Bioinformatics, 37(24), 2021, 4744–4755
doi: 10.1093/bioinformatics/btab531
Advance Access Publication Date: 16 July 2021
Original Paper

OXFORD

Genetics and population analysis

LINADMIX: evaluating the effect of ancient admixture events on modern populations

Lily Agranat-Tamir[1,2], Shamam Waldman[3], Naomi Rosen[1], Benjamin Yakir[2,*], Shai Carmi[3,*] and Liran Carmel [1,*]

[1]Department of Genetics, The Alexander Silberman Institute of Life Sciences, Faculty of Science, The Hebrew University of Jerusalem, Edmond J. Safra Campus, Givat Ram, Jerusalem 9112102, Israel, [2]Department of Statistics and Data Science, The Hebrew University of Jerusalem, Jerusalem 9112102, Israel and [3]Braun School of Public Health and Community Medicine, The Hebrew University of Jerusalem, Jerusalem 9112102, Israel

*To whom correspondence should be addressed.
Associate Editor: Russell Schwartz

Received on November 8, 2020; revised on June 25, 2021; editorial decision on June 30, 2021; accepted on July 15, 2021

Abstract

Motivation: The rise in the number of genotyped ancient individuals provides an opportunity to estimate population admixture models for many populations. However, in models describing modern populations as mixtures of ancient ones, it is typically difficult to estimate the model mixing coefficients and to evaluate its fit to the data.

Results: We present LINADMIX, designed to tackle this problem by solving a constrained linear model when both the ancient and the modern genotypes are represented in a low-dimensional space. LINADMIX estimates the mixing coefficients and their standard errors, and computes a *P*-value for testing the model fit to the data. We quantified the performance of LINADMIX using an extensive set of simulated studies. We show that LINADMIX can accurately estimate admixture coefficients, and is robust to factors such as population size, genetic drift, proportion of missing data and various types of model misspecification.

Availability and implementation: LINADMIX is available as a python code at https://github.com/swidler/linadmix.

Contact: liran.carmel@huji.ac.il or shai.carmi@mail.huji.ac.il or benjamin.yakir@mail.huji.ac.il

Supplementary information: Supplementary data are available at *Bioinformatics* online.

1 Introduction

The advent of ancient DNA (aDNA) sequencing technologies gave rise to an ever-growing number of ancient genomes, thus providing an opportunity to tackle historical questions from a genetic perspective. To date, many ancient humans have been sequenced at ≈1.2 M single-nucleotide polymorphisms (SNPs) targeted using a capture array. However, sequencing depth is typically low, resulting in many uncalled SNPs in each sample. Even for called SNPs, only a single allele is usually reported, randomly selected from among the relevant reads. Therefore, many tools designed to study ancestry patterns in modern genomic data need to be adapted or re-evaluated for aDNA data.

A popular tool for the analysis of past admixture events is *qpAdm* (Haak *et al.*, 2015). *qpAdm*, which is based on f_4-*statistics* (Patterson *et al.*, 2012), tests whether a specific population, called the *target*, can be modeled as a mixture of other pre-specified populations, called the *sources*. In addition, it estimates the relative contribution of every source, called the *mixing coefficients*, as well as their standard errors. Technically, *qpAdm* considers a *left* list of populations, containing the target and source populations and a *right* list of populations sometimes called *outgroups*. When both target and sources are ancient populations, *qpAdm* is the most frequently used method to test such admixture models. However, the use of *qpAdm* is not recommended for a modern target population (Harney et al., 2021). Two other methods that are very common in aDNA analyses are *smartpca* (Patterson *et al.*, 2006) and ADMIXTURE (Alexander *et al.*, 2009). However, these methods do not directly test for admixture between populations, and are designed to examine individual genomes rather than whole populations. *smartpca* typically projects ancient samples onto a plane spanned by the first principal components of modern individuals. This allows for a qualitative examination of the relationship between populations. The ADMIXTURE algorithm takes a set of samples, each can be either ancient or modern, and estimates the proportion of each genome descending from each of *K* hypothetical ancestral populations.

Many methods have been developed to estimate the structure (e.g. Gaspar and Breen, 2019), ancestry (e.g. Taravella Oill *et al.*, 2021) and geography (e.g. Battey *et al.*, 2020) of modern populations and individuals based on their DNA. These methods use modern genomes, and hence enjoy full genotype information. Many methods use information on allele frequency, but some also use

other types of information. To give a few examples, ChromoPainter is a haplotype-based method (Lawson *et al.*, 2012); and iAdmix (Bansal and Libiger, 2015) finds the relative contribution of given source populations to an individual using both allele frequencies of the source populations and sequence reads of the target individual. GRAF-pop (Jin *et al.*, 2019) is an exception in the sense that it integrates data from many studies or even different SNP arrays. GRAF-pop then finds the proportion of ancestry an individual has from each source. ARCHes (Noto *et al.*, 2020) is based on local ancestry inference. However, it is haplotype-based, and therefore less suited to aDNA. DyStruct (Joseph and Pe'er, 2019) analyzes both modern and ancient samples. However, it does not test specific models but rather infers shared ancestry, similarly to ADMIXTURE. As opposed to ADMIXTURE, DyStruct takes drift into account and therefore makes use of temporally sampled populations.

Many methods perform ancestry inference based on the output of PCA or ADMIXTURE as features. Examples include PopInf (Taravella Oill *et al.*, 2021), which assigns ancestries to individuals based on PCA output, and GPS (Elhaik *et al.*, 2014), which takes the ADMIXTURE output, along with the geographic location of reference populations, to find the geographic origin of individuals.

Taken together, the existing set of tools, some of which we have highlighted above, is incomplete for analyzing admixture models when the target is a modern population, but the source populations are ancient. We developed LINADMIX with the intention to offer a remedy to this situation. We offer formal ways to test the validity of the examined model and to estimate the mixing coefficients. To this end, we use ADMIXTURE as a low-dimensional representation of the target and source populations, and find the mixing coefficients by solving a linear mixing model using least squares. reAdmix (Kozlov *et al.*, 2015) is a similar approach, but it seeks the best sources out of a set of reference populations. In addition, reAdmix neither provides standard errors nor an assessment of the validity of the model.

We introduced the main features of LINADMIX in a previous publication (Agranat-Tamir *et al.*, 2020), where we studied modern Levantine populations as mixtures of ancient populations. In that paper, we have shown preliminary simulations demonstrating the accuracy of LINADMIX in a specific scenario. Here, we evaluate the performance of LINADMIX using a comprehensive simulation study that allows us to quantify how the performance of LINADMIX depends on various model parameters. In addition, we devise a formal test for the validity of the admixture model. We find that LINADMIX performs well even for high levels of missing data in the ancient source populations. In addition, we show that while ADMIXTURE does not model genetic drift, LINADMIX is robust to drift. We also show that LINADMIX can distinguish between valid and invalid models, and correctly estimates no contribution from source populations that were not used to simulate the target. Finally, we show that while the overall performance of LINADMIX is better when the sources are highly diverged (as expected), it performs well even when the source populations are genetically similar.

2 Materials and methods

2.1 Simulating admixed genomes

To test the performance of LINADMIX, we simulated admixed genomes (target populations) in predetermined mixing proportions. We used modern genomes as the basic building blocks of the simulations and down-sampled some genomes to mimic ancient DNA data. Details are given in Supplementary Text S1.

2.2 Calculating F_{ST} values

We calculated F_{ST} values between pairs of populations using *smartpca* (Patterson *et al.*, 2006) on the Human Origins array data.

2.3 The LINADMIX algorithm

We provide here a detailed description of the three parts of the algorithm.

2.3.1 Estimating the mixing coefficients

The ADMIXTURE algorithm describes a set of samples as a mixture of K ancestral hypothetical populations, whose allele frequencies are also estimated by the algorithm (Alexander *et al.*, 2009). For each sample l, ADMIXTURE estimates the fraction of its genome, $q^l(k)$, that is contributed by ancestral population k. Clearly, $\sum_k q^l(k) = 1$. Therefore, ADMIXTURE represents a sample l as a vector of length K,

$$q^l = (q^l(1), q^l(2) \dots q^l(K))^T.$$

LINADMIX builds on this representation. Similar to previous works (Kozlov *et al.* 2015, Leslie *et al.*, 2015) it seeks to explain the distribution of genotypes of a target population as a linear mixture of given source populations, using the ADMIXTURE representations of both target and sources.

