Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Genetics & Development

Methods for detecting introgressed archaic sequences

Sriram Sankararaman[1,2,3]

Analysis of genome sequences from archaic and modern humans have revealed multiple episodes of admixture between highly-diverged population groups. Statistical methods that attempt to localize DNA segments introduced by these events offer a powerful tool to investigate recent human evolution. We review recent advances in methods for detecting introgressed sequences.

Addresses
[1] Department of Computer Science, University of California, Los Angeles, CA 90095, United States
[2] Department of Human Genetics, David Geffen School of Medicine, University of California, Los Angeles, CA 90095, United States
[3] Department of Computational Medicine, University of California, Los Angeles, CA 90095, United States

Corresponding author: Sankararaman, Sriram (sriram@cs.ucla.edu)

Current Opinion in Genetics & Development 2020, **62**:85–90

This review comes from a themed issue on **Genetics of human origin**

Edited by **Sarah Tishkoff** and **Joshua Akey**

https://doi.org/10.1016/j.gde.2020.05.026

Introduction

The sequencing of archaic human genomes [1–3] coupled with the availability of genome sequences from diverse present-day human populations [4–10] have revealed multiple episodes of gene flow between archaic and modern humans (see [11] for a review). Beyond genome-wide proportions, a number of studies have attempted to characterize how introgressed archaic DNA is distributed along the genome [1,12–16,17•,18••,19•,20,21]. Analysis of maps of archaic introgression have yielded novel insights into human evolution and biology [12,13,15,16,22,23,17•,24–31].

A key analytical problem in these studies is the identification of introgressed archaic DNA segments in the genomes of present-day humans, that is, segments of DNA segregating in the genomes of present-day humans that trace their ancestry to archaic hominins (Figure 1). Methods to identify introgressed segments rely on diverse population genetic statistics that rely on allele frequency configurations across populations, sequence divergence, and linkage disequilibrium (LD). Efforts to combine multiple statistics have typically relied on statistical models such as hidden Markov models (HMMs) and conditional random fields (CRF) that account for correlation in ancestry across SNPs [24].

Proposed methods differ in the types of information that they rely on. HMM and CRF-based methods for detecting Neanderthal segments in a target population (e.g. non-Africans) relied on the Neanderthal genome as an archaic reference and west African genomes as a modern human reference assumed to be unadmixed [1,12,14] (Figure 1). While leveraging the availability of unadmixed archaic and modern human reference genomes can lead to increased sensitivity to detect introgressed segments, the inferences from these methods can be biased when their assumptions are violated. Methods such as the S*-statistic [13] that do not require access to an archaic reference can be applied to detect introgressed segments from as-yet-undiscovered archaic populations.

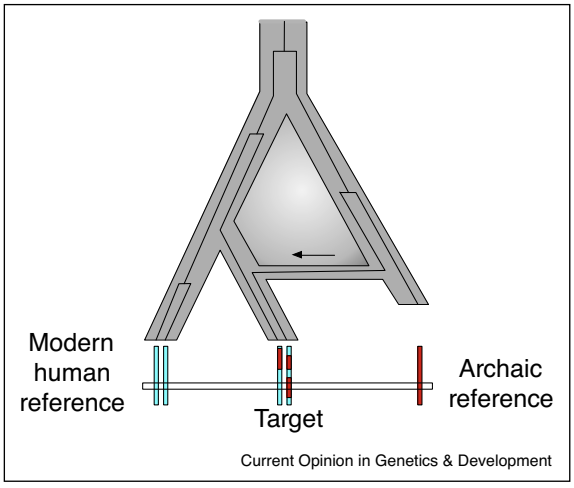
Recent developments

S* and Sprime

Introgression from a highly-diverged population introduces novel mutations into the hybrid population that remain tightly linked if the introgression is not too ancient. The S* approach [32] searches for extended sequences of tightly linked SNPs that are likely to represent introgressed segments. Variants of this approach differ in the size of the genomic region, set of SNPs, and number of target individuals analyzed, and the score function optimized [32,13,16]. The versions of Vernot and Akey 2014 and Vernot *et al.* 2016 used a sliding window appropriate for the time scale of Neanderthal and Denisovan introgression [16]. These approaches use an unadmixed modern human reference population to filter out SNPs unlikely to be introgressed, specifically removing SNPs at which Yoruba genomes carry the derived allele. Importantly, this approach does not require an archaic reference genome allowing the identification of ghost introgression, that is, introgression from an unknown archaic group. In applications of S* to Neanderthal or Denisovan admixture, the segments are matched to a reference archaic genome to decrease the false discovery rate.

