locus) alignments in preparation for *DIP*. Alternatively, *DIP* takes single-locus alignments, bypassing the window partitioning step. *DIP* outputs descriptive statistics and PDF figures similar to Fig. 8.

Simulations of sequence evolution

We generated whole genome alignments in which introgression has occurred in some (but not all) loci, and in which donor and recipient taxa for each introgressed locus are known. To accomplish this, we simulated sequence evolution of loci 5000 nucleotides in length in a four-taxon system (three in-group taxa, *P1, P2*, and *P3* and an outgroup, *O*) (Fig. 1). All simulations were performed with *ms* (Hudson 2002) and *seq-gen* (Rambaut and Grassly 1997) implemented in *R* v3.5.0 with *phyclust* v0.1-22 (Chen 2011) similar to (Martin et al. 2015). *Ms* was used to generate a coalescence tree, which was passed to *seq-gen* in order to generate a sequence alignment. A portion of the loci were simulated to have evolved along a path of simple speciation. In the absence of ILS, the gene trees for these loci should match the speciation history, ((*P1,P2*)*P3*)*O*) (Fig. 1A). These loci, denoted as species topology loci, were simulated with the following *R* commands:

```
ret.msSP<-ms(nsam = 4, nreps = 1, opts = "-T -t 50 -I 4 1 1 1 1 -ej 4 2 1
-ej 8 3 1 -ej 12 4 1 -r 5 5000")

seqsSP<-seqgen(opts = "-mHKY -l5000 -s 0.01", newick.tree = ret.msSP[3])
```

In the above *ms* call, the -T argument directs *ms* to output gene trees, one of which is used as input for *seq-gen*. The -t argument sets the *theta* value used by *ms*, which was held constant across all simulations. The arguments -I 4 1 1 1 1 indicate that four populations were simulated with one individual sampled from each, which was also held constant across all simulations.

Each -ej command represents a speciation event (in a forward-time context), the first number following the -ej flag being the timing of the event and the two following numbers being the two daughter populations arising from the speciation. The -r argument indicates the rate of recombination and the final number indicates the length of the segments being simulated by *ms.* However, for this simulation strategy we only input one tree into seq-gen, essentially simulating non-recombining loci (however, see below for our explicit treatment of recombination).

Loci with instantaneous unidirectional introgression occurring between *P2* and *P3* were also simulated. Introgression trees (transferred in either direction) will have the topology, (*P3,P2*)*P1*)*O*), and thus differ from the species tree. The direction of introgression for an individual locus was indicated by 'donor taxon' and 'recipient taxon' as in the following *R* command:

```
ret.msIG <- ms(nsam = 4, nreps = 1, opts= "-T -t 50 -I 4 1 1 1 1 -ej 4 2
1 -ej 8 3 1 -ej 12 4 1 -es 2 <recipient taxon> 0.4 -ej 2 5 <donor taxon>
-r 5 5000")

seqsIG<-seqgen(opts = "-mHKY -l5000 -s 0.01", newick.tree = ret.msIG[3])
```

We replicated the above commands for species and introgressed topology loci to create datasets representing simulated whole-genome alignments composed of a total of 5000 loci (Fig. S1). The argument in the above command that specify introgression are the -es argument and the final -ej command. We define the proportion of all loci in the genome resulting from simulated introgression in either direction as *pINT* and the proportion of introgressed genes that were transferred in the *P3⇒P2* direction as *p(P3⇒P2).* Because a single locus can only be transferred in one direction or the other, the proportion of loci transferred in the *P2⇒P3* direction, *p(P2⇒P3),* is 1 – *p(P3⇒P2).* Whole genome alignments with known values of *p(IG)* and *p(P3⇒P2)* were used to test the performance of *DIP*. We performed parameter scans by simulating genome alignments with varying combinations of *p(IG)* and *p(P3⇒P2)* (See Fig. S1).

Recognizing that the above simulation strategy does not realistically model recombination, we also employed an alternative simulation strategy in which we simulate whole chromosomes (rather than individual loci) while allowing for varying levels of recombination.

Introgression in the presence of recombination was simulated with the following *ms* command in *R*.

```
ms(nsam = 4, nreps = 1, opts = T -t 50 -I 4 1 1 1 1 -ej 4 2 1 -ej 8 3 1 -
ej 12 4 1 -es 1 <recipient taxon> <pINT> -ej 1 5 <donor taxon> -r
<recombination rate> 12500000)
```

The output files from the above *ms* command (run twice in cases of bidirectional introgression—once for each direction of introgression) were combined into a single file, which was input to *seq-gen* in order to generate a whole chromosome alignment. *Seq-gen* was called from the command line with the following command:

```
seq-gen -mHKY -l 25000000 -s 0.01 -p <number of haplotype blocks from ms>
< <ms_output_file> > <seqgen output file name> 2> <file name to store
haplotype block positions>
```

Whole chromosome alignments were replicated five times for each parameter value and *DIP* analyses were performed with the 5000-bp partitioning approach applied elsewhere in this manuscript.

The default branch length parameters used for Fig. 3 and Fig. 5 are T_{INT}=1, T_{α}=4, T_{β}=8, and T_{γ}=12 measured in coalescent units of 4N generations (see Fig. 1). To explore the effects of divergence times, we multiplied all branch length parameters by a range of different scaling factor (SF) values. For example, SF=0.1 results in the following node depths: T_{INT}=0.1, T_{α}=0.4, T_{β}=0.8, and T_{γ}=1.2.

As an additional means of exploring the effects of speciation and introgression timing, we also varied the timing of introgression in proportion to the most recent speciation even (relative introgression time). The timing of introgression was set relative to the T_{α} speciation time. For example, under default SF described in the previous paragraph with T_{α}=4, a relative introgression time of 0.8 translates to T_{INT}=3.2. For parameter scans involving branch lengths, we generated point estimates of $\Delta\Delta K$ and $\Delta\Delta\Delta K$ from five replicate genomes for each condition.

Classification of introgressed and non-introgressed loci

The first step in all versions of *DIP* is sorting loci to distinguish the loci that were introgressed from those that follow the species branching order (i.e. classification). Using simulated data affords us omniscience at this step (i.e. we know whether each locus was originally simulated as introgressed or not). However, unless specifically stated, we did not make use of the known history of simulated loci. Instead, *DIP* infers the introgression status of loci based on the topology of a neighbor joining gene tree inferred for each locus using *Ape* v5.2 (Paradis et al. 2004). Loci displaying the ((*P1,P2*)*P3*)*O*) topology are marked as non-introgressed loci. Loci displaying the ((*P2,P3*)*P1*)*O*) topology (introgressed topology) are designated as introgressed loci. Any loci displaying the alternative topology, ((*P1,P3*)*P2*)*O*), which are not produced by speciation or introgression, are omitted from *1×DIP* and *2×DIP* but used by *3×DIP* to calculate a correction factor (see below).

Inferring introgression directionality with 1×DIP

We calculated the pairwise divergences, K_{23}, K_{12}, and K_{13} (as indicated in Fig. 1A) for each locus using the *dist.dna* command from the *Ape* package with default settings. Pairwise divergences, K_{23}, K_{12}, and K_{13} are named for the taxa involved in the distance calculation. For example, K_{23} measures the divergence of *P2* and *P3* (see Fig. 1). ΔK_{23}, ΔK_{12}, and ΔK_{13} were calculated based on difference in mean *K* values between SP and introgression loci as shown in *Eqs. 1-3*. To test for significance, bootstrapped distributions were obtained by resampling (with replacement) loci from the genome to achieve genome alignments equal in number of loci to the original genome alignment. One thousand such replicates were performed, recalculating ΔK_{23}, ΔK_{12}, and ΔK_{13} for each replicate. *P*-values for the significance of ΔK values were calculated as the proportion of replicates for which $\Delta K \leq 0$. For the parameter scan of *1×DIP* (Fig. 3D), inference of a significant directional profile required that all three measures, ΔK_{23}, ΔK_{12}, and ΔK_{13}, adhere to their expected profile with a significant ($p<0.05$) *p*-value for each (with the exception of cases in which the expectation is $\Delta K=0$).

Inferring introgression directionality with 2×DIP and 3×DIP

$\Delta\Delta K$ was calculated from ΔK_{12}, and ΔK_{13} as described in *Eq. 4*. The bootstrap resampling scheme described in the previous paragraph was used to assess the significance of *2×DIP*. $\Delta\Delta K$ was calculated for each replicate and *p*-values were obtained from the proportion of replicates for

which $\Delta\Delta K$ overlapped zero (multiplied by two for a two-sided test). Like *2×DIP, 3×DIP* makes use of $\Delta\Delta K$ to indicate the directionality of introgression. However, *3×DIP* also introduces $\Delta\Delta K_{ALT}$, which is calculated according to *Eq. 5*. $\Delta\Delta\Delta K$ is obtained from the difference between $\Delta\Delta K$ and $\Delta\Delta K_{ALT}$ (see *Eq. 6*).

The rationale for the *3xDIP* correction is that the observed value of $\Delta\Delta K$ may be viewed as a weighted average of 1) a corrected value ($\Delta\Delta\Delta K$) that is based only on the loci that truly experienced a history of introgression and 2) a spurious signal ($\Delta\Delta K_{ILS}$) arising from the unknown number of loci that exhibit an introgression topology that is actually the result of ILS (N_{ILS}).