Formally, we look at the ADMIXTURE representation of n_i samples from population i and compute a representation vector of the population by taking the average of the samples' representation vectors: $q_i = (q_i(1) \dots q_i(K))^T = \frac{1}{n_i}\sum_{l=1}^{n_i} q_i^l$. Here q_i^l is the ADMIXTURE representation of the lth individual in the ith population.

Let q_t be the ADMIXTURE representation of a *target population* and let $q_1, \dots, q_n$ be the ADMIXTURE representations of n *source populations*. Under our model, the target population is a mixture of the sources:

$$q_t = \alpha_1 q_1 + \alpha_2 q_2 + \dots + \alpha_n q_n \quad (1)$$

where $\alpha_1, \dots, \alpha_n$ are the *mixing coefficients*. We estimate these mixing coefficients as the solutions of the constrained non-negative least squares problem:

$$\min_{\alpha} ||Q_s \alpha - q_t||_2^2, \text{ subject to } \sum_{i=1}^{n} \alpha_i = 1 \text{ and } \alpha_i \geq 0, \quad (2)$$

where $\alpha = (\alpha_1, \dots, \alpha_n)^T$ and Q_s is the $K \times n$ matrix $Q_s = (q_1, \dots, q_n)$. In other words, we search for the coefficients $\alpha_1, \dots, \alpha_n$ that most closely satisfy Equation (1) simultaneously for each of the K ADMIXTURE components.

2.3.2 Estimating the standard errors

We use parametric bootstrapping to find the standard errors of the mixing coefficients $\alpha = (\alpha_1, \dots, \alpha_n)^T$. To this end, let $\left\{q_i^l\right\}$ be the ADMIXTURE representations (of both target and sources individuals) that are used to produce α and let $\left\{\xi_i^l\right\}$ be the theoretical ancestral components (the parameters estimated by ADMIXTURE). We generate a series of bootstrap copies $\left\{q_i^{l,(b)}\right\}$ of the $\left\{q_i^l\right\}$ vectors and compute the solution $\alpha^{(b)}$ for each copy. The standard error of α is obtained from the empirical standard deviation of the bootstrap copies. In the results reported in this manuscript, we used 1000 bootstrap repetitions.

We assume that each ξ_i^l is normally distributed, namely $\xi_i^l \sim N(\xi_i, \Sigma_i)$, where ξ_i is the ADMIXTURE parameter of the population, and Σ_i is the covariance matrix representing the variation in the ADMIXTURE profiles between individuals in the population. We assume that in each bootstrap repetition b and for each sample l from population i, the vector $q_i^{l,(b)}$ is drawn from $N(q_i, \hat{\Sigma}_i)$, where $\hat{\Sigma}_i = cov(q_i^1, \dots, q_i^{n_i})$ is the sample $K \times K$ covariance matrix between the K ADMIXTURE components. However, we found empirically that for small samples, the components of $\hat{\Sigma}_i$ are underestimated. As a heuristic approach to obtain a more unbiased estimate of $\hat{\Sigma}_i$, we drew $\left\{q_i^{l,(b)}\right\}$ from $N\left(q_i, \hat{\Sigma}_i + \hat{\Sigma}^l\right)$. Here, $\hat{\Sigma}^l = \left[I\left(q_i^l\right)\right]^{-1}$, where I is the $K \times K$ Fisher information matrix associated with the maximum-likelihood estimation of the ADMIXTURE components. Here, $\hat{\Sigma}^l$ is the variance of the estimated ADMIXTURE components around their true value (Supplementary

Text S2). We found empirically (Supplementary Text S3) that the combined variance performs better and have thus used it as default in our simulations. We present additional details in Supplementary Text S4.

2.3.3 Hypothesis testing

To provide a measure of the validity of any specific model, we developed a way to compute *P*-values for the null hypothesis that the mixture model is correct, namely that the ADMIXTURE representation of the target population is a weighted average of the ADMIXTURE representations of the source populations. The null hypothesis is H_0: $\xi_t = \Xi_s\alpha$ for some $\alpha = (\alpha_1, \ldots, \alpha_n)^T$ satisfying $\sum_{i=1}^{n} \alpha_i = 1$ and $\alpha_i \geq 0$, where ξ_t and Ξ_s are the parameters estimated by ADMIXTURE. ξ_t is the expectation of the distribution of the ξ_i^l vectors of the target population, and Ξ_s is the matrix of expectations of the distributions of the ξ_i^l vectors of the source populations. In practice, we take the α estimated by the constrained non-negative least squares (Equation 2). To compute a *P*-value, we again perform a parametric bootstrap where we produce a large number of copies $\left\{q_i^{l,(b)}\right\}$ from the null distribution H_0.

Specifically, we denote $q_0 = Q_s\alpha$ and define a test statistic

$$W = \sum_k \frac{\left(q_t(k) - q_0(k)\right)^2}{q_0(k)}.$$

for the observed q_t and Q_s and the estimated α.

In each bootstrap run, we draw, as described above, the source vectors that make the matrix $Q_s^{(b)}$, as well as the target vector $q_t^{(b)}$. The vectors of the target population are generated from $N\left(q_0^{(b)}, \hat{\Sigma}_t\right)$ or, in the case of small samples as in this paper, from $N\left(q_0^{(b)}, \hat{\Sigma}_t + \hat{\Sigma}^l\right)$, where $q_0^{(b)} = Q_s^{(b)}\alpha$. Finally, in each iteration we calculate the test statistic

$$W^{(b)} = \sum_k \frac{\left(q_t^{(b)}(k) - q_0^{(b)}(k)\right)^2}{q_0^{(b)}(k)}.$$

The *P*-value is computed as

$$p = \frac{\left\{W^{(b)} \geq W\right\} + 1}{B + 1}, \tag{3}$$

where B is the number of bootstrap repetitions (all results reported in this manuscript are based on 10 000 repetitions). Some more implementation details are given in Supplementary Text S5.

3 Results

3.1 An overview of LINADMIX and the validation approach

We developed LINADMIX to test and estimate the parameters of admixture models in which the target is a modern population and the sources are ancient. Specifically, LINADMIX tests the validity of a hypothesis that a given target population can be represented as a weighted average of a given set of source populations. In addition, LINADMIX finds the relative contributions of the different sources to the target population, which we refer to as the *mixing coefficients*, as well as their standard errors.

We conducted a series of simulation studies in order to assess LINADMIX's ability to distinguish between valid and invalid models and to estimate mixing coefficients in various scenarios, which differed in the quality of the genetic data, the time since admixture and the relation between the sources. We generated artificial genomes that mimic the given scenario, ran ADMIXTURE, applied LINADMIX to the simulated genomes, and compared its estimations to the parameters that were used in the simulation.

To generate the artificially admixed target genomes, we mixed segments from samples from a set of modern source populations, in predetermined proportions, and with the segments' lengths corresponding to a predetermined time since admixture (20–160 generations ago). For each scenario, we generated several simulated target populations with different mixing proportions of the sources. To avoid any risk of overfitting, when running ADMIXTURE prior to applying LINADMIX, we only used samples that were not used to simulate the targets.

The source genomes we used in the simulations were modern, whereas our goal was to study admixture of ancient sources. Thus, we randomly deleted one of the alleles in each marker of each genome to be used as a source. This way, we generated pseudo-haploids, which is a typical data structure in ancient DNA. We also randomly deleted some of the SNPs to mimic fractions of missing data realistic in ancient DNA analyses.

We ran ADMIXTURE on the source and target samples, and then applied LINADMIX to produce estimated mixing coefficients, their standard errors and a *P*-value testing the validity of the model. We measured the estimation error as the absolute difference (in percentage points) between the LINADMIX estimate of the mixing coefficient and the actual mixing proportions. These absolute differences were used to produce two summary statistics that measure the accuracy of estimation: the average and the maximum (of these differences) over the different simulated populations in each scenario. In addition, we computed the average of the standard errors estimated by LINADMIX over the different populations. Finally, we evaluated whether the observed *P*-value was consistent with the simulations, namely that it is higher than a threshold (0.05) when simulating from the null model and lower than the threshold when the model is invalid.

3.2 The effect of missing data in the source populations

Ancient DNA has high levels of missing data, manifested as pseudo-haploid genotype calls and complete absence of many SNPs. We simulated these two forms of missing data in the genomes of the source populations. The proportion of missing SNPs in previously reported ancient genomes (Agranat-Tamir *et al.*, 2020) was on average 68% ± 27%, and we have therefore evaluated the performance of LINADMIX using proportions of 0.4, 0.6, 0.8 and 0.9 missing SNPs.