A recent version of the S* approach (Sprime) [18••] proposes improvements that include not using a fixed genomic window, allowing for low levels of archaic introgression in the modern human reference population, and a scoring function that accounts for the frequency of the introgressed haplotype, variation in mutation and recombination rates. Sprime identifies an introgressed segment, removes the putative introgressed SNPs associated with the segment, and proceeds to identify an introgressed segment among the remaining SNPs.

Figure 1



Inference of archaic introgressed segments in a target population: Introgression leads to the genome of the target population having a mosaic of archaic (red) and non-archaic (blue) ancestries. In the typical setting for identifying introgressed segments in a target population, genomes are assumed to be available from the archaic population (archaic reference) as well as a modern human population that does not share the introgression event (modern human reference). For example, studies of Neanderthal introgression would study a non-African population as target assuming that African populations represent a modern human reference. The lengths of the introgressed segments depend on the introgression fraction as well as the time since introgression. The genealogy at one of the genomic loci shows one of the key signals used to identify introgressed segments, that is, introgressed segments in the target population tend to be closer to the archaic genome relative to the modern human genome.

A related approach to identify ghost introgressed haplotypes searches for regions with ten or more SNPs in complete LD (pairwise $r^2 = 1$) with derived allele frequency below 0.1 in the Denisovan and Neanderthal genomes in addition to length and distance cutoffs [20].

Probabilistic models

Generative models

Correlation of ancestry across SNPs have motivated the use of HMMs [33] to detect introgressed segments. The hidden state in HMMs represents the ancestry of an individual at a genomic location while observed states model the data. A version of the HMM to infer archaic segments [14,34] modeled the configuration of derived allele frequency in the test haplotype, archaic and reference populations given the ancestry state is archaic or non-archaic.

diCal-admix [17•] is a HMM-based method that infers introgressed segments while explicitly accounting for demographic history. To detect Neanderthal introgression in non-Africans, the method uses a demographic model relating non-Africans, Africans and Neanderthals. The hidden state represents the African or Neanderthal haplotype that a target non-African genome coalesces with and the time of coalescence which can then be related to the density of mutations observed at a given locus. The model parameters are derived under the relevant demographic models [35]. The posterior probability of the hidden state at each locus is computed using a forward–backward algorithm [33] which is then used to derive the posterior probability of Neanderthal ancestry.

An archaic reference-free HMM-based method [36•] leverages the observation that introgressed segments contain a high density of variants found in the target genome but absent in an unadmixed modern human reference. The hidden state represents the ancestry within a genomic window while the observations are the density of private variants. HMM parameters are related to the population genetic parameters. Maximum likelihood estimates of the parameters are computed using the Baum–Welch algorithm and introgressed segments are identified using posterior decoding.

ChromoPainter [37] uses a HMM to describe each target haplotype as a copy of a panel of donor haplotypes. The method computes the probability that the target haplotype copies from each of the donor populations at a SNP. The model parameters are estimated using an expectation-maximization algorithm. Originally designed to study more recent admixture within human populations [38], the method has been recently applied to infer Denisovan segments in Papuans [19•].

Discriminative models

While HMMs model the joint probability of the hidden ancestry states and the observations, an alternate statistical approach directly models the conditional probability of the ancestry state given the observations. Parameter estimation for these models requires simulating labeled data under appropriate demographic models. CRFs [39] for identifying introgressed archaic segments [12] are an example of this approach. More recently, Durvasula and Sankararaman proposed ArchIE, a logistic regression predictor to detect introgressed haplotypes without an archaic reference by combining diverse population genetic summary statistics [40•]. The parameters of the logistic regression are estimated from labeled data generated from demographic simulations. The summary statistics include frequency-based features, distance-based features, and the S* statistic.

Methods that do not assume an unadmixed modern human reference

A limitation of previous methods is their assumption that the modern human reference lacks archaic ancestry. IBDMix [41•] detects introgressed segments using an archaic genome but does not rely on an unadmixed modern human reference. This method identifies introgressed segments as those shared identical-by-descent (IBD) between a modern human and archaic genome. It

computes log-odds (LOD) score at a given SNP of the likelihood that the target modern human and archaic individuals share an allele IBD and then uses dynamic programming to identify segments that maximize the sum of the LOD score.