Eq. 7

$$\Delta\Delta K = \left(\frac{N_{INT} - N_{ILS}}{N_{INT}}\right)\Delta\Delta\Delta K + \left(\frac{N_{ILS}}{N_{INT}}\right)\Delta\Delta K_{ILS}$$

Based on the expected symmetry of ILS, we can use $\Delta\Delta K_{ALT}$ and N_{ALT} as estimates of $\Delta\Delta K_{ILS}$ and N_{ILS}, respectively.

Eq. 8

$$\Delta\Delta K = \left(\frac{N_{INT} - N_{ALT}}{N_{INT}}\right)\Delta\Delta\Delta K + \left(\frac{N_{ALT}}{N_{INT}}\right)\Delta\Delta K_{ALT}$$

Solving *Eq. 8* for $\Delta\Delta\Delta K$ yields *Eq. 6* (see Results). This approach is based on substantial simplifying assumptions. For example, it does not account for the misidentification of loci that have a true history of introgression but exhibit the species or ALT topology because of ILS (see Discussion). As for $\Delta\Delta K$ above, significance of $\Delta\Delta\Delta K$ is obtained from resampled whole-genome alignments.

Hominin data analysis

To generate whole-chromosome alignments from the hominin dataset for *DIP*, Chromosome I sequencing data for two Neanderthal, one Denisovan, and two modern human samples from (Prüfer et al. 2014) were downloaded from http://cdna.eva.mpg.de/neandertal/ as VCF files. The human reference genome (hg19) (International Human Genome Sequencing Consortium 2001), which was originally used for read mapping during the creation of VCF files, was obtained from http://hgdownload.cse.ucsc.edu/goldenPath/hg19/.

Structural variation (indel) information was trimmed from VCF files, using *VCFtools v. 0.1.13* (Danecek et al. 2011) and *Tabix* (Li et al. 2009) with the following commands:

```
vcftools --gzvcf Chrom1_with_indels.vcf.gz --remove-indels --recode --
recode-INFO-all --out Chrom1_SNPs_only.vcf

bgzip Chrom1_SNPs_only.vcf

tabix -p vcf Chrom1_SNPs_only.vcf.gz
```

Whole-chromosome consensus sequence was extracted from VCF files using *BCFtools v1.6* (Li et al. 2009) with the command below. For heterozygous sites, by default *bcftools consensus* applies the alternative variant (i.e. the variant that does not match the reference genome) to the consensus sequence for the given sample (see https://samtools.github.io/bcftools/bcftools.html). It should be noted that heterozygosity information may be lost at this step, which was necessary to match the format of the phylogenetic data generated in our simulations.

```
cat hg19_chrom1.fa | bcftools consensus Chrom1_SNPs_only.vcf.gz >
Chrom_1_consensus.fa
```

We used the reference chimpanzee genome (PanTro5) (The Chimpanzee Sequencing Consortium 2005) as an outgroup. We downloaded a MAF alignment of chromosome one from PanTro5 and hg19 from: http://hgdownload.cse.ucsc.edu/goldenpath/hg19/vsPanTro5/axtNet/. We converted this file to FASTA format using Galaxy tools (Afgan et al. 2018) available at https://usegalaxy.org/. Finally, the consensus sequence from each hominin sample and chimpanzee was concatenated into a whole-chromosome multiple sequence alignment in FASTA

format. This five-taxon alignment was pruned to contain four taxa according to each TSS (see Fig. 8) and then divided into single-locus alignments 5000 bp in length, which were used as input to *DIP*.

ACKNOWLEDGMENTS

This work was funded by NSF grant IOS-1444490 to M.A.B as well as NSF grant MCB-1733227 to D.B.S. We thank M.W. Hahn, M.J. Sanderson, R.A. Mosher, A.D.L Nelson, K. Dew-Budd, K. Palos, A.E. Baniaga, S.M. Lambert, G. Achaz, and one anonymous reviewer for helpful comments and discussion.

FIGURE LEGENDS

Fig. 1. Expected divergence under simulated introgression
The species *P1, P2, P3*, and *O* were used for simulation analyses. (**A**) The species branching order. Introgression between species *P2* and *P3* is indicated with a double-sided dotted arrow. Default values used during all simulations, unless specified otherwise, are: T_{IG}=1, T_{α}=4, T_{β}=8, and T_{γ}=12 in coalescent units (4*N* generations) (Hudson 2002). (**B**) A gene tree depicting a gene that was introgressed *P3⇒P2*. (**C**) A gene tree depicting a gene that was introgressed *P2⇒P3*. ΔK values are calculated based on changes in mean divergence between pairs of taxa in the set of trees with the speciation topology *vs.* the set of introgression trees (see *Eq. 1-3*). Note that the expected profiles of ΔK values for *P3⇒P2* introgression differs from that of *P2⇒P3* introgression, forming the basis for the *DIP* test (see Main Text and Fig. 2).

Fig. 2. Workflow of the *DIP* test.
Point estimates of ΔK_{23}, ΔK_{12}, ΔK_{13} are calculated from whole genomes, which are then resampled to yield distributions of ΔK_{23}, ΔK_{12}, ΔK_{13}. Unidirectional *P3⇒P2* introgression is indicated by the profile, $\Delta K_{23} > 0$, $\Delta K_{12} > 0$, and $\Delta K_{13} = 0$. Unidirectional *P2⇒P3* introgression is indicated by $\Delta K_{23} > 0$, $\Delta K_{12} = 0$, and $\Delta K_{13} > 0$. Bidirectional introgression is indicated by $\Delta K_{23} > 0$, $\Delta K_{12} > 0$, and $\Delta K_{13} > 0$. All other profiles are considered inconclusive regarding the occurrence and directionality of introgression. *P*-values for testing whether each ΔK value significantly differs from 0 are obtained from the proportion of replicates for which $\Delta K \leq 0$. Colors reflect the black, red, and gray genealogical histories from Fig. 1. In this illustration, all introgression loci are in the *P3⇒P2* (red) direction. But we use the red/gray dashed lines for showing the distribution of introgression loci because, in general, the set of introgression loci can contain *P3⇒P2* loci, *P2⇒P3* loci, or both.

Fig. 3. *DIP* analysis of simulated introgression.
Genomes were simulated according to steps 1-3 in Fig. S1, under unidirectional *P2⇒P3* introgression (**A**), symmetrical bidirectional *P3⇔P2* introgression (**B**), and unidirectional *P3⇒P2* introgression (**C**). Simulation parameters are as follows: (**A**), n = 5000, *pINT* = 0.5, *p(P3⇒P2)* = 0; (**B**), n = 5000, *pINT* = 0.5, *p(P3⇒P2)* = 0.5; (**C**), n = 5000, *pINT* = 0.5,

p(P3⇒P2) = 1. *DIP* was applied to each genome to yield profiles of ΔK_{23}, ΔK_{12}, ΔK_{13}. ** indicates significant departure from 0 ($p < 0.01$). (**D**) A plot scanning simulation parameters, proportion of the genome that was introgressed (*pINT*) (y-axes) and proportion of introgressed loci transferred in each direction (*p(P3⇒P2)*) (x-axis). Each square in the plot indicates the *DIP* results obtained from five replicated simulated genome alignments. Red boxes indicate the profile consistent with *P3⇒P2* introgression (see panel C). Gray boxes indicate the profile consistent with *P2⇒P3* introgression (see panel A). The shading of the boxes corresponds the percentage of replicates that indicate a given profile, as specified by the key to the right of the plot. Unshaded boxes indicate zero replicates yielded a significant unidirectional profile (i.e. all replicates yield the bidirectional introgression profile (see panel B)).

Fig. 4. Workflow of the *2×DIP* test.
(**Top**) A point estimate of $\Delta\Delta K$ is calculated from a whole genome alignment from ΔK_{12} and ΔK_{13} values. (**Bottom**) A sampling distribution of $\Delta\Delta K$ is calculated from resampled gene alignments (bootstrapping) obtained from the original genome. If the majority of $\Delta\Delta K$ replicates are > 0, it is an indication of asymmetric *P3⇒P2* introgression. In this case, the proportion of $\Delta\Delta K$ replicates < 0 determines the *p*-value (doubled for a two-sided test) for asymmetric *P3⇒P2* introgression. Asymmetric *P2⇒P3* introgression is indicated by the opposite pattern.

Fig. 5. *2×DIP* analysis of simulated introgression.
Genomes were simulated according to steps 1-3 in Fig. S1. Genomes were simulated under unidirectional *P2⇒P3* introgression (**A**), symmetrical bidirectional *P3⇔P2* introgression (**B**), and unidirectional *P3⇒P2* introgression (**C**). Simulation parameters are as follows: (**A**), *n* = 5000, *pINT* = 0.5, *p(P3⇒P2)* = 0; (**B**), *n* = 5000, *pINT* = 0.5, *p(P3⇒P2)* = 0.5; (**C**), *n* = 5000, *pINT* = 0.5, *p(P3⇒P2)* = 1. *2×DIP* was applied to each genome to yield a sampling distribution of $\Delta\Delta K$. ** indicates significant departure from 0 ($p < 0.01$). (**D**) A plot scanning *pINT* and *p(P3⇒P2)* as in Fig. 3D. Red boxes indicate significant ($p<0.05$) *P3⇒P2 2×DIP* signature (see panel C). Gray boxes indicate significant ($p<0.05$) *P2⇒P3 2×DIP* signature (see panel A). Five replicate genomes were simulated for each parameter value. The shading of the boxes corresponds the percentage of replicates for which *2×DIP* significantly indicated a directional signature, as specified by the key to the right of the plot. Unshaded boxes indicate zero replicates yielded a significant directional signature (i.e. all five replicates failed to reject the null hypothesis of symmetrical introgression (see panel B)).