For the simulations, we considered two distinct pairs of source populations. One pair was Jordanians and English and the other pair was Spanish and Russians. Both pairs are genetically distinct, with the Spanish and Russians being somewhat more closely related (F_{ST} values of 0.013 and 0.008 for Jordanians-English and Spanish-Russians, respectively; Supplementary Table S1). For each pair we considered three different mixing proportions of 0.2:0.8, 0.5:0.5 and 0.8:0.2. The simulations assumed that 50 generations have passed since the admixture event that formed the target population.

As expected, we observed that the performance of LINADMIX deteriorated with increasing proportion of missing data (Fig. 1, Supplementary Table S2). In both scenarios (Jordanians-English and Spanish-Russians), LINADMIX was able to correctly assess the validity of the models as long as the fraction of missing SNPs did not exceed 60%. At higher fractions of missing data, LINADMIX showed inferior performance when the source populations were Jordanians and English. However, it demonstrated a very stable performance when the source populations were Spanish and Russians, obtaining low estimation errors and higher *P*-values even for 80% missing values (in 2/3 models) and for 90% missing values (in 1/3 models; Fig. 1, Supplementary Table S2).

In summary, the performance of LINADMIX was good in the simulations, producing maximum and average errors of ∼4% points and ∼2% points (for English-Jordanians and Spanish-Russians admixtures, respectively) for up to 60% missing data. Based on these results we chose 0.6 as the proportion of missing SNPs in the source samples for the rest of the simulations.

3.3 Admixture of two closely related populations

An admixture event is more difficult to detect when the source populations are genetically similar. To evaluate the performance of

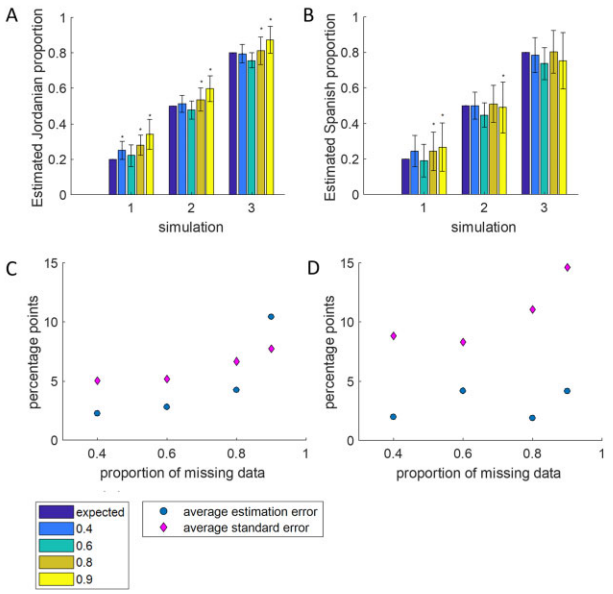


Fig. 1. The effect of the proportion of missing data in the source populations on LINADMIX. The estimated Jordanian (**A**) or Spanish (**B**) contributions in three simulations where either Jordanians and English (A) or Spanish and Russians (B) were mixed in varying proportions. Estimations were carried out using different proportions of missing SNPs in the sources as indicated in the bottom-left legend. An asterisk indicates that the *P*-value of the model was below 0.05. (**C, D**) LINADMIX performance, measured in several ways, for the different proportions of missing SNPs on the Jordanian-English (C) and the Spanish-Russian (D) simulations

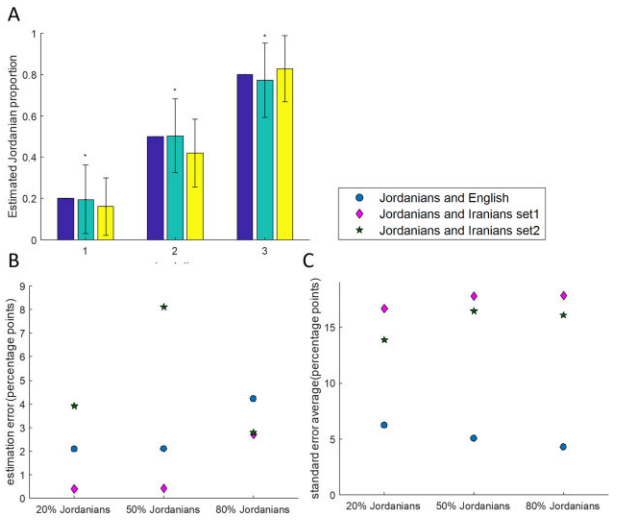


Fig. 2. Admixture of closely related populations. (**A**) The estimated Jordanian contribution in three simulations where Jordanians and Iranians were mixed in varying proportions. An asterisk indicates that the *P*-value of the model was below 0.05. (**B, C**) LINADMIX performance, measured in several ways, comparing the simulations of Jordanians-English with those of Jordanians-Iranians. In the y axis are (B) the absolute error of estimation, and (C) the standard errors of the estimates

LINADMIX in this setting, we simulated target populations with mixed modern Jordanians and Iranians ancestries (Fig. 2, Supplementary Table S3). These source populations are more related to each other in comparison with both the Jordanian-English and the Spanish-Russian pairs (Supplementary Table S1).

We considered again 0.2:0.8, 0.5:0.5 and 0.8:0.2 admixture ratios. Here, the estimation errors were low—a maximum of 2.7% points and an average of 1.2% points, but the *P*-values were small (0.0005–0.0019), indicating a lack of fit to the model. To determine whether this is a general observation, we performed another set of three simulations. In the second set, the *P*-values were much higher (above 0.05), but so were the estimation errors, with a maximum of 8.1% points and an average of 4.9% points (Fig. 2, Supplementary Table S3). Notably, we expect such estimation differences across simulations, as the standard errors of the estimators are high, on average 17.4% and 15.4% for the two simulation sets, respectively (Fig. 2, Supplementary Table S3). These high standard errors reflect the similarity of the sources.

In summary, for closely related source populations, LINADMIX was still able to estimate the mixing coefficients, albeit with large standard errors.

3.4 The effect of time since admixture

The time that passed since an admixture event may affect the target genome via recombination, which shortens the length the segments that originated from each source population, as well as via genetic drift that results in a random change of allele frequencies.

3.4.1 The effect of segment lengths

Segment lengths are not expected to affect LINADMIX, as LINADMIX relies on the output of ADMIXTURE, which only uses SNP frequencies and not haplotype information. Indeed, we found that segment lengths do not affect LINADMIX (Supplementary Text S6). Therefore, in all subsequent simulations we assumed that the admixture event occurred 50 generations ago.

3.4.2 Genetic drift

To examine the effect of drift, we ran LINADMIX with source populations that are related, but not identical, to the source populations that were used to simulate the target genomes. The divergence between the true source populations and the proxy populations that were used as sources in LINADMIX mimics the situation where genetic drift took place in the source populations since the actual time of admixture.

We chose Syrians or Palestinians as proxies for Jordanians, and Germans or French as proxies for English (Supplementary Table S1). We used four individuals from each replacement population as input to LINADMIX, as this was the number of individuals used in the original English-Jordanians admixture simulations.

We found that using a proxy source population had only a minor effect on LINADMIX (Fig. 3, Supplementary Table S4A). Replacing English by Germans did not affect the estimates. Replacing English by French had a larger effect, increasing the maximum estimation error from 4.4 to 7.6 percentage points. Likewise, the average estimation error increased from 2.9 to 6.8 percentage points. (Fig. 3A, E, Supplementary Table S4A). Compared to the original Jordanian-English model, the Syrian-English model reduced the average error of the estimates by ∼1% point, whereas the Palestinian-English model increased the average error by ∼1% point (Fig. 3B, E, Supplementary Table S4A). In all cases the standard errors were only slightly affected. In all tests but two (10/12), the *P*-values were above 0.05, compatible with a fit of the proxy models (Supplementary Table S4A). We also merged the English with Germans to see how well the merged population performs as a source and found the results to be closer to (though not as good as) those obtained using English as source than to those obtained using Germans as source (Fig. 3A, Supplementary Table S4A). However, the standard errors were smaller in the merged population than with either of the separate source populations.

Next, we simultaneously replaced both original source populations by a proxy (Fig. 3C, E, Supplementary Table S4A). Overall, the results did not vary much, suggesting that LINADMIX is robust with respect to genetic drift, at least in this setup. The worst estimates of the mixing coefficients were obtained when Palestinians were used as a proxy to Jordanians, and French were used as a proxy to English, resulting in an average estimation error of 8.2% points.