Insights into human history from the analysis of introgressed sequences

Analysis of introgressed sequences are revealing the complex web of interactions within and between the ancestors of present-day and archaic humans. By analyzing the joint distribution of the frequency of introgressed Neanderthal fragments in present-day Europeans and East Asians, Villanea and Schraiber [42•] find support for multiple episodes of gene flow from Neanderthals in both European and East Asian populations. Analysis of introgressed Denisovan DNA in genomes from mainland Eurasia and Oceania have revealed introgression from at least three distinct Denisovan populations [18••,19•]. Present-day Papuans carry introgressed DNA from two deeply diverged Denisovan populations that are distantly related to the sequenced Denisovan genome [2] while East Asians carry an additional component of Denisovan ancestry that is more closely related. A pygmy population in Flores Island, Indonesia, was found to carry introgressed Denisovan DNA at lower levels relative to Melanesians, introgressed Neanderthal DNA at levels intermediate to those found in East Asians and Melanesians, but no DNA from other archaic hominins such as Homo Erectus or Homo Floresiensis [43]. These analyses suggest that the pygmy populations in Flores inherited Denisovan DNA from a population east of Wallace's line with subsequent dilution due to East Asian-related admixture. A recent study of archaic introgression from Iceland detected substantial introgressed Denisovan DNA which could be explained by Denisovan introgressed either into the ancestors of non-Africans or into the introgressing Neanderthals [10].

Attempts to infer Neanderthal and Denisovan introgression have generally assumed that west African populations harbor negligible archaic ancestry. Inferences from IBDMix, a method that does not rely on this assumption, suggests higher levels of Neanderthal ancestry in African populations than previously thought likely due to recent gene flow from from the ancestors of present-day Europeans into African populations as well as earlier gene flow from early modern humans into Neanderthals [41•]. Accounting for the increased sharing of introgressed Neanderthal sequences between Africans and Europeans leads Chen et al. [41•] to observe that the proportions of Neanderthal ancestry across non-African populations is more uniform [41•] than previously suggested [44,2], an observation that is consistent with a single Neanderthal introgression event into the ancestors of present-day non-Africans. Searching for diverged haplotypes unlikely to derive from Neanderthals or Denisovans, Wall *et al.* [20] attempted to discover ghost introgression by searching for diverged haplotypes unlikely to derive from Neanderthals or Denisovans. They document substantial ghost introgression in Khoesan and Central African Pygmy populations but find no evidence for such an introgression in the Andaman, in contrast to Mondal *et al.* [45]. Durvasula and Sankararaman [21] identify divergent haplotypes in the genomes of west African populations that are unlikely to trace their ancestry to Neanderthals or Denisovans. While these and other studies [46,47] have offered valuable clues into archaic ancestry in Africa, pinpointing the sources of this ancestry remains challenging in the absence of representative archaic genomes.

Outlook

Going forward, it will become increasingly important to develop methods that can accurately infer introgressed DNA from unknown archaic hominins while not making strong assumptions about population history . Methods that jointly infer demographic history and aspects of introgression (global parameters such as introgression fraction, timing and the location of introgressed segments) could formalize the support for introgression events and estimates of introgressed segments. Current approaches for detecting introgressed segments typically consider a limited number of statistics that summarize either patterns of allele frequency or LD. The accuracy of these methods can be improved by the inclusion of novel statistics that are sensitive to introgression . Recent work [48•] proposes a novel approach to compute moments of a wide range of two-locus statistics(including commonly used measures of LD) under complex demographic scenarios. For example, their work fits demographic models to a novel statistic that measures LD between low-frequency variants (the Dz-statistic) to provide evidence for archaic admixture. Novel summaries of the site frequency spectrum [49,21,50] have been applied genome-wide with the aim of fitting models of demographic history: it would be of interest to explore the utility of these statistics to identify introgressed segments. Since the local genealogy contains all the information to determine the ancestry of a given target haplotype, recent developments that allow inference of local genealogies from large samples offer another rich source of informative summary statistics [51,52]. In parallel to the choice of informative statistics, statistical models that can combine diverse statistics to yield calibrated inferences are essential. Supervised machine learning approaches have shown promise in making population genetic inferences by combining a large number of statistics with complex dependencies [53]. For example, deep learning approaches have been applied to identify complex demographic histories from genome-wide summary statistics [54,42] while random forests have been used to detect recent introgressed loci in Drosophila [55]. It would be of interest to adapt these approaches to study archaic introgression. The proliferation of methods to detect introgression (see Table 1) also necessitate systematic studies of their relative strengths in terms of accuracy (sensitivity,