Fig. 6. Exploration of branch length parameters used during genome simulation.
The default branch lengths used during all previous simulations (T_{IG}=1, T_{α}=4, T_{β}=8, and T_{γ}=12) were multiplied by branch-length scaling factors. For all plots, five replicate genomes were simulated for each scaling factor value. *pINT* = 0.5 was used for all simulations. *DIP* was performed on each replicate; individual points on plots represent point estimates of $\Delta\Delta K$ and $\Delta\Delta\Delta K$ (jittered for clarity). Genomes were simulated with asymmetric introgression favoring *P3⇒P2* (**A-C**), symmetric bidirectional introgression (**D-F**), and asymmetric introgression favoring *P2⇒P3* (**G-I**). Omniscient *2×DIP* (**A, D, and G**), standard *2×DIP* (**B, E, and H**), and *3×DIP* (**C, F, and I**) were performed. $\Delta\Delta\Delta K$ data points are absent at higher scaling factors because this adjusted version of $\Delta\Delta K$ can only be calculated when there are at least some loci with the unexpected topology (ALT loci) as a result of topology misclassification or ILS.

Fig. 7. Characterization of *DIP* bias under short branch conditions.
Genomes were simulated with different values of *p(P3⇒P2)* (x axis) and different branch length scaling factors (SF) (point colors). See Fig. 6 for description of SF. Purple, SF = 0.1; Orange, SF = 0.2; Green, SF = 0.3. As in Fig. 6, Omniscient *2×DIP* (**A**), standard *2×DIP* (**B**), and *3×DIP* (**C**) were performed. Five replicate genomes were analyzed for each condition. *pINT* = 0.5 was used for all simulations.

Fig. 8. *DIP* analysis of hominin introgression. *DIP* was performed on whole-chromosome alignments of chromosome 1 using two different taxon sampling schemes (TSS). (**A**) Depiction of the samples used in TSS1. (**B**) Neighbor-joining gene-tree topologies from individual loci. (San.,French),Nean.), green; (French, Nean.),San), orange; (San, Nean.),French), purple. (**C-E**) Results from *1×DIP* (**C**), *2×DIP* (**D**), and *3×DIP* (**E**) applied to TSS1 alignment. (**F**) Depiction of the sampled used in TSS2. (**G**) Neighbor-joining gene-tree topologies from individual loci. (Deni.,Nean.),French), green; (Nean.,French),Deni.), orange; (Deni.,French),Nean.), purple. (**H-J**) Results from *1×DIP* (**H**), *2×DIP* (**I**), and *3×DIP* (**J**) applied to TSS2 alignment. ** indicates significant departure from 0 ($p < 0.01$).

FIGURES

Fig. 1

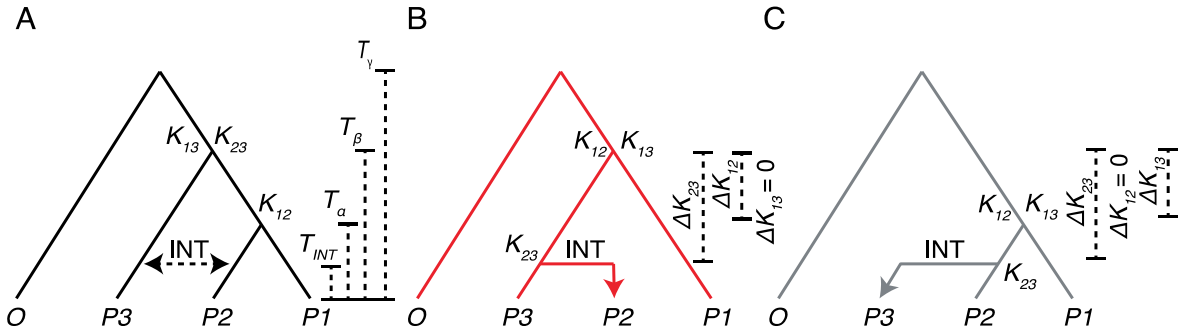


Fig. 2

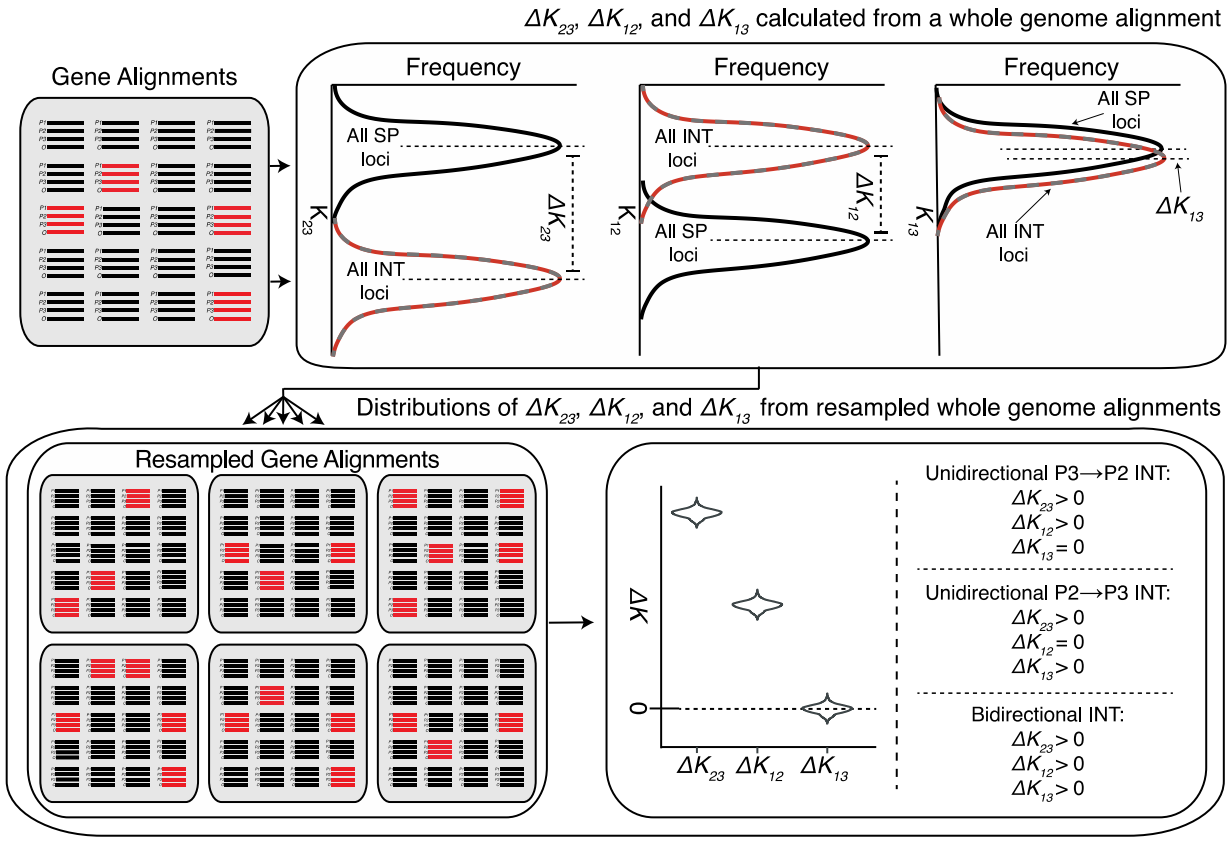

ΔK23, ΔK12, and ΔK13 calculated from a whole genome alignment
Gene Alignments
Frequency
All SP loci
All INT loci
K23
ΔK23
Frequency
All INT loci
All SP loci
K12
ΔK12
Frequency
All SP loci
All INT loci
K13
ΔK13
Distributions of ΔK23, ΔK12, and ΔK13 from resampled whole genome alignments
Resampled Gene Alignments
ΔK
0
ΔK23 ΔK12 ΔK13
Unidirectional P3⟶P2 INT:
ΔK23 > 0
ΔK12 > 0
ΔK13 = 0
Unidirectional P2⟶P3 INT:
ΔK23 > 0
ΔK12 = 0
ΔK13 > 0
Bidirectional INT:
ΔK23 > 0
ΔK12 > 0
ΔK13 > 0