We repeated the above tests for the scenario that the target is an admixture of closely related populations. We used the Jordanian-Iranian simulations for this purpose, taking again Syrians and

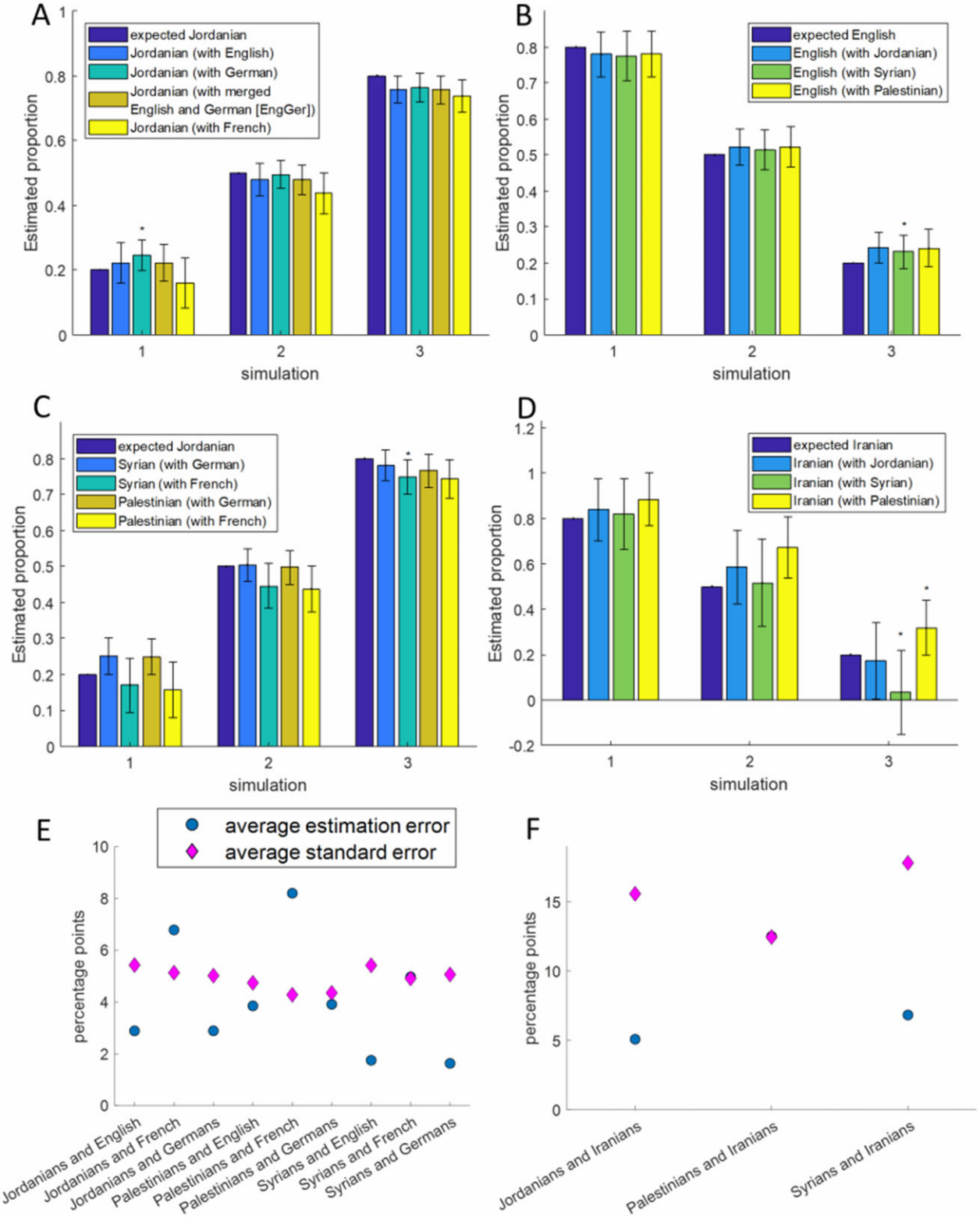


Fig. 3. The effect of genetic drift on LINADMIX. Target populations were generated from an admixture of Jordanians and English (A,B,C,E) or Jordanians and Iranians (D,F). (A) The estimated Jordanian contribution when a proxy source population, or a merged source population of English with the proxy population German ('EngGer'), is given as input replacing the original English population (the results with English are given for comparison.) (B) The estimated English contribution when a proxy source population is given as input replacing the original Jordanian population (the results with Jordanians are given for comparison.) (C) LINADMIX estimations when both original source populations are replaced by proxy populations. (D) The estimated Iranian contribution when a proxy source population is given as input replacing the original Jordanian population (the results with Jordanians are given for comparison). Asterisks above bars in A-D represent that the p-value of the model was below 0.05.(E,F) LINADMIX performance, measured in several ways, for different proxy source populations. Each source population consisted of four individuals.

Palestinians as proxies for Jordanians (Fig. 3D, F, Supplementary Table S4B). Here, replacing Jordanians with Syrians had a small effect, increasing by ∼2% points both the average estimation error and the average standard error. Replacing the Jordanians with the less related Palestinians had a larger effect on the estimation, increasing the average error by ∼7.5% points. The only models associated with *P*-values lower than 0.05 were those with a mixture of 80% Jordanians and 20% Iranians. Thus, using a proxy source is also possible when the target population is a mixture of closely related populations, although a less diverged proxy is more recommended.

Finally, we tested how genetically remote a population could be in order to be still an adequate proxy source population. We found that populations with up to $F_{ST} = 0.003$ are good proxies for sources when the admixed populations are not closely related, and that this number reduces to $F_{ST} = 0.002$ when the admixed populations are closely related (Supplementary Text S7).

In conclusion, we found LINADMIX to perform robustly with respect to both the number of generations since the admixture event and to drift in allele frequencies between the true and modeled source populations.

3.5 The effect of the number of individuals in the source populations

Sample size is an important parameter that could potentially affect LINADMIX. To look at the effect of varying sample sizes on LINADMIX we chose two source populations with relatively large sample sizes—25 for French and 38 for Palestinians. We mixed French and Palestinians in proportions of 0.2:0.8, 0.5:0.5 and 0.8:0.2 as described above, and ran LINADMIX with increasing sample sizes of the French and Palestinian sources: For each French-Palestinian target we performed three sets of LINADMIX runs, each with sample sizes (of both French and Palestinians) of 2, 4, 8, 12, 16 and 20. The individuals for each set were chosen randomly (but all individuals in a smaller sample were contained in larger ones).

Generally, as expected, larger sample sizes led to more accurate results (Fig. 4, Supplementary Table S5). Nonetheless, in terms of estimation, even with only 2 individuals in the source populations, the average estimation error did not exceed 6% points, only 3% points higher than with the largest sample size of 20 individuals. However, the main hinderance of the smaller sample sizes, in particular that of 2, was the greater tendency to reject the models: for the smallest sample size of 2, three out of the nine models (three models for each of the three sets) were rejected, and for the second smallest sample size (4) one model out of the nine was rejected. For all other sample sizes, all *P*-values were higher than 0.05. The standard errors ranged between 5.3% points to 9.1% points with no distinct correlation with the sample size (Supplementary Table S5).

In conclusion, whereas estimation is robust for any source sample size, the use of small sample sizes (2 individuals, and to a lesser extent 4 individuals) increases the risk of rejecting true models.

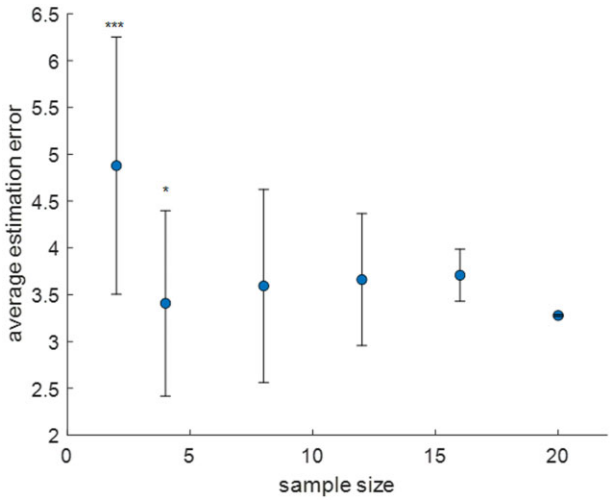


Fig. 4. The effect of sample size on LINADMIX. French and Palestinians were mixed in three different proportions. For each sample size of the sources, LINADMIX was run three times (with different individuals in the sources) on all target populations. The graph shows the average estimation error across the three targets and three sets of runs versus the number of individuals in the source populations. The standard deviations of the estimation errors across the three sets of runs are indicated. The number of asterisks above points in the graph represents the number of models where the *P*-value was lower than 0.05

3.6 Misspecification of the model

An admixture model provided to LINADMIX may be misspecified in two main ways. First, a population that did contribute to the target population may be omitted from the list of source populations. Second, a population that did not contribute to the target population may be included.