Table 1

URLs of software for detecting introgressed archaic sequences

Method	URL
Sprime [18••]	https://github.com/browning-lab/sprime
diCal-admix [17•]	http://dical-admix.sourceforge.net
Skov et al. PLoS Genetics 2018 [36•]	https://github.com/LauritsSkov/Introgression-detection
ArchIE [40•]	https://github.com/sriramlab/ArchIE
IBDmix [41•]	https://github.com/PrincetonUniversity/IBDmix
ChromoPainter [37]	http://paintmychromosomes.com

false discovery rate, false positive rates), types of data required and modeling assumptions. Recent efforts to curate a library of population genetic models can be leveraged in efforts to benchmark these methods [56]. Studies that have compared inferences from multiple methods [17•,19•] recommend using a consensus approach to increase the reliability of inferences.

Multiple studies have documented the effects of natural selection on introgressed archaic DNA with evidence for purifying selection shown to reduce the genome-wide proportions [57,58] and altering the genomic distribution of archaic ancestry [12,13,23,22] as well as evidence for adaptive benefits of archaic alleles in specific loci [59]. These results motivate the need for methods that can infer introgression jointly with models of selection [34]. The ability to measure genetic variation in modern human populations through time offers another powerful source of information to understand the dynamics of introgression [60]. Availability of ancient genomes from Africa [61,62,47] opens up the possibility of understanding the sources of archaic ancestry. Methods that effectively leverage this data would likely need to account for the error processes and biases unique to ancient DNA data. We anticipate that high-resolution maps of introgression obtained by applying sophisticated methods to growing collections of ancient and modern genome sequences will lead to novel insights into human history and biology.

Conflict of interest statement

Nothing declared.

Acknowledgements

SS is supported in part by NIH grants R35GM125055 and an Alfred P Sloan Research Fellowship.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of special interest

1. Prüfer K, Racimo F, Patterson N, Jay F, Sankararaman S, Sawyer S *et al.*: **The complete genome sequence of a Neanderthal from the Altai Mountains**. *Nature* 2014, **505**:43-49 http://dx.doi.org/10.1038/nature12886.

2. Meyer M, Kircher M, Gansauge MT, Li H, Racimo F, Mallick S *et al.*: **A High-Coverage Genome Sequence from an Archaic Denisovan Individual**. *Science* 2012, **338**:222-226 http://dx.doi.org/10.1126/science.1224344.

3. Prüfer K, Filippo Cd, Grote S, Mafessoni F, Korlevic P, Hajdinjak M *et al.*: **A high-coverage Neandertal genome from Vindija Cave in Croatia**. *Science* 2017:eaao1887 http://dx.doi.org/10.1126/science.aao1887.

4. Mallick S, Li H, Lipson M, Mathieson I, Gymrek M, Racimo F *et al.*: **The Simons Genome Diversity Project: 300 genomes from 142 diverse populations**. *Nature* 2016, **538**:201 http://dx.doi.org/10.1038/nature18964.

5. Malaspinas AS, Westaway MC, Muller C, Sousa VC, Lao O, Alves I *et al.*: **A genomic history of Aboriginal Australia**. *Nature* 2016, **538**:207-214.

6. Pagani L, Lawson DJ, Jagoda E, Mörseburg A, Eriksson A, Mitt M *et al.*: **Genomic analyses inform on migration events during the peopling of Eurasia**. *Nature* 2016, **538**:238-242.

7. GenomeAsia100K Consortium: **The GenomeAsia 100K Project enables genetic discoveries across Asia**. *Nature* 2019, **576**:106.

8. The 1000 Genomes Project Consortium: **A global reference for human genetic variation**. *Nature* 2015, **526**:68-74 http://dx.doi.org/10.1038/nature15393.