Fig. 3

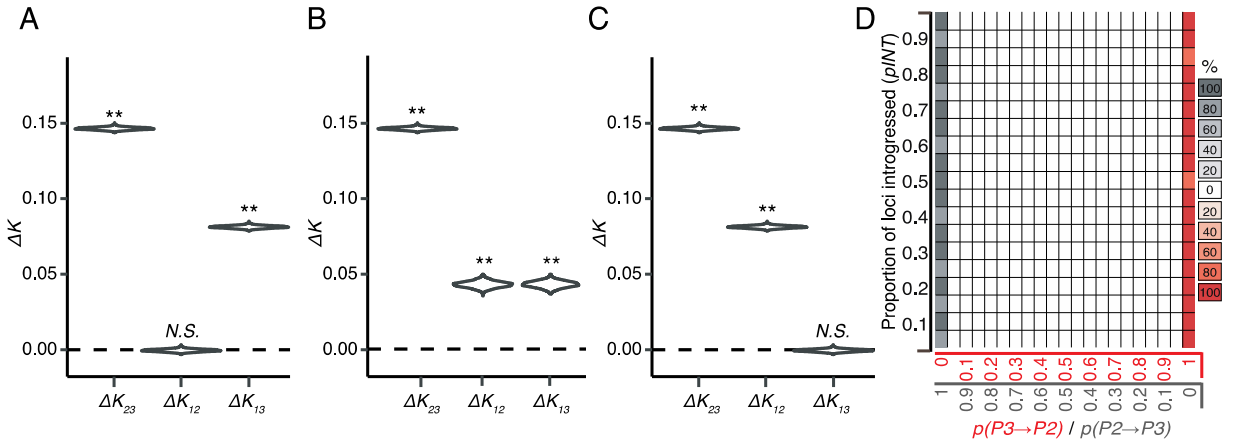

A
B
C
D
ΔK
0.15
0.10
0.05
0.00
**
N.S.
ΔK23
ΔK12
ΔK13
Proportion of loci introgressed (pINT)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
%
100
80
60
40
20
0
20
40
60
80
100
p(P3→P2) / p(P2→P3)

Fig. 4

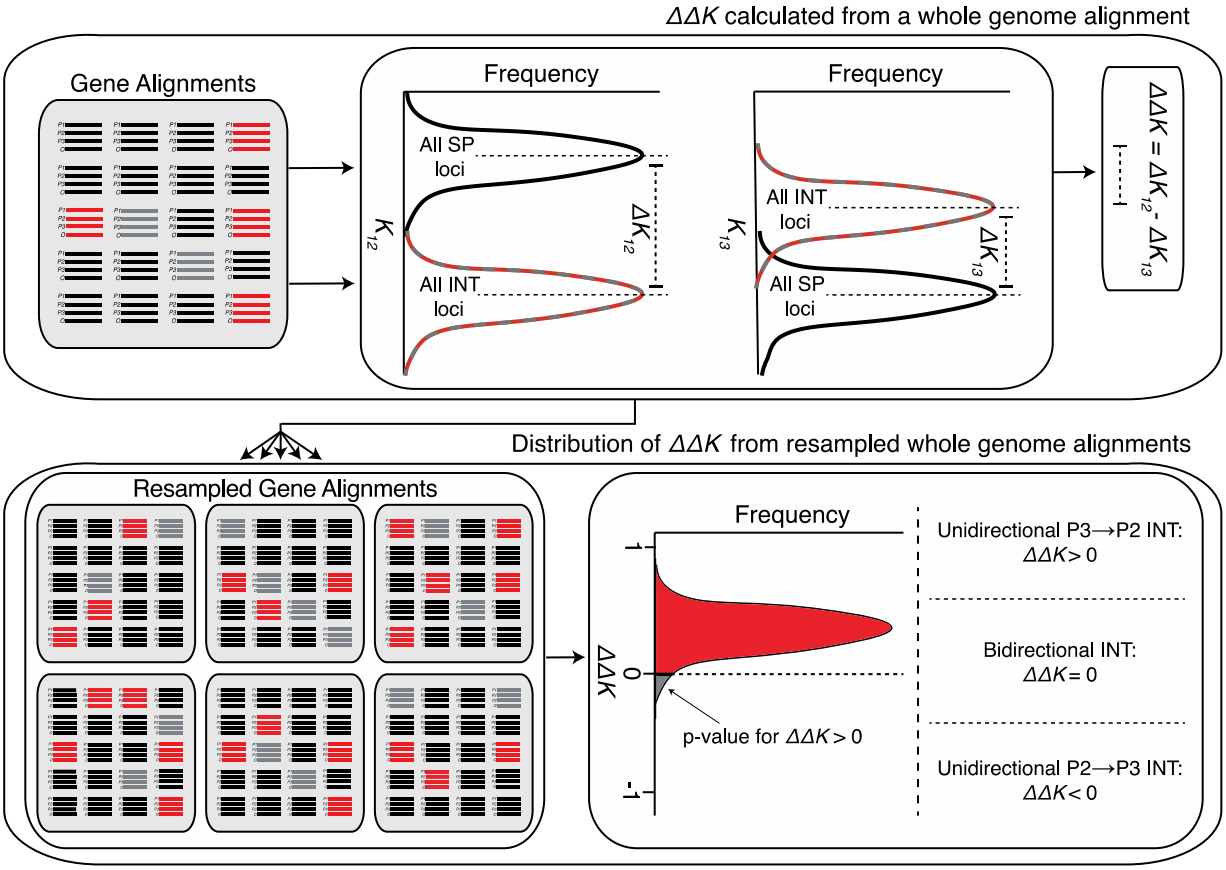

ΔΔK calculated from a whole genome alignment
Gene Alignments
Frequency
All SP loci
All INT loci
K12
ΔK12
Frequency
All INT loci
All SP loci
K13
ΔK13
ΔΔK = ΔK12 - ΔK13
Distribution of ΔΔK from resampled whole genome alignments
Resampled Gene Alignments
Frequency
1
0
-1
ΔΔK
p-value for ΔΔK > 0
Unidirectional P3→P2 INT:
ΔΔK > 0
Bidirectional INT:
ΔΔK = 0
Unidirectional P2→P3 INT:
ΔΔK < 0

Fig. 5

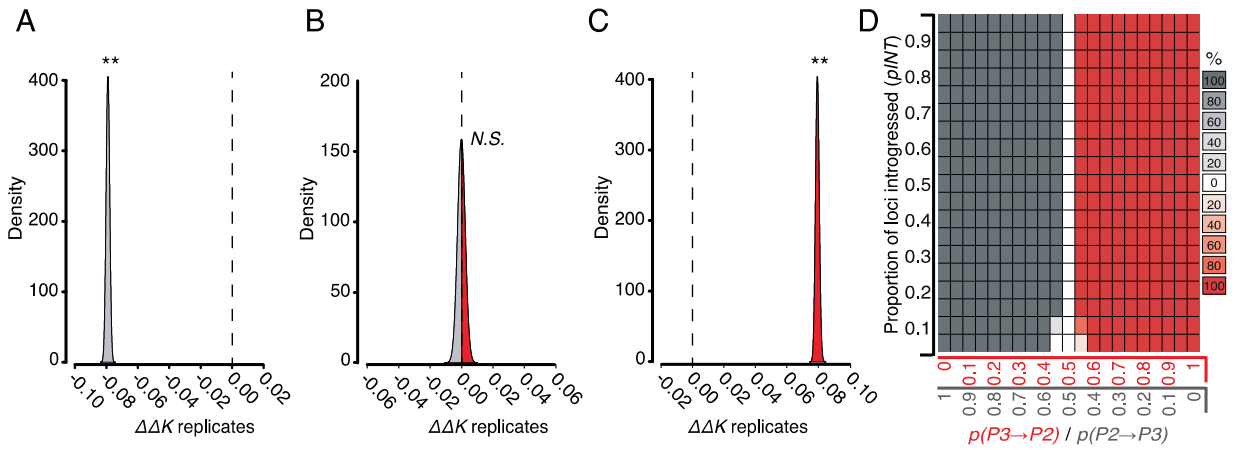

A
B
C
D
**
N.S.
**
Density
ΔΔK replicates
Proportion of loci introgressed (pINT)
%
p(P3→P2) / p(P2→P3)