To account for the first scenario, we replaced the original source population with another population that is genetically remote. In this case, we expect the *P*-values reported by LINADMIX to fall below the 0.05 threshold, indicating that the model does not fit the data. To account for the second scenario, we included a third, unrelated, source population in addition to the original two sources populations. In this setting, we expect the estimator of the mixing coefficient of the redundant source population not to be significantly different from zero.

3.6.1 The effect of omitting a true source population

We first considered the Jordanian-English admixture model. To test the effect of omitting a true source population, we replaced one of the true source populations by a population that is genetically remote. English were replaced by either Somalis, Syrians or Palestinians, whereas Jordanians were replaced by either Somalis, Germans or French (Supplementary Table S1). As expected, all six models were shown to be incompatible with the data (*P*-values lower than 0.05; Table 1). Thus, LINADMIX has the desired property that the replacement of a true source population by a close proxy provides robust estimation, while the use of a remote population yields an invalid model.

Next, to further explore the fact that closeness between population is a continuum, we tried to replace the source populations with another population that—while not a good proxy—is still not too diverged. We performed the same analysis as above, but replaced Jordanians with an ancient population that is genetically not very remote—Iranian Chalcolithic (Iran_ChL) (Agranat-Tamir *et al.*, 2020). The average proportion of missing SNPs for Iran_ChL is 0.43. In addition, we replaced English with Russians, which are relatively closely related ($F_{ST} = 0.005$, Supplementary Table S1). In all cases but one, LINADMIX detected that the model is incompatible with the data (Table 1). The single exception was where English were replaced by Russians and the target population had the lowest proportion of English (20%, *P*-value 0.34).

3.6.2 Effect of a redundant source

We first considered the Jordanian-English admixture model. We used the two original source populations, and added either Somalis, Iran_ChL, or Russians (Fig. 5, Supplementary Table S6A). LINADMIX correctly identified that the Somali population did not contribute to the target population. The highest estimated Somali mixing coefficient was 1.1% (with 2% standard error). The estimated contributions of both English and Jordanians were similar to those obtained when the Somali samples were not included as a source population (Fig. 5A–C). When the less genetically remote Iran_ChL or the related Russians were used as source populations, LINADMIX was still largely able to detect the true sources and gave only small weights to these populations. In all cases, the estimated mixing coefficients of these redundant sources did not deviate significantly from zero. In addition, the mixing coefficients estimated for the true source populations were robust and similar to those obtained in the analysis that did not include a redundant source population (Fig. 5A–C).

Table 1. Invalid models

Sources in the model	F_{ST} of alternative and original populations	Average error (percentage points)	Maximum error (percentage points)	Average standard error (percentage points)	*P*-value (target 1)	*P*-value (target 2)	*P*-value (target 3)
Jordanian with Somali	0.075	49.66	80.00	2.35	0.0001	0.0001	0.0001
Jordanian with Syrian	0.014	42.24	56.72	40.38	0.0001	0.0001	0.0001
Jordanian with Palestinian	0.016	35.12	50.00	43.12	0.0001	0.0001	0.0001
Jordanian with Russian	0.005	7.40	15.26	3.79	0.0001	0.0028	0.3409
English with Somali	0.049	31.94	50.91	2.54	0.0002	0.0001	0.0001
English with German	0.013	50.00	80.00	44.09	0.0032	0.0001	0.0001
English with French	0.011	50.00	80.00	39.67	0.0072	0.0001	0.0001
English with Iran_ChL	0.013	7.74	14.53	5.58	0.0400	0.0002	0.0001

Note: The target population is a mixture of Jordanian and English, different invalid models. *P*-values under 0.05 are shaded.
Target 1: Jordanian 20% English 80%
Target 2: Jordanian 50% English 50%
Target 3: Jordanian 80% English 20%

Next, we wished to test the scenario that a population that is genetically similar to a true source is added as an additional source. It is important to understand how LINADMIX treats such redundant sources: does it allocate the entire mixing coefficient to either of the two similar source populations, or does it split the mixing coefficients between them? To this end, we took targets produced by the Jordanian-English model, and added to the Jordanian and English sources either Palestinians, Syrians, Germans or French as a third source. In all cases, we used an 'ancient' version of the redundant sources, with a proportion of 0.6 missing data.

With two exceptions, LINADMIX estimated the mixing coefficient of the redundant population as no greater than 3.72%, with an average of 0.66% (Fig. 5D–F, Supplementary Table S6B). The standard errors of the redundant population increased when the true mixing coefficient of the related source increased, and were very similar to those of the related source (Fig. 5D–F, Supplementary Table S6B). The average of the standard errors of the mixing coefficients of the third population in these cases was much lower (5.2%). In two exceptional cases, the redundant population either replaced the true source or, in combination with the true source, gave a reasonable estimate (a combined estimate of 0.77 from both related sources versus 0.8 coefficient of the actual source). Importantly, as would be expected based on the fact that the true sources are in the set of sources given to LINADMIX, all *P*-values are above the 0.05 threshold.

We tested whether LINADMIX treats redundant sources the same way when the target is either a mixture of closely related populations or of particularly remote populations. As representatives of closely related populations, we used the Jordanian-Iranian simulations (Fig. 5G–I, Supplementary Table S6C). For remote populations, we performed simulations mixing Somalis and English (F_{ST} = 0.075, details in Supplementary Text S8). Generally, the results are similar to those with the Jordanian-English targets. Nonetheless, when the true sources are distant, LINADMIX is slightly more prone to give the redundant source some contribution (Fig. 5J–L, Supplementary Table S6D).

We conclude that when two related populations are provided as source populations, the mixing coefficient estimates provided by LINADMIX are characterized by higher standard errors. Nonetheless, in most cases LINADMIX assigns a near zero ancestry to the redundant source. This robustness is particularly advantageous when using ancient source populations, as sparse sampling makes it more likely that related ancient populations are used as sources.

3.6.3 Different proportions of missing data in the redundant source populations

We next wished to test whether LINADMIX is biased when the redundant sources have lower proportions of missing data. We repeated the previous analysis on the Jordanian-English targets, while replacing the 'ancient' Syrians, Palestinians, Germans and French with the original modern genomes (Fig. 6). English and Jordanians remained 'ancient' with 60% missing SNPs. LINADMIX's results in this case were very similar to those when the redundant input sources were 'ancient'. In nine out of twelve models tested, the inferred contribution of the redundant source was lower, or no more than 2.2% points higher, compared to the 'ancient' version of the redundant source (Table 2). In the remaining three cases, LINADMIX increased the contribution of the redundant sources by 5.6% to 7.4% points (Table 2). Considering the average across the three Jordanian-English targets, the differences between the modern sources estimates to the 'ancient' sources estimates range from 0.44% points (Syrian) to 2.61% points (Palestinian) (Fig. 6D). Thus, varying proportions of missing data in the sources do not strongly affect LINADMIX (Fig. 6, Table 2).

3.7 The effect of a source with low contribution

Low contribution of sources to a target (small mixing coefficients) are expected to be harder to infer both in terms of estimation and in terms of model fit. To test how LINADMIX performs in such cases we simulated the admixture of Jordanians with either English or Iranians in mixing proportions of 0.9:0.1 and 0.95:0.05 [Fig. 7 (0.8:0.2 mixing proportions are shown for comparison) Table 3 and Supplementary Table S7]. In the Jordanian-English simulations, the estimation was not substantially affected—the absolute estimation error increased at most by 1.2 percentage points, the standard errors of the estimates were not affected, and the *P*-values were in accordance with the model fit. In the Jordanian-Iranian simulations, while the absolute estimation errors were up to 5% points, the Iranian contribution was not recognized when it was 5%. In addition, because the standard errors of the estimates in the Jordanian-Iranian simulations are about 15%, it would be difficult to determine (without prior knowledge) whether low estimated contributions of Iranians do indeed reflect true Iranian contributions.