9. Bergstrom A, McCarthy SA, Hui R, Almarri MA, Ayub Q, Danecek P *et al.*: **Insights into human genetic variation and population history from 929 diverse genomes**. *Science* 2020, **367**.

10. Skov L, Macia MC, Sveinbjornsson G, Mafessoni F, Lucotte EA, Einarsdottir MS *et al.*: **The nature of Neanderthal introgression revealed by 27,566 Icelandic genomes**. *Nature* 2020:1-6.

11. Vattathil S, Akey JM: **Small amounts of archaic admixture provide big insights into human history**. *Cell* 2015, **163**:281-284.

12. Sankararaman S, Mallick S, Dannemann M, Prüfer K, Kelso J, Pääbo S *et al.*: **The genomic landscape of Neanderthal ancestry in present-day humans**. *Nature* 2014, **507**:354-357 http://dx.doi.org/10.1038/nature12961.

13. Vernot B, Akey JM: **Resurrecting surviving Neandertal lineages from modern human genomes**. *Science* 2014, **343**:1017-1021 http://dx.doi.org/10.1126/science.1245938.

14. Seguin-Orlando A, Korneliussen TS, Sikora M, Malaspinas AS, Manica A, Moltke I *et al.*: **Genomic structure in Europeans dating back at least 36,200 years**. *Science* 2014, **346**:1113-1118 http://dx.doi.org/10.1126/science.aaa0114.

15. Sankararaman S, Mallick S, Patterson N, Reich D: **The combined landscape of Denisovan and Neanderthal ancestry in present-day humans**. *Curr Biol* 2016, **26**:1241-1247 http://dx.doi.org/10.1016/j.cub.2016.03.037.

16. Vernot B, Tucci S, Kelso J, Schraiber JG, Wolf AB, Gittelman RM *et al.*: **Excavating Neandertal and Denisovan DNA from the genomes of Melanesian individuals**. *Science* 2016:aad9416 http://dx.doi.org/10.1126/science.aad9416.

17. • Steinrücken M, Spence JP, Kamm JA, Wieczorek E, Song YS: **Model-based detection and analysis of introgressed Neanderthal ancestry in modern humans**. *Mol Ecol* 2018, **27**:3873-3888

The authors propose a method to detect introgressed segments under an explicit demographic model.

18. •• Browning SR, Browning BL, Zhou Y, Tucci S, Akey JM: **Analysis of human sequence data reveals two pulses of archaic Denisovan admixture**. *Cell.* 2018, **173** 53–61.e9

This study introduces a powerful method to detect introgression that does not require on an archaic genome. Analyzing introgressed Denisovan segments from this method, they find evidence for genetic ancestry from two deeply diverged Denisovan lineages.

19. • Jacobs GS, Hudjashov G, Saag L, Kusuma P, Darusallam CC, Lawson DJ *et al.*: **Multiple deeply divergent Denisovan ancestries in Papuans**. *Cell* 2019, **177**:1010-1021

This study builds a map of introgressed Denisovan DNA in genomes of individuals from Island Southeast Asia to show that Denisovans introgressed at least thrice in the history of modern humans.

20. Wall JD, Ratan A, Stawiski E, Kim HL, Kim C, Gupta R *et al.*: **Identification of African-specific admixture between modern and archaic humans**. *Am J Human Genet* 2019, **105**:1254-1261.

21. Durvasula A, Sankararaman S: **Recovering signals of ghost archaic admixture in the genomes of present-day Africans**. *Sci Adv* 2020. (in press).

22. Harris K, Nielsen R: **The genetic cost of Neanderthal introgression**. *Genetics* 2016, **203**:881-891 http://dx.doi.org/10.1534/genetics.116.186890.

23. Juric I, Aeschbacher S, Coop G: **The strength of selection against Neanderthal introgression**. *PLOS Genet* 2016, **12**: e1006340 http://dx.doi.org/10.1371/journal.pgen.1006340.

24. Racimo F, Sankararaman S, Nielsen R, Huerta-Sánchez E: **Evidence for archaic adaptive introgression in humans**. *Nat Rev Genet* 2015, **16**:359-371 http://dx.doi.org/10.1038/nrg3936.

25. Gittelman RM, Schraiber JG, Vernot B, Mikacenic C, Wurfel MM, Akey JM: **Archaic hominin admixture facilitated adaptation to out-of-Africa environments**. *Curr Biol* 2016, **26**:3375-3382.