Fig. 6

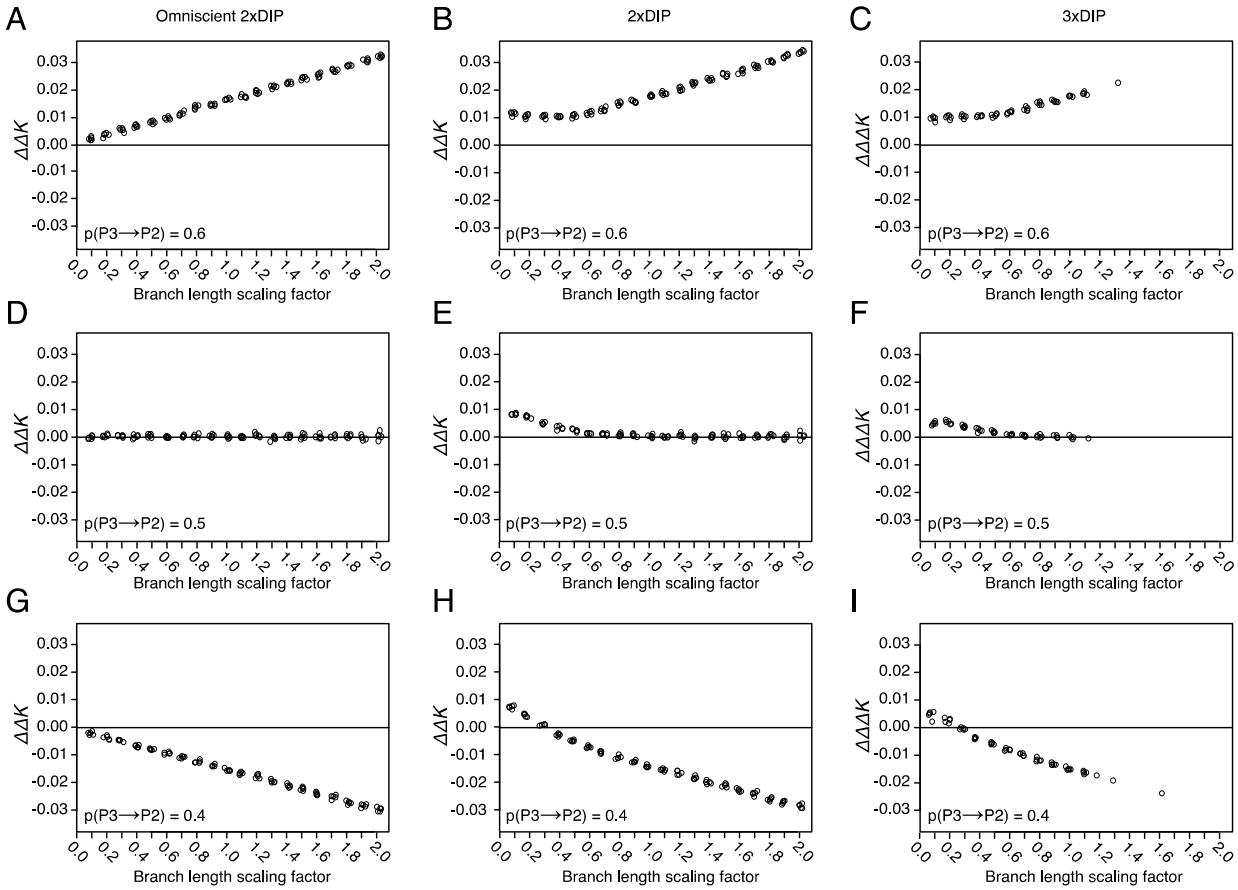

A
Omniscient 2xDIP
B
2xDIP
C
3xDIP
D
E
F
G
H
I
ΔΔK
ΔΔΔK
p(P3⟶P2) = 0.6
p(P3⟶P2) = 0.5
p(P3⟶P2) = 0.4
Branch length scaling factor

Fig. 7

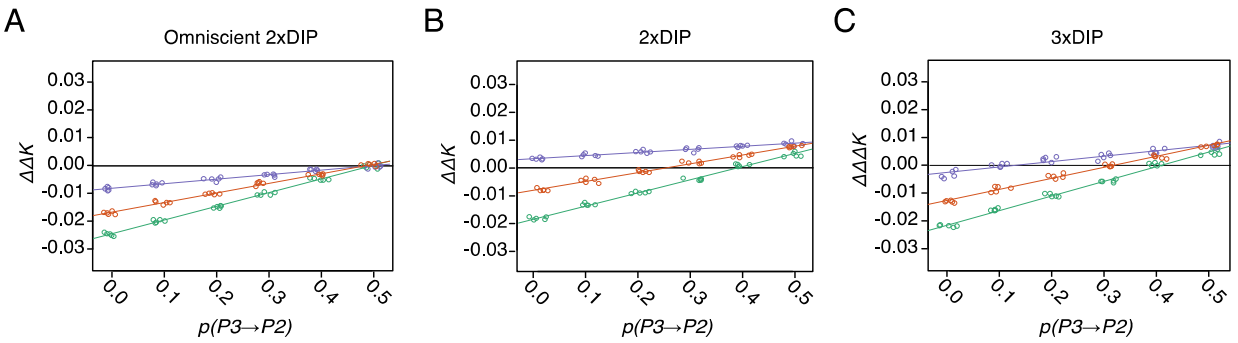
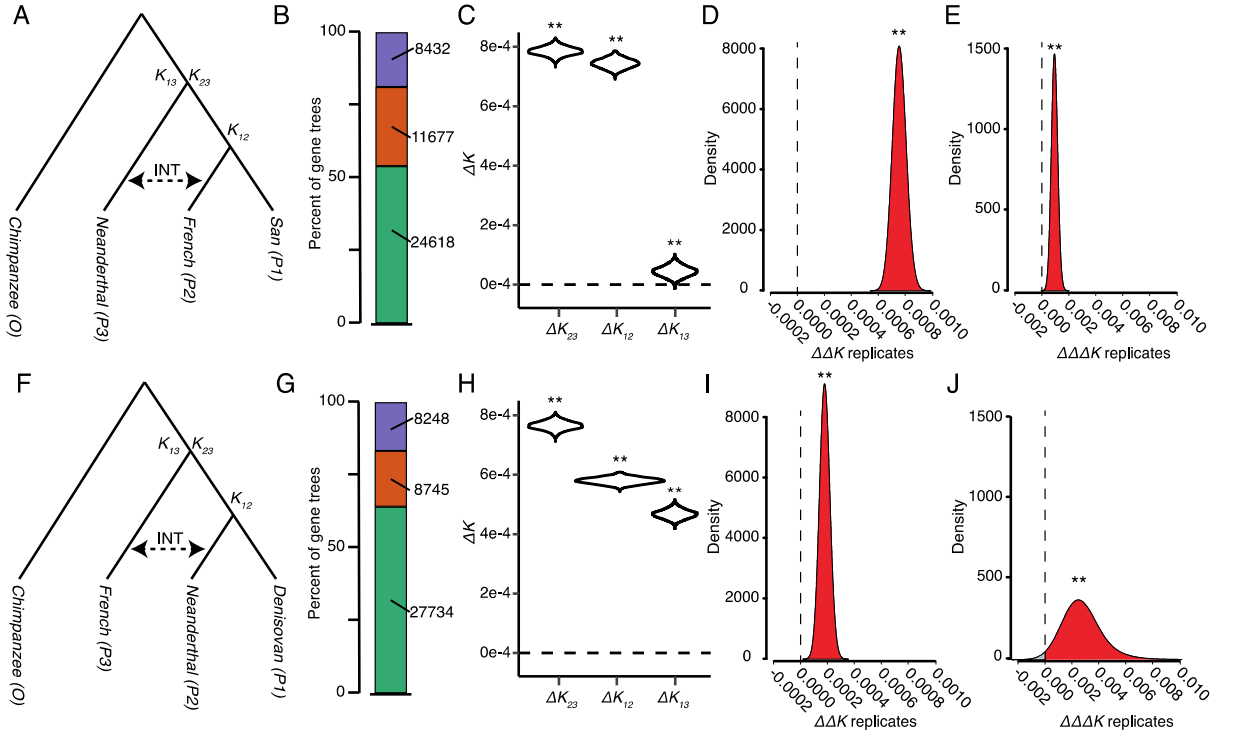

A
Omniscient 2xDIP
B
2xDIP
C
3xDIP
ΔΔK
0.03
0.02
0.01
0.00
-0.01
-0.02
-0.03
0.0
0.1
0.2
0.3
0.4
0.5
p(P3→P2)

Fig. 8


A
B
C
D
E
F
G
H
I
J
K_{13}
K_{23}
K_{12}
INT
Chimpanzee (O)
Neanderthal (P3)
French (P2)
San (P1)
French (P3)
Neanderthal (P2)
Denisovan (P1)
Percent of gene trees
8432
11677
24618
8248
8745
27734
ΔK
ΔK_{23}
ΔK_{12}
ΔK_{13}
Density
ΔΔK replicates
ΔΔΔK replicates
**

CITATIONS

Afgan E, Baker D, Beek M Van Den, Bouvier D, Chilton J, Clements D, Coraor N, Guerler A, Hillman-jackson J, Hiltemann S, et al. 2018. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. Nucleic Acids Res. 46:537–544.

Barton NH, Hewitt GM. 1985. Analysis of hybrid zones. Annu. Rev. Ecol. Syst. 16:113–148.

Chen W-C. 2011. Overlapping Codon model, Phylogenetic Cluetering, and Alternative Partial Expectation Conditional Maximization Algorithm. Ph.D. Diss., Iowa Stat Univ.

Currat M, Ruedi M, Petit RJ, Excoffier L. 2008. The hidden side of invasions: Massive introgression by local genes. Evolution (N. Y). 62:1908–1920.

Danecek P, Auton A, Abecasis G, Albers CA, Banks E, DePristo MA, Handsaker RE, Lunter G, Marth GT, Sherry ST, et al. 2011. The variant call format and VCFtools. Bioinformatics 27:2156–2158.

Dannemann M, Andrés AM, Kelso J. 2016. Introgression of Neandertal- and Denisovan-like Haplotypes Contributes to Adaptive Variation in Human Toll-like Receptors. Am. J. Hum. Genet. 98:22–33.

Dasmahapatra KK, Walters JR, Briscoe AD, Davey JW, Whibley A, Nadeau NJ, Zimin A V., Hughes DST, Ferguson LC, Martin SH, et al. 2012. Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. Nature 487:94–98.