A possible way to test whether a population that is given a small contribution by LINADMIX is indeed a valid source is to look at the *P*-value of a model that omits this population from the sources. We performed this and found that in the Jordanian-Iranian simulations, when LINADMIX is given a model where the only source is the Jordanian population, it gives *P*-values higher than 0.05 when the Iranian contribution is 10% or less, indicating that Jordanians are sufficient to explain the target (Table 3). In contrast, in the Jordanian-English simulations, LINADMIX gives *P*-values lower than 0.05 when Jordanians are the only source provided, even with 5% English contribution to the target. Therefore, when the target is composed of relatively distant populations, LINADMIX is able to detect very low contributions of source populations. However, when the target is composed of two closely related populations, lower contributions are harder to detect.

3.8 Admixture of more than two populations

We next wished to examine how LINADMIX performs when the target population is an admixture of more than two populations. We started with three source populations. To this end, we simulated an admixture of English, Jordanians and Iranians, thus combining both distant and closely related populations. Three different mixing

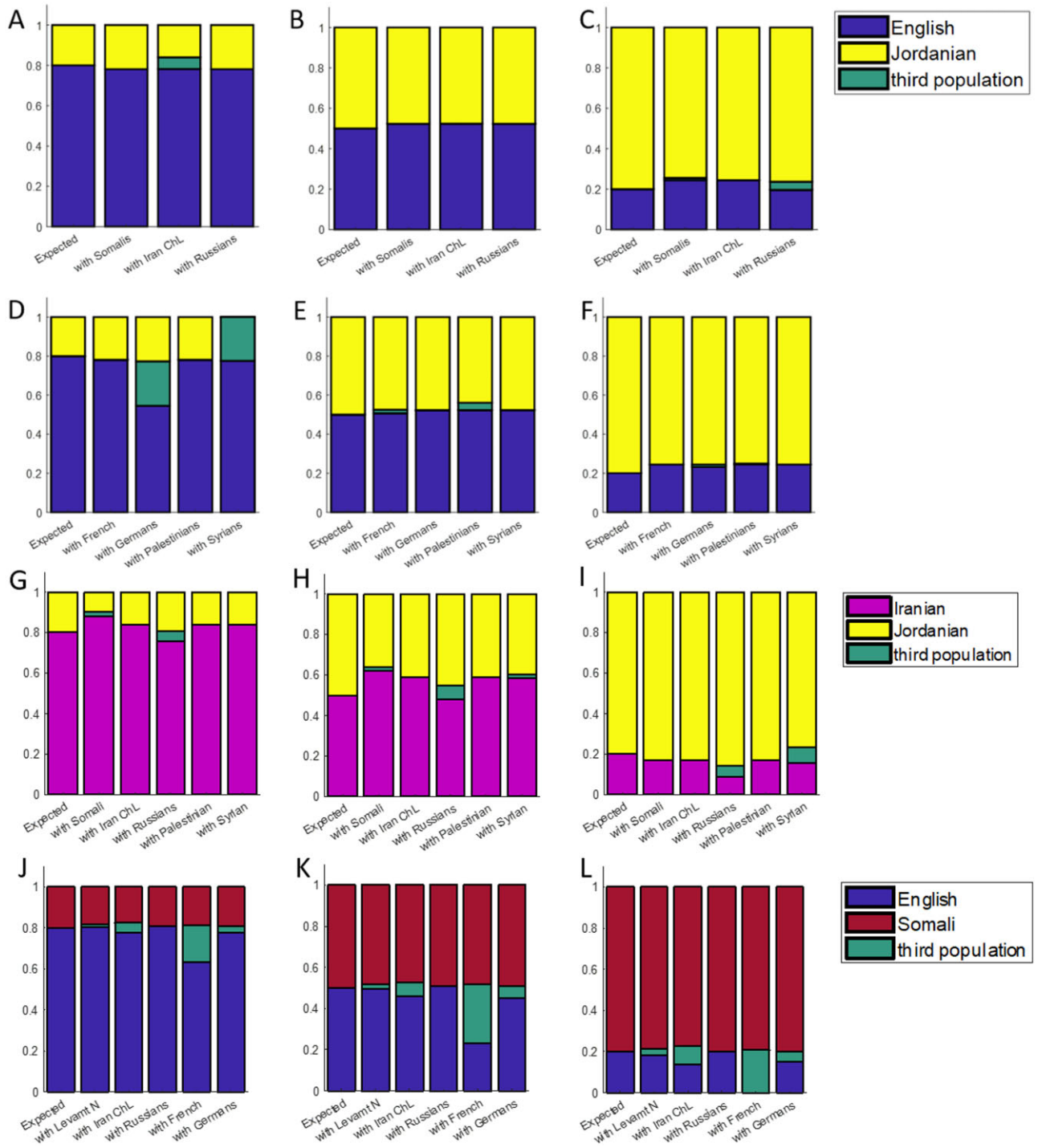


Fig. 5. LINADMIX with three source populations for targets that are mixtures of two populations in different proportions. The third population was either artificial ancient with a proportion of 0.6 missing data or true ancient (Iran_ChL and Levant_N). (**A–F**) The targets are mixtures of Jordanians and English. (A, D) 20% Jordanians in the target; (B, E) 50% Jordanians; and (C, F) 80% Jordanians. (**G–I**) The targets are mixtures of Jordanians and Iranians. (G) 20% Jordanians; (H) 50% Jordanians; and (I) 80% Jordanians. (**J–L**) The targets are mixtures of Somalis and English. (J) 20% Somali; (K) 50% Somali; (L) 80% Somali

proportions for Jordanian-Iranian-English triplets were considered: 0.1:0.6:0.3, 0.3:0.1:0.6 and 0.6:0.3:0.1, and the simulations were repeated five times.

All the computed *P*-values exceeded 0.05, indicating that in all cases the model was compatible with the data. The highest estimation error was 9.9% points and the average of the estimation errors was 4.1% points (Fig. 8A–C, E, Supplementary Table S8). Specifically, the average estimation errors of the two closely related populations (Jordanians and Iranians) was about 5%, whereas that of the more distant English source was only 2.4%. A similar pattern was observed for the average standard errors of the mixing coefficients. The overall average was 7.7%, but a higher average of 8.9% was obtained for the more closely related populations, whereas a lower average of 5% was observed for the more distant population.

We then wished to examine a scenario of four source populations in equal proportions, adding Russians to the Jordanians, Iranians and English, thus having two pairs of closely related populations. The estimation accuracy is not very different from that with three

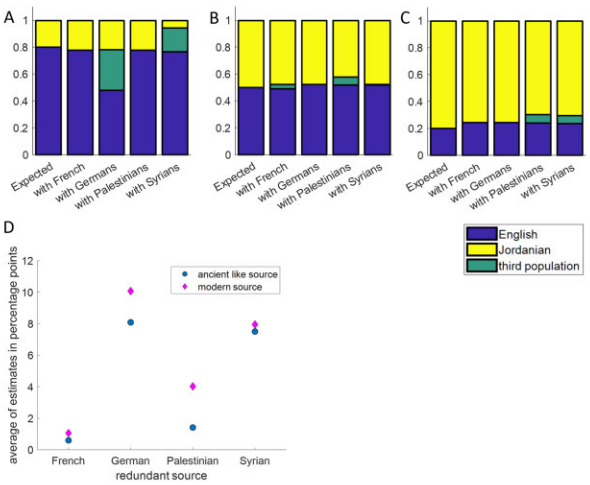


Fig. 6. LINADMIX with three source populations for targets that are a mixture of Jordanians and English in different proportions. (**A–C**) The estimates by LINADMIX when the third (redundant) source population is modern. (**D**) The average of the estimates of the redundant sources across the three Jordanian-English simulations when it is either "ancient" with 60% missing data (blue discs) or modern (pink diamonds)

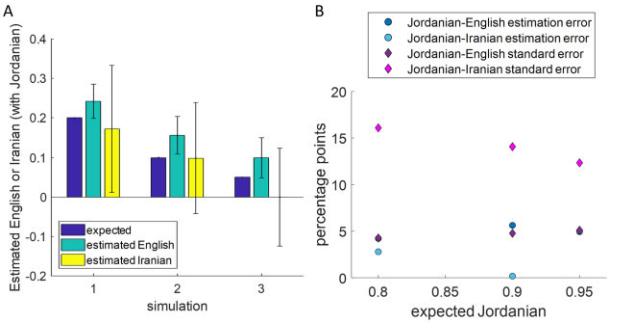


Fig. 7. Sources with low contribution. (**A**) The estimated contribution of either English (turquoise) or Iranian (yellow) in three simulations where Jordanians were admixed with either English (turquoise) or Iranians (yellow). (**B**) The absolute estimation errors (discs) and standard errors (diamonds) where Jordanians were admixed with either English (dark colors) or Iranians (light colors)

source populations, albeit with a slightly elevated maximum error (10.81 versus 9.92 percentage points), and with somewhat elevated standard errors (Fig. 8D, E, Supplementary Table S8).