26. Dannemann M, Andrés AM, Kelso J: **Introgression of Neandertal-and Denisovan-like haplotypes contributes to adaptive variation in human Toll-like receptors**. *Am J Hum Genet* 2016, **98**:22-33.

27. Quach H, Rotival M, Pothlichet J, Loh YHE, Dannemann M, Zidane N *et al.*: **Genetic adaptation and Neandertal admixture shaped the immune system of human populations**. *Cell* 2016, **167**:643-656.

28. Enard D, Petrov DA: **Evidence that RNA viruses drove adaptive introgression between Neanderthals and modern humans**. *Cell* 2018, **175**:360-371.

29. Simonti CN, Vernot B, Bastarache L, Bottinger E, Carrell DS, Chisholm RL *et al.*: **The phenotypic legacy of admixture between modern humans and Neandertals**. *Science* 2016, **351**:737-741 http://dx.doi.org/10.1126/science.aad2149.

30. McCoy RC, Wakefield J, Akey JM: **Impacts of Neanderthal-introgressed sequences on the landscape of human gene expression**. *Cell* 2017, **168** http://dx.doi.org/10.1016/j.cell.2017.01.038 916–927.e12.

31. Dannemann M, Kelso J: **The contribution of Neanderthals to phenotypic variation in modern humans**. *Am J Hum Genet* 2017, **101**:578-589.

32. Plagnol V, Wall JD: **Possible ancestral structure in human populations**. *PLOS Genet* 2006, **2**:e105 http://dx.doi.org/10.1371/journal.pgen.0020105.

33. Rabiner L, Juang B: **An introduction to hidden Markov models**. *IEEE ASSP Mag* 1986, **3**:4-16.

34. Racimo F, Marnetto D, Huerta-Sánchez E: **Signatures of archaic adaptive introgression in present-day human populations**. *Mol Biol Evol* 2017, **34**:296-317.

35. Steinrücken M, Kamm J, Spence JP, Song YS: **Inference of complex population histories using whole-genome sequences from multiple populations**. *Proc Natl Acad Sci U S A* 2019, **116**:17115-17120.

36. • Skov L, Hui R, Shchur V, Hobolth A, Scally A, Schierup MH *et al.*: **Detecting archaic introgression using an unadmixed outgroup**. *PLOS Genet* 2018, **14**:e1007641 http://dx.doi.org/10.1371/journal.pgen.1007641
The authors propose a HMM-based method to detect introgressed segments without using an archaic reference.

37. Lawson DJ, Hellenthal G, Myers S, Falush D: **Inference of population structure using dense haplotype data**. *PLoS Genet* 2012, **8**.

38. Hellenthal G, Busby GB, Band G, Wilson JF, Capelli C, Falush D *et al.*: **A genetic atlas of human admixture history**. *Science* 2014, **343**:747-751.

39. Lafferty JD, McCallum A, Pereira FCN: **Conditional random fields: probabilistic models for segmenting and labeling sequence data**. In *Proceedings of the Eighteenth International Conference on Machine Learning. ICML'01*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 2001, 282-289.

40. • Durvasula A, Sankararaman S: **A statistical model for reference-free inference of archaic local ancestry**. *PLoS Genet* 2019, **15**: e1008175
The authors propose a logistic regression classifier to detect introgressed segments without using an archaic reference.

41. • Chen L, Wolf AB, Fu W, Li L, Akey JM: **Identifying and interpreting apparent Neanderthal ancestry in African individuals**. *Cell* 2020
This study introduces a IBD-based method to detect introgression that does not require an unadmixed modern human reference genome Applying this method, the authors find more Neanderthal ancestry in Africans than previous estimates and substantially smaller differences in Neanderthal ancestry across Europeans and East Asians.

42. • Villanea FA, Schraiber JG: **Multiple episodes of interbreeding between Neanderthal and modern humans**. *Nat Ecol Evol* 2019, **3**:39-44
The authors analyze the joint distribution of introgressed segments in Europe and East Asia to show that Neanderthals introgressed more than once in the history of each population.

43. Tucci S, Vohr SH, McCoy RC, Vernot B, Robinson MR, Barbieri C *et al.*: **Evolutionary history and adaptation of a human pygmy population of Flores Island, Indonesia**. *Science* 2018, **361**:511-516.