Degnan JH, Rosenberg N a. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends Ecol. Evol. 24:332–340.

Degnan JH, Rosenberg NA. 2013. Discordance of species trees with their most likely gene trees: A unifying principle. Mol. Biol. Evol. 30:2709–2713.

Durand EY, Patterson N, Reich D, Slatkin M. 2011. Testing for Ancient Admixture between Closely Related Populations. Mol. Biol. Evol. 28:2239–2252.

Eaton DAR, Hipp AL, González-Rodríguez A, Cavender-Bares J. 2015. Historical introgression among the American live oaks and the comparative nature of tests for introgression. Evolution (N. Y). 69:2587–2601.

Eaton DAR, Ree RH. 2013. Inferring Phylogeny and Introgression using RADseq Data: An Example from Flowering Plants (Pedicularis: Orobanchaceae). Syst. Biol. 62:689–706.

Eriksson A, Manica A. 2012. Effect of ancient population structure on the degree of polymorphism shared between modern human populations and ancient hominins. Proc. Natl. Acad. Sci. 109:13956–13960.

Feder JL, Xie X, Rull J, Velez S, Forbes A, Leung B, Dambroski H, Filchak KE, Aluja M. 2005. Mayr, Dobzhansky, and Bush and the complexities of sympatric speciation in Rhagoletis. Proc. Natl. Acad. Sci. U. S. A. 102:6573–6580.

Figueiró H V, Li G, Trindade FJ, Assis J, Pais F, Fernandes G, Santos SHD, Hughes GM, Komissarov A, Antunes A, et al. 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci. Adv. 3:1–14.

Fontaine MC, Pease JB, Steele A, Waterhouse RM, Neafsey DE, Sharakhov I V., Jiang X, Hall AB, Catteruccia F, Kakani E, et al. 2015. Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science 80. 347:1–6.

Forsythe ES, Nelson AD, Beilstein MA. Biased gene retention in the face of massive nuclear introgression obscures species relationships (In revision). bioRxiv. https://www.biorxiv.org/content/early/2018/10/18/197087?%3Fcollection=

760 Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W,
761 Fritz MH-Y, et al. 2010. A Draft Sequence of the Neandertal Genome. Science (80-.).
762 328:710–722.
763 Hibbins MS, Hahn MW. 2019. The timing and direction of introgression under the multispecies
764 network coalescent. Genetics 211:1059–1073.
765 Hudson R. 2002. Ms a Program for Generating Samples Under Neutral Models. Bioinformatics
766 18:337–338.
767 Hudson RR, Kaplan NL. 1985. Statistical properties of the number of recombination events in
768 the history of a sample of DNA sequences. Genetics 111:147–164.
769 Huson DH, Kl T, Lockhart PJ, Steel M a, Klopper T, Lockhart PJ, Steel M a, Kl T, Lockhart PJ,
770 Steel M a. 2005. Reconstruction of Reticulate Networks from Gene Trees. Res. Comput.
771 Mol. Biol. Proc. 3500:233–249.
772 International Human Genome Sequencing Consortium. 2001. Initial sequencing and analysis of
773 the human genome. Nature 409:860–921.
774 Joly S, McLenachan PA, Lockhart PJ. 2009. A Statistical Approach for Distinguishing
775 Hybridization and Incomplete Lineage Sorting. Am. Nat. 174:E54–E70.
776 Kuhlwilm M, Gronau I, Hubisz MJ, De Filippo C, Prado-Martinez J, Kircher M, Fu Q, Burbano
777 HA, Lalueza-Fox C, De La Rasilla M, et al. 2016. Ancient gene flow from early modern
778 humans into Eastern Neanderthals. Nature 530:429–433.
779 Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R.
780 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078–2079.
781 Liu KJ, Dai J, Truong K, Song Y, Kohn MH, Nakhleh L. 2014. An HMM-Based Comparative
782 Genomic Framework for Detecting Introgression in Eukaryotes. PLoS Comput. Biol. 10.
783 Maddison WP, Knowles LL. 2006. Inferring Phylogeny Despite Incomplete Lineage Sorting.
784 Syst. Biol. 55:21–30.
785 Mallet J, Besansky N, Hahn MW. 2016. How reticulated are species? BioEssays 38:140–149.
786 Martin SH, Davey JW, Jiggins CD. 2015. Evaluating the use of ABBA-BABA statistics to locate
787 introgressed loci. Mol. Biol. Evol. 32:244–257.
788 Orive ME, Barton NH. 2002. Associations between cytoplasmic and nuclear loci in hybridizing
789 populations. Genetics 162:1469–1485.
790 Paradis E, Claude J, Strimmer K. 2004. APE: Analyses of phylogenetics and evolution in R
791 language. Bioinformatics 20:289–290.
792 Pease JB, Hahn MW. 2015. Detection and Polarization of Introgression in a Five-Taxon
793 Phylogeny. Syst. Biol. 64:651–662.
794 Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S, Heinze A, Renaud G,
795 Sudmant PH, De Filippo C, et al. 2014. The complete genome sequence of a Neanderthal
796 from the Altai Mountains. Nature 505:43–49.
797 Rambaut A, Grassly NC. 1997. Seq-Gen: an application for the Monte Carlo simulation of DNA
798 sequence evolution along phylogenetic trees. Cabios 13:235–238.
799 Rieseberg LH, Soltis DE. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants.
800 Evol. trends Plants 5:65–84.
801 Rieseberg LH, Whitton J, Linder CR. 1996. Molecular marker incongruence in plant hybrid
802 zones and phylogenetic trees. Acta Bot. Neerl. 45:243–262.
803 Rosenzweig BK, Pease JB, Besansky NJ, Hahn MW. 2016. Powerful methods for detecting
804 introgressed regions from population genomic data. Mol. Ecol.:2387–2397.
805 Schrider D, Ayroles J, Matute DR, Kern AD. 2018. Supervised machine learning reveals

introgressed loci in the genomes of Drosophila simulans and D. sechellia. PLoS Genet. 10:1–29.
Slon V, Mafessoni F, Vernot B, de Filippo C, Grote S, Viola B, Hajdinjak M, Peyrégne S, Nagel S, Brown S, et al. 2018. The genome of the offspring of a Neanderthal mother and a Denisovan father. Nature.
Soltis DE, Soltis PS. 2003. The role of phylogenetics in comparative genetics. Plant Physiol. 132:1790–1800.
Stebbins GL. 1968. The Significance of Hybridization for Plant Taxonomy and Evolution. Taxon 18:26–35.
Stenz NWM, Larget B, Baum DA, Ané C. 2015. Exploring tree-like and non-tree-like patterns using genome sequences: An example using the inbreeding plant species Arabidopsis thaliana (L.) heynh. Syst. Biol. 64:809–823.
Than C, Ruths D, Nakhleh L. 2008. PhyloNet: a software package for analyzing and reconstructing reticulate evolutionary relationships. BMC Bioinformatics 9:322.
The Chimpanzee Sequencing Consortium. 2005. Initial sequence of the chimpanzee genome and comparison with the human genome. Nature 437:69–87.
Whitney KD, Randell R a, Rieseberg LH. 2006. Adaptive introgression of herbivore resistance traits in the weedy sunflower Helianthus annuus. Am. Nat. 167:794–807.
Yakimowski SB, Rieseberg LH. 2014. The role of homoploid hybridization in evolution: A century of studies synthesizing genetics and ecology. Am. J. Bot. 101:1247–1258.
Zheng Y, Janke A. 2018. Gene flow analysis method, the D-statistic, is robust in a wide parameter space. BMC Bioinformatics 19:1–19.