We next tested whether, in these more complex admixtures, LINADMIX can detect that a source population is missing from the set of sources in the model. To this end, we looked at the *P*-values computed by LINADMIX when one of the sources was missing from the model (Table 4). Similarly to what we observed for two population admixtures, LINADMIX returns *P*-values above 0.05 when one of the two closely related source populations (Jordanians or Iranians) had only 10% contribution to the target population (Table 4). A *P*-value of 0.051 was observed when the English were missing from the sources and the target contained 10% English (Table 4). The only case where LINADMIX failed at detecting a missing source population when the contribution to the target was higher than 10% was when Iranians were missing as a source for a target to which they contributed 30%. For the four-way admixture Jordanian-Iranian-English-Russian, LINADMIX could not recognize a missing source population when it has a closely related source population present in the set of sources in the model (Supplementary Table S9). When both populations of a closely related pair were missing from the model, LINADMIX returned *P*-values below 0.05, as expected.

In summary, the estimation accuracy in a setting that involves three source populations, two that are closely related in addition to a more distant third population, or in a setting that involves two pairs of closely related populations, is comparable to the accuracy achieved for an admixture of only two populations. The rise in complexity compromises LINADMIX's ability to detect a missing source from the model, even though it is still able to detect a missing source when it is genetically distant from other sources and contributes more than 10% to the target population. Therefore, LINADMIX shows robustness in estimation to an increase in the complexity of the admixture model.

3.9 Nested models

Finally, we tested LINADMIX when a series of source populations was added to the model on top of the original source populations (or their proxies). To this end, we used a target population that is a three-way admixture of Jordanians, Iranians and English. We provided LINADMIX with a series of nested models, adding in sequence gradually more genetically distant source populations. In total we generated 11 models for each Jordanian-Iranian-English target ranging from the original three source populations and up to a total of ten source populations. We found (Fig. 9, Supplementary Fig. S1, Supplementary Table S10) that as long as the combined contribution of a source and its proxies is considered, LINADMIX maintains high accuracy of estimation. The estimates of the mixing coefficients of non-relevant or closely related but not proxy populations Somalis and Russians average on 1.40% points with a maximum of 4.80% points. Adding Spanish, who are close to the English ($F_{ST} = 0.003$) but not as close as the Germans and French ($F_{ST} = 0.001$), the average slightly rises to 1.56%, however the maximum is 12.36%. The median in both cases is 0.83%.

Notably, in the vast majority of tests, the standard errors of the non-relevant populations were at least half of the estimate, indicating that the mixing coefficient is not significantly different from zero. In addition, when closely related populations are

Table 2. Estimation with full and missing data

Redundant source provided to LINADMIX	Difference between modern estimate and 'ancient' estimate		
	Target 1	Target 2	Target 3
French	0.0001	0.0137	0.0000
German	0.0736	−0.0020	−0.0121
Palestinian	0.0000	0.0220	0.0563
Syrian	−0.0456	0.0032	0.0558

Note: The difference in the estimation of the mixing coefficient of a redundant source with full data (modern) and a redundant source with 0.6 missing positions and in haploid form ('ancient').

Target 1: Jordanian 20% English 80%

Target 2: Jordanian 50% English 50%

Target 3: Jordanian 80% English 20%

Table 3. Model fit with a missing source population

Populations in the target	Sources in the model	*P* values with expected Jordanian		
		0.8	0.9	0.95
Jordanian and English	Jordanian and English	0.6017	0.4435	0.3029
	Jordanian	0.0001	0.0001	0.0002
Jordanian and Iranian	Jordanian and Iranian	0.0693	0.2369	0.0548
	Jordanian	0.0176	0.2199	0.0558

Note: *P*-values below 0.05 are shaded.

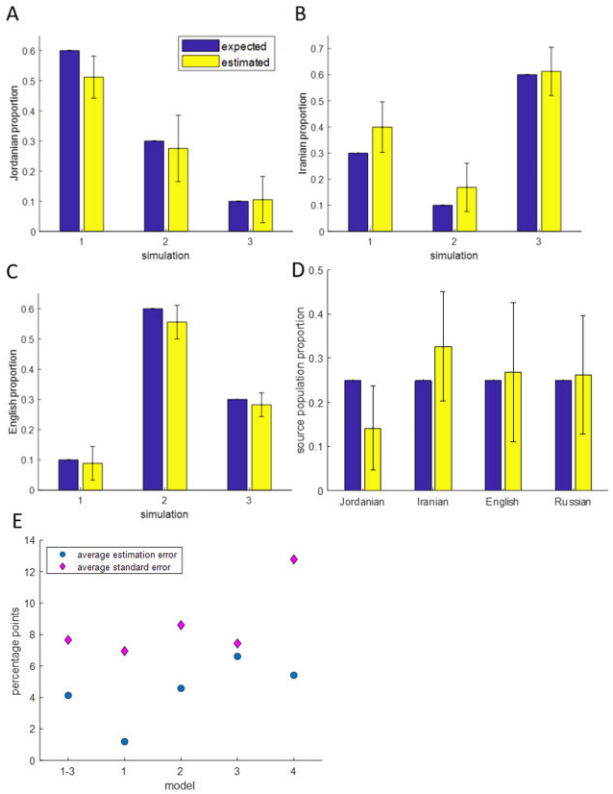


Fig. 8. Admixture of three or four populations. The Jordanian (**A**) Iranian (**B**) and English (**C**) expected and estimated contributions in three simulations of Jordanian, Iranian and English admixture. (**D**) The expected and estimated contributions of Jordanians, English, Iranians and Russians in a simulation of their admixture. (**E**) The average (absolute) estimation errors and average standard errors across all sources in the indicated models. Model 1—10% Jordanian, 60% Iranians, 30% English. Model 2—30% Jordanian, 10% Iranian, 60% English. Model 3—60% Jordanian, 30% Iranian, 10% English. Model 4—25% Jordanian, 25% Iranian, 25% English, 25% Russian. 1–3 stands for: across models 1, 2 and 3

simultaneously given as sources their standard errors occasionally rise. Therefore, it is advisable not to provide LINADMIX with closely related source populations but rather to try and either merge them or to choose the one that is most fitting.

3.10 ADMIXTURE perturbations

Because LINADMIX relies on the output of ADMIXTURE it is important to understand how perturbations in the ADMIXTURE run affect LINADMIX. We discussed this in detail in Agranat-Tamir *et al.*, (2020) without referring to the effect on the calculation of empirical *P*-values as this calculation is introduced here. We now complete the discussion about ADMIXTURE perturbations referring to the effect on the calculation of empirical *P*-values. Full details are given in Supplementary Text S9. The main conclusions are that using a sub-optimal number K of ADMIXTURE reference populations is not advisable. In addition, it is recommended to have a set of background populations that (while thorough for the study) does not have a very high optimal K.

4 Discussion

The main ideas behind LINADMIX have been briefly described in our previous work (Agranat-Tamir *et al.*, 2020). Here, we provide a comprehensive set of simulations testing the performance of LINADMIX under several settings and parameter values. Moreover, we introduce a technique to compute an empirical *P*-value that measures how well the estimated model fits the data.