44. Wall JD, Yang MA, Jay F, Kim SK, Durand EY, Stevison LS *et al.*: **Higher Levels of Neanderthal Ancestry in East Asians than in Europeans**. *Genetics* 2013, **194**:199-209 http://dx.doi.org/10.1534/genetics.112.148213.

45. Mondal M, Casals F, Xu T, Dall'Olio GM, Pybus M, Netea MG *et al.*: **Genomic analysis of Andamanese provides insights into ancient human migration into Asia and adaptation**. *Nat Genet* 2016, **48**:1066-1070.

46. Lorente-Galdos B, Lao O, Serra-Vidal G, Santpere G, Kuderna LF, Arauna LR *et al.*: **Whole-genome sequence analysis of a Pan African set of samples reveals archaic gene flow from an extinct basal population of modern humans into sub-Saharan populations**. *Genome Biol* 2019, **20**:1-15.

47. Lipson M, Ribot I, Mallick S, Rohland N, Olalde I, Adamski N *et al.*: **Ancient West African foragers in the context of African population history**. *Nature* 2020:1-6.

48. • Ragsdale AP, Gravel S: **Models of archaic admixture and recent history from two-locus statistics**. *PLoS Genet* 2019, **15**: e1008204
The authors propose methods to compute expectation of a broad family of two-locus statistics, that generalize commonly used statistics that measure LD, under complex demographic models. They then use these statistics for model fitting and discover the need to consider models of archaic admixture.

49. Chen H, Green RE, Pääbo S, Slatkin M: **The joint allele-frequency spectrum in closely related species**. *Genetics* 2007, **177**:387-398 Available from: http://www.genetics.org/content/177/1/387.

50. Rogers AR: **Legofit: estimating population history from genetic data**. *BMC Bioinformatics* 2019, **20**:526.

51. Kelleher J, Wong Y, Wohns AW, Fadil C, Albers PK, McVean G: **Inferring whole-genome histories in large population datasets**. *Nat Genet* 2019, **51**:1330-1338.

52. Speidel L, Forest M, Shi S, Myers SR: **A method for genome-wide genealogy estimation for thousands of samples**. *Nat Genet* 2019, **51**:1321-1329.

53. Schrider DR, Kern AD: **Supervised machine learning for population genetics: a new paradigm**. *Trends Genet* 2018, **34**:301-312.

54. Mondal M, Bertranpetit J, Lao O: **Approximate Bayesian computation with deep learning supports a third archaic introgression in Asia and Oceania**. *Nat Commun* 2019, **10**:1-9.

55. Schrider DR, Ayroles J, Matute DR, Kern AD: **Supervised machine learning reveals introgressed loci in the genomes of Drosophila simulans and *D. sechellia***. *PLoS Genet* 2018, **14**: e1007341.

56. Adrion JR, Cole CB, Dukler N, Galloway JG, Gladstein AL, Gower G *et al.*: **A community-maintained standard library of population genetic models**. *bioRxiv* 2019 http://dx.doi.org/10.1101/2019.12.20.885129.

57. Fu Q, Posth C, Hajdinjak M, Petr M, Mallick S, Fernandes D *et al.*: **The genetic history of ice age Europe**. *Nature* 2016, **534**:200-205.

58. Petr M, Pääbo S, Kelso J, Vernot B: **Limits of long-term selection against Neandertal introgression**. *Proc Natl Acad Sci U S A* 2019, **116**:1639-1644.

59. Huerta-Sánchez E, Jin X, Bianba Z, Peter BM, Vinckenbosch N, Liang Y *et al.*: **Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA**. *Nature* 2014, **512**:194-197.

60. Slatkin M, Racimo F: **Ancient DNA and human history**. *Proc Natl Acad Sci U S A* 2016, **113**:6380-6387.

61. Llorente MG, Jones ER, Eriksson A, Siska V, Arthur K, Arthur J *et al.*: **Ancient Ethiopian genome reveals extensive Eurasian admixture in Eastern Africa**. *Science* 2015, **350**:820-822.

62. Skoglund P, Thompson JC, Prendergast ME, Mittnik A, Sirak K, Hajdinjak M *et al.*: **Reconstructing prehistoric African population structure**. *Cell* 2017, **171**:59-71.