SUPPLEMENTAL INFORMATION

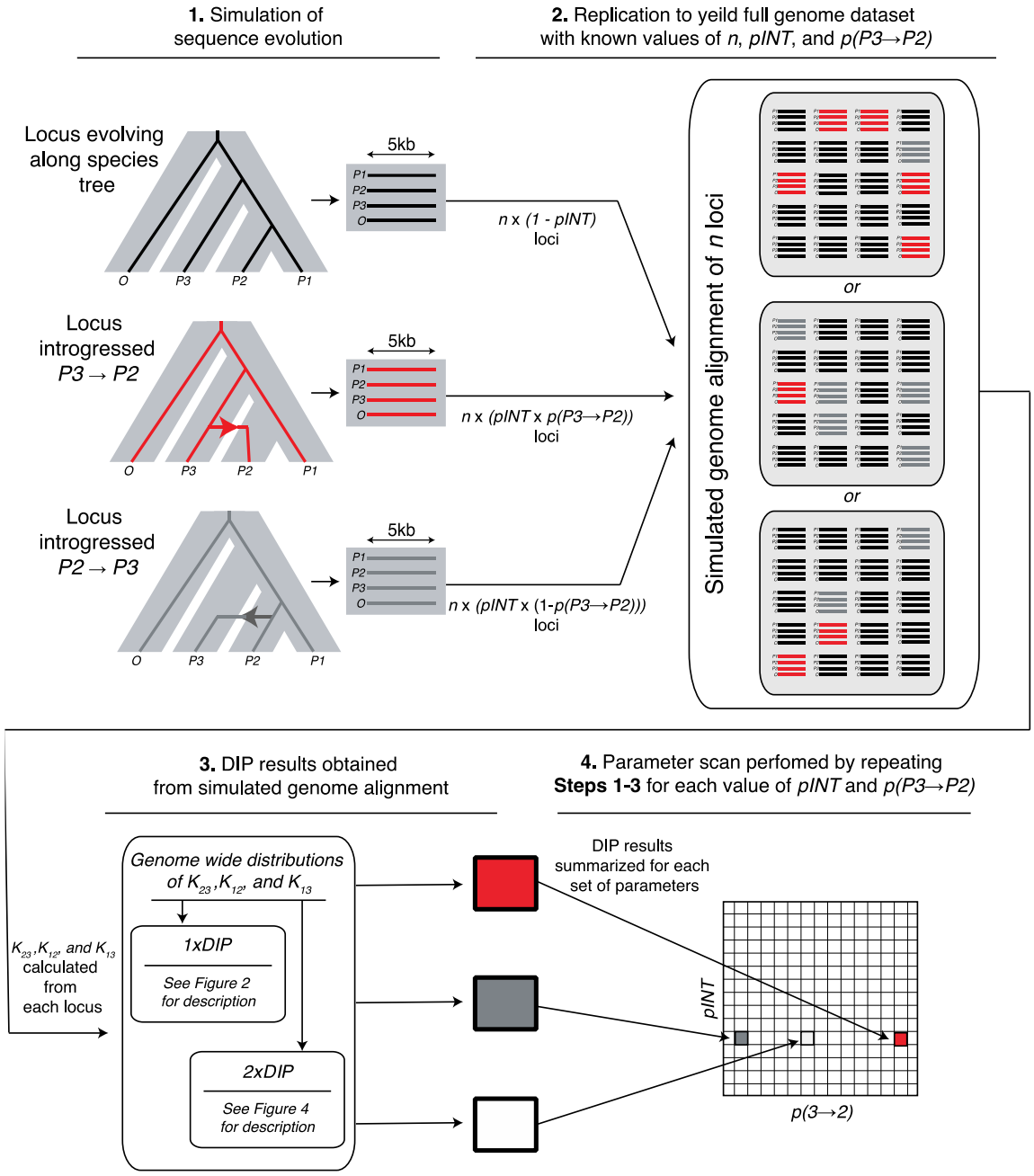


Fig. S1. Schematic of the workflow used to simulate introgression across a genome and perform *DIP*. (1) Each locus is evolved along the species tree or along a path of introgression and used to generate a 5000-bp alignment using *ms* and *seq-gen* similar to (Martin et al. 2015). **(2)** Step 1 was repeated to yield a full genome of n=5000 loci in which n x *pINT* loci were introgressed and the remaining loci evolved along the species tree. For example, a genome in which half of all genes were not transferred while the other half were transferred *P3⇒P2* would be generated with: n=5000, *pINT* = 0.5, *p(P3⇒P2)* = 1.0. **(3)** Different steps in the *DIP* pipeline are performed on the simulated genome. **(4)** Steps 1-3 are repeated for each combination of *pINT* and *p(P3⇒P2)*. Each pixel in a parameter scan graph represents one or more runs of Steps 1-3.

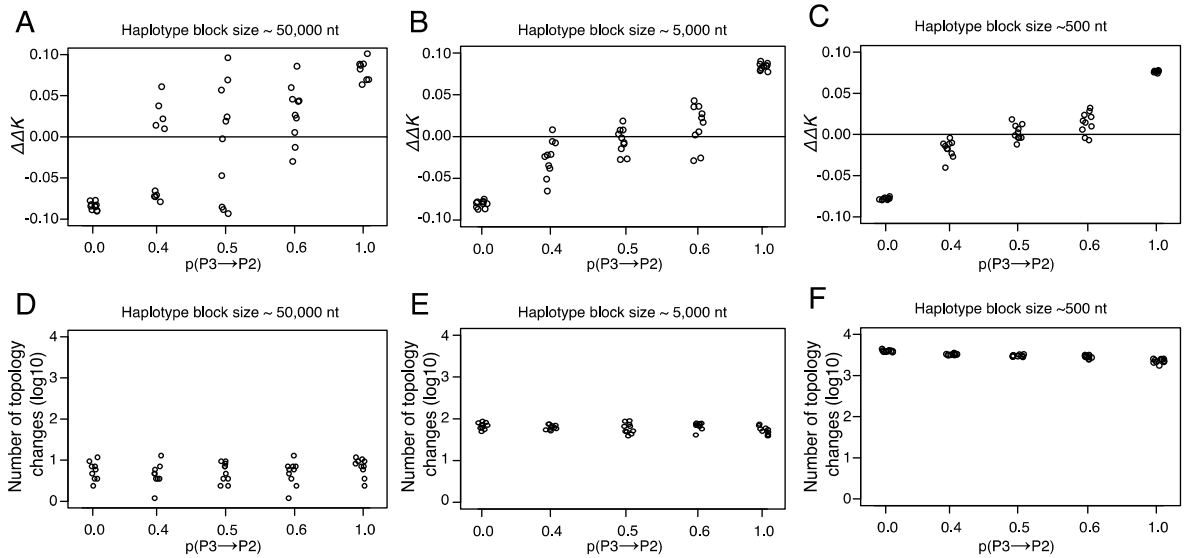


Fig. S2. 2xDIP analyses of introgression simulated with recombination. Genomes were simulated using an alternative simulation strategy that includes recombination (see Methods). Three different recombination rates were used, r=7 (A and D), r=70 (B and E), and r=700 (C and F), resulting in haplotypes of different sizes. Five replicate genomes were simulated for each value of *p(P3⇒P2)*. To provide a understanding of the haplotype blocks with diagnostic synapomorphies that exist at different levels of recombination, the number of instances in which neighboring blocks exhibit different topologies (i.e. topology changes) were summed (D-F). Note that number of topology changes differs from haplotype blocks because some SNPs occurred along branches that don't define the topology (i.e. not all SNPs represent diagnostic synapomorphies).

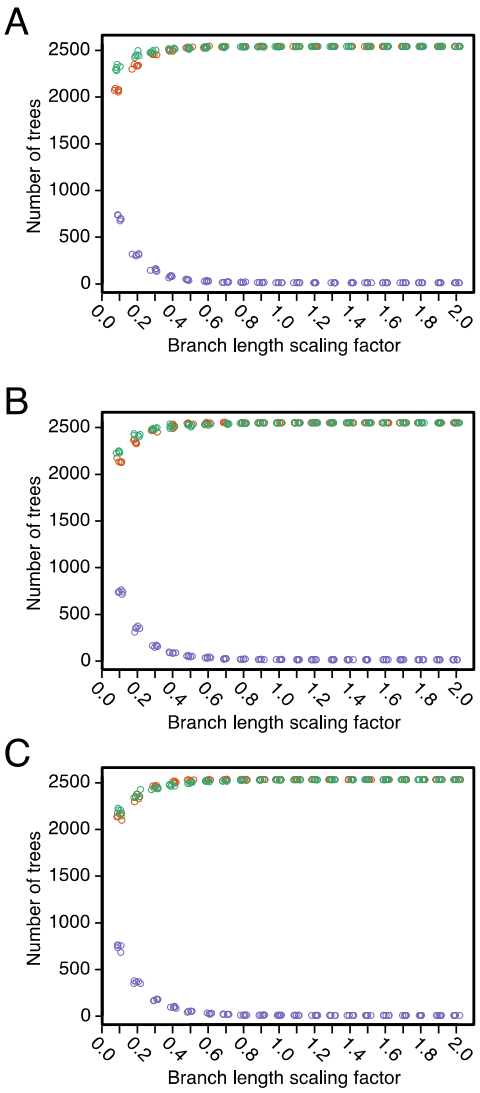


Fig. S3. Gene tree topologies inferred from simulated genomes. Gene tree counts for genomes simulated with different branch lengths (x-axes) and *p(P3⇒P2)* values of 0.6 (**A**), 0.5 (**B**), and 0.4 (**C**). Each point represents the number of trees displaying a given topology from a replicate genome. ((*P1,P2*),*P3*)*, orange;* ((*P2,P3*),*P1*)*, green;* ((*P1,P3*),*P2*)*, purple.* These same simulated genomes were analyzed in Fig. 6.