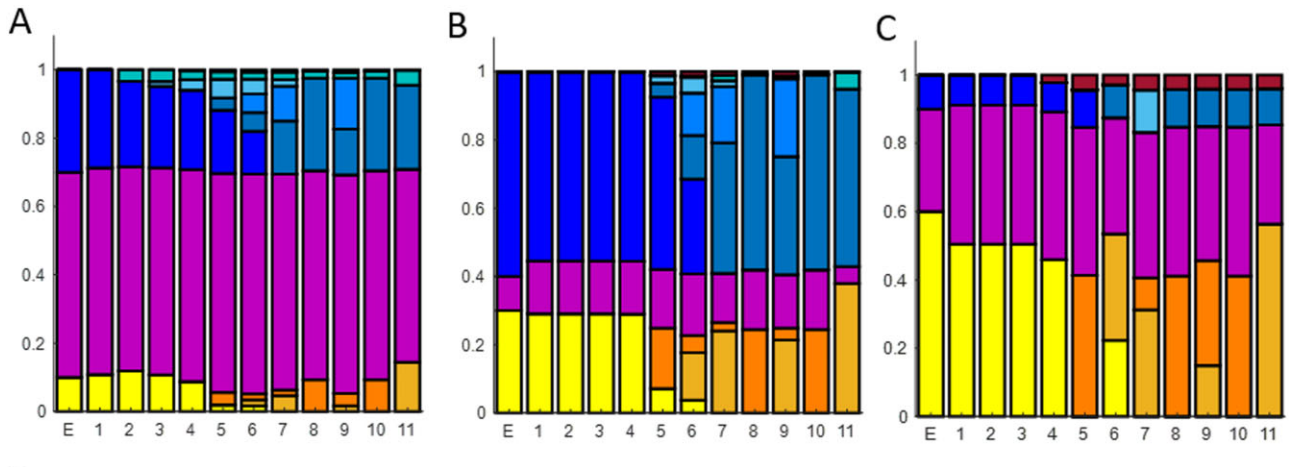
We showed that LINADMIX can accurately estimate mixing coefficients and confirm model validity with up to 60% missing data. Higher proportions of missing data are characterized by an increased estimation error, although LINADMIX keeps performing reasonably well even up to 80% missing data (with a maximum error of 8% points in the simulations that were considered). The main effect of such high proportions of missing data is the failure of LINADMIX (in some cases) to distinguish between valid and invalid models. However, even for proportions of missing data as high as 80%, LINADMIX may be used to estimate mixing coefficients, as long as there is an independent indication for the validity of a model. Even at 90% missing data, LINADMIX was able to provide rough estimates of the mixing coefficients (maximum estimation error of 14% points, averaging at 7.3% points). Importantly, LINADMIX has remained roughly unbiased even for samples with either low or high proportions of missing data. Therefore, there seems to be no need for down-sampling of data to equalize the proportions of missing data across the source populations.

As expected, source populations with a large number of individuals yield better results. We realize that in practice, large samples from ancient populations are rare, and researchers often proceed with whatever data is available. On the other hand, one should be mindful that the larger the available number of individuals in a

Table 4. models with a missing source

Source populations provided to LINADMIX	*P*-values expected Jordanian:Iranian:English		
	0.1:0.6:0.3	0.3:0.1:0.6	0.6:0.3:0.1
Iranian and English	0.3030	0.0080	0.0006
Jordanian and English	0.0001	0.3750	0.1272
Jordanian and Iranian	0.0001	0.0001	0.0514

Note: *P*-values below 0.05 are shaded.



D

	1	2	3	4	5	6	7	8	9	10	11
Jordanian	+	+	+	+	+	+	-	-	-	-	-
Syrian	-	-	-	-	-	+	+	-	+	-	+
Palestinian	-	-	-	-	+	+	+	+	+	+	-
Iranian	+	+	+	+	+	+	+	+	+	+	+
English	+	+	+	+	+	+	-	-	-	-	-
German	-	-	-	-	+	+	+	+	+	+	+
French	-	-	-	-	-	+	+	-	+	-	-
Spanish	-	-	+	+	+	+	+	+	-	-	-
Russian	-	+	+	+	+	+	+	+	+	+	+
Somali	-	-	-	+	+	+	+	+	+	+	+

Fig. 9. The effect of nested models. (**A–C**) The expected (*E*) and estimated contribution of the indicated source populations to targets composed of different mixing proportions of Jordanians, Iranians and English in 11 models consisting of nested lists of sources as indicated in (**D**)

sample, the more likely it is to capture the variability in the entire population.

One of the main criticisms of ADMIXTURE is that is does not account for genetic drift that occurred after the admixture event (Lawson *et al.*, 2018). A possible explanation for this property, as well as the robustness to missingness of data, is the fact that the proportions of the hypothetical ancestral populations are computed as averages across many sites. Genetic drift is modeled to be random and is thus canceled out by averaging. In our context this property is actually very useful because it guarantees the relative robustness of our procedure to using in the model source populations that are closely related, but not identical, to the actual source populations that produced the target population.

Here, we show that LINADMIX is indeed very robust to drift, as we showed that it can successfully recognize drifted populations as the sources and estimate their mixing coefficients accurately. This is particularly important for ancient DNA studies, as sampling is sparse, and we cannot expect to sample the precise source populations. Given the continuous nature of the genetic relatedness of populations, LINADMIX demonstrates desired properties. Source populations that are closely related to the true source population, yet are still remote enough not to behave as a drifted source population (Iran_ChL with respect to Jordanians for example), were not recognized as adequate sources, and the model is rejected with a low *P*-value. Even closer populations, such as Russians to English, are only recognized as alternatives when the true source's proportion in the target is low.

Besides drift, genetic variability is shaped over time by recombination. We showed that, as expected, LINADMIX is insensitive to recombination as it is ultimately based on allele frequencies, and not on haplotype information. We did not observe any notable change in LINADMIX's performance between 20 and 160 generations since admixture.

LINADMIX can distinguish the contributions of genetically similar populations almost as well as for distantly related ones. However, when the sources are genetically similar, the standard errors of the mixing coefficients are higher. Thus, it is not advisable to use LINADMIX to distinguish between two very close populations. Yet, when both the true source and a proxy were modeled as potential sources, LINADMIX tended to prefer the true source over its proxy.

An important issue is the ability to recognize sources with low contributions to the target population. We suggest two tests for the significance of an estimate. First, to look at a contribution as positive (rather than zero) if it is higher than a certain number of (say, two) standard errors. Second, to look at model validity without the source (or proxies to that source). We found that when the target is composed of remote sources, it is possible to detect contributions as low as 10%. One should be mindful that if the target is composed of related sources, the high standard errors would make it difficult to detect small contributions. This is also true for complex admixture models that involve related sources.

Acknowledgements

The authors thank the Israel Science Foundation for funding. L.C. is the Snyder Granadar chair in Genetics.

Funding

This work was supported by the Israel Science Foundation [ISF grant 1009/17 to L.C. and B.Y.; ISF grant 407/17 to S.C.].

Conflict of Interest: none declared.

References

Agranat-Tamir,L. *et al.* (2020) The genomic history of the bronze age southern levant. *Cell*, **181**, 1146–1157.e11.

Alexander,D.H. *et al.* (2009) Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.*, **19**, 1655–1664.

Bansal,V. and Libiger,O. (2015) Fast individual ancestry inference from DNA sequence data leveraging allele frequencies for multiple populations. *BMC Bioinformatics*, **16**, 4.

Battey,C. *et al.* (2020) Predicting geographic location from genetic variation with deep neural networks. *Elife*, **9**, 1–22.

Elhaik,E. *et al.* (2014) Geographic population structure analysis of worldwide human populations infers their biogeographical origins. *Nat. Commun.*, **5**, 3513.

Gaspar,H.A. and Breen,G. (2019) Probabilistic ancestry maps: a method to assess and visualize population substructures in genetics. *BMC Bioinformatics*, **20**, 13.

Haak,W. *et al.* (2015) Massive migration from the steppe was a source for Indo-European languages in Europe. *Nature*, **522**, 207–211.

Harney,É. *et al.* (2021) Assessing the performance of qpAdm: a statistical tool for studying population admixture. *Genetics*, **217, iyaa045**.

Jin,Y. *et al.* (2019) GRAF-pop: a fast distance-based method to infer subject ancestry from multiple genotype datasets without principal components analysis. *G3 Genes Genomes Genet.*, **9**, 2447–2461.

Joseph,T.A. and Pe'er,I. (2019) Inference of population structure from time-series genotype data. *Am. J. Hum. Genet.*, **105**, 317–333.

Kozlov,K. *et al.* (2015) Differential Evolution approach to detect recent admixture. *BMC Genomics*, **16**, S9.

Lawson,D.J. *et al.* (2018) A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nat. Commun.*, **9**, 3258.

Lawson,D.J. *et al.* (2012) Inference of population structure using dense haplotype data. *PLoS Genet.*, **8**, e1002453.

Leslie,S. *et al.*; International Multiple Sclerosis Genetics Consortium. (2015) The fine-scale genetic structure of the British population. *Nature*, **519**, 309–314.

Noto,K. *et al.* (2020) Ancestry inference using reference labeled clusters of haplotypes. *bioRxiv*, 2020.09.23.310698.

Patterson,N. *et al.* (2012) Ancient admixture in human history. *Genetics*, **192**, 1065–1093.

Patterson,N. *et al.* (2006) Population structure and eigenanalysis. *PLoS Genet.*, **2**, e190.

Taravella Oill,A.M. *et al.* (2021) PopInf: an approach for reproducibly visualizing and assigning population affiliation in genomic samples of uncertain origin. J. Comput. Biol., **28**, 296–303. 10.1089/cmb.2019.0434 33074720.