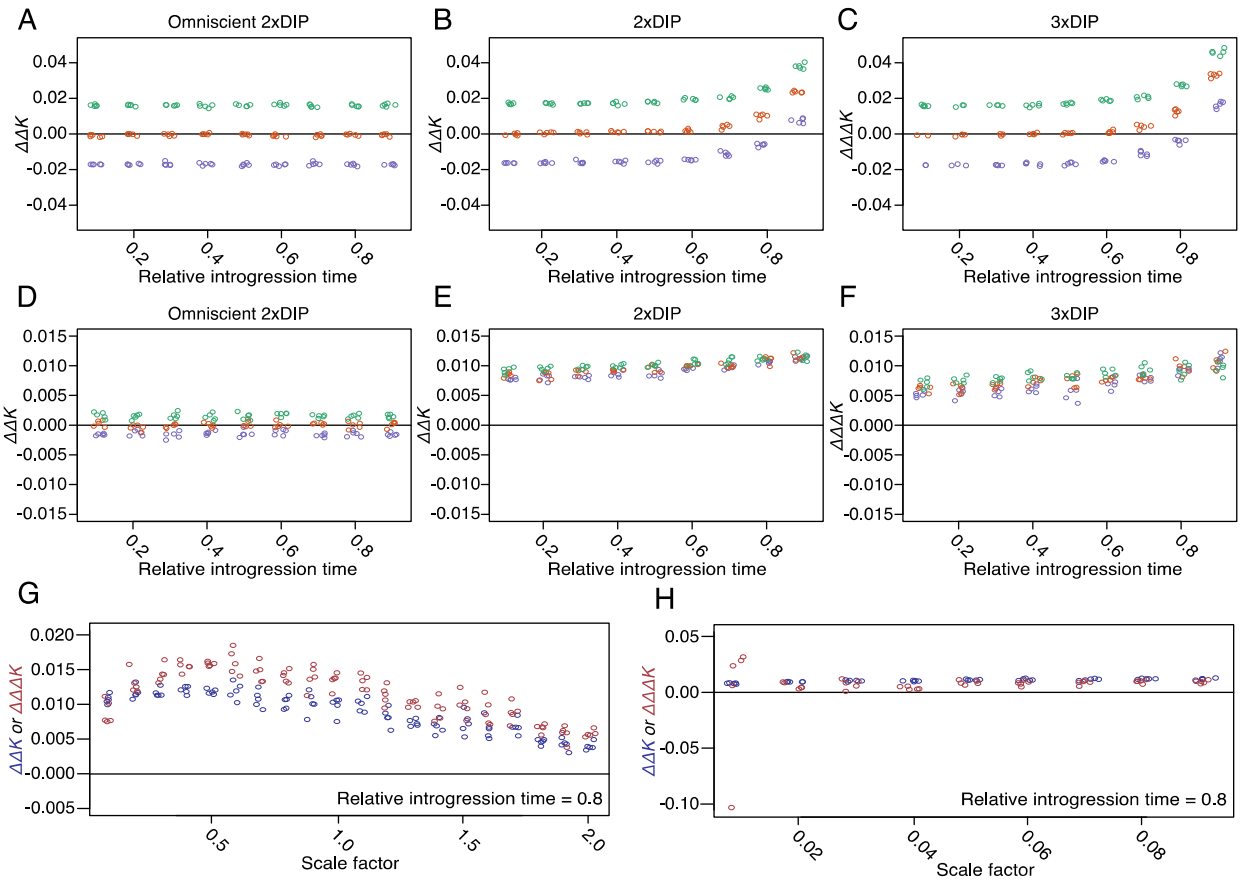


Fig. S4. Simulations of relative introgression timing. (A-F) Genomes were simulated using two different scaling factors, SF=1 (A-C) and SF=0.1 (D-F) to determine the timing of speciation events. While the timing of all speciation events adhere to the given scaling factor, the timing of the introgression event (T_{INT}) was set as a fraction of the timing of the most recent speciation event (T_{α}) (i.e. relative introgression time). High relative introgression time indicates introgression occurred directly following speciation whereas low relative introgression time indicates introgression occurred a long duration of time following speciation. Three different types of asymmetrical/symmetrical bidirectional introgression were simulated for each relative introgression time, *p(P3⇒P2)*= 0.4 (purple), *p(P3⇒P2)* = 0.5 (orange), and *p(P3⇒P2)* = 0.6 (green). Omniscient *2×DIP* (A and D), non-omniscient *2×DIP* (B and E), and *3×DIP* (C and F) were performed on all genomes. (G-H) *ΔΔK* (blue) and *ΔΔΔK* (dark red) measurements for genomes simulated with *p(P3⇒P2)* = 0.5 and different scaling factors. Five replicate genomes were simulated for each parameter value.

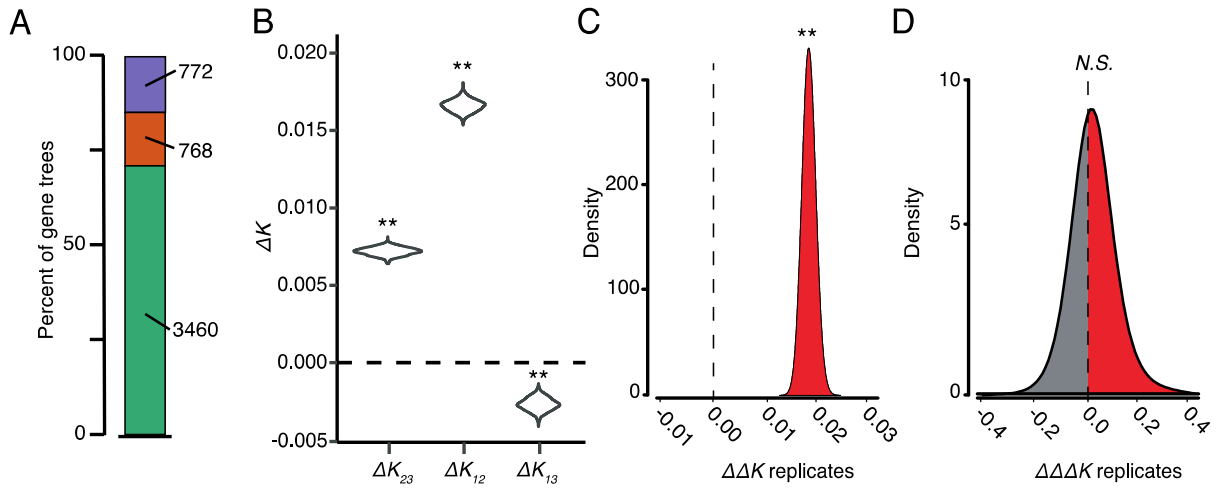


Fig. S5. *DIP* analysis of a genome with incomplete lineage sorting but no introgression. A genome alignment was simulated with *pINT* set to zero using the scaling factor 0.1 (see Fig. 1 and Fig. 6). Therefore, all loci with topologies that conflict with species tree are the result of ILS and not introgression (**A**) The topologies of neighbor joining trees inferred from 5000 simulated loci. ((*P1,P2*),*P3*), *green;* ((*P2,P3*),*P1*), *orange;* ((*P1,P3*),*P2*), *purple.* (**B-D**) 1×*DIP* (**B**), 2×*DIP* (**C**) and 3×*DIP* (**D**) analysis of the genome alignment.

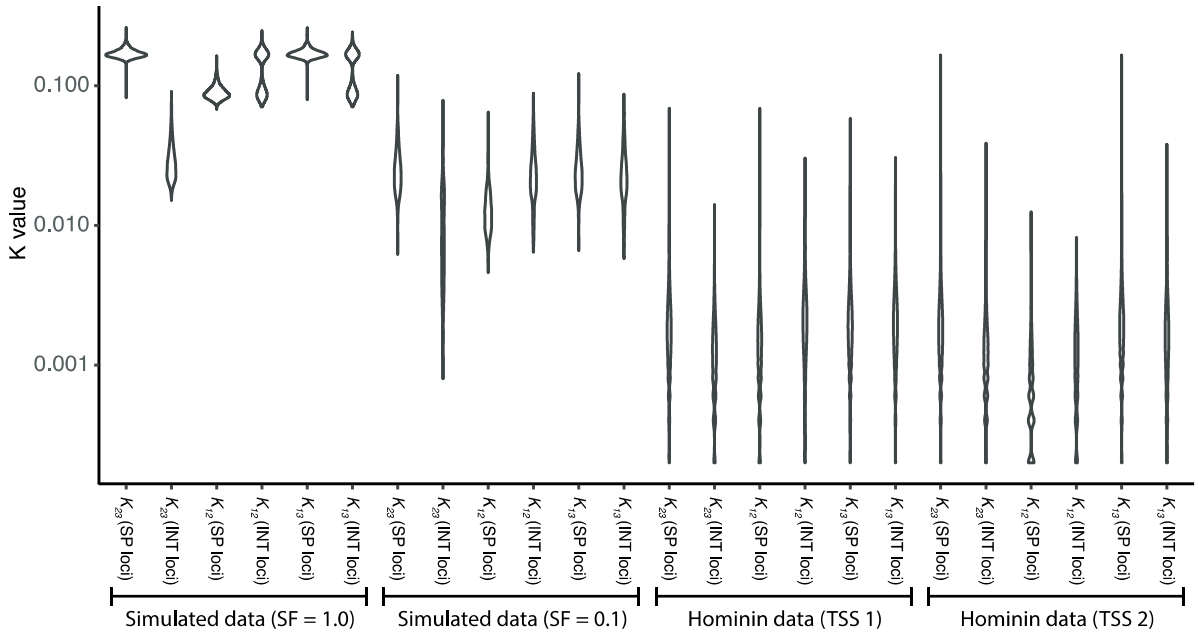


Fig. S6. Sequence divergence measures from simulated and Hominin data. Violin plot showing distributions of pairwise divergence values for inferred loci displaying the species (SP) and introgressed (INT) topology (see Fig. 1 and 2). Both simulated datasets were simulated with *pINT*=0.5 and *p(P3⇒P2)*=0.5.