Trends in Genetics

CellPress
REVIEWS

Review

The Genomics of Human Local Adaptation

Jasmin S. Rees,[1,2] Sergi Castellano,[2,3] and Aida M. Andrés [1,*]

Modern humans inhabit a variety of environments and are exposed to a plethora of selective pressures, leading to multiple genetic adaptations to local environmental conditions. These include adaptations to climate, UV exposure, disease, diet, altitude, or cultural practice and have generated important genetic and phenotypic differences amongst populations. In recent years, new methods to identify the genomic signatures of natural selection underlying these adaptations, combined with novel types of genetic data (e.g., ancient DNA), have provided unprecedented insights into the origin of adaptive alleles and the modes of adaptation. As a result, numerous instances of local adaptation have been identified in humans. Here, we review the most exciting recent developments and discuss, in our view, the future of this field.

Highlights

Local adaptation has critically contributed to the (modest) genetic and phenotypic differentiation that exists among human groups, including in health-related traits that contribute to population health disparities.

Local adaptation has happened on alleles of diverse origin (on new, pre-existing, and introgressed alleles) and through diverse mechanisms (monogenic and polygenic adaptation).

Ancient DNA will play a key role in our understanding of local adaptation by improving inferences of past events. Further, it has revealed the importance of adaptive introgression, by which modern humans acquired adaptative alleles from archaic humans.

Novel analysis methods will improve our power to identify targets of local adaptation, especially those with weak signatures. Combining genetic and environmental information promises to improve the identification of genomic targets and the corresponding selective force.

Local Adaptation in Human Evolution

Local adaptation occurs when, due to genetic change, individuals from a population have higher average fitness in their local environment than those from other populations of the same species [1]. It is driven by natural selection that differs among populations and, over time, leads to genetic and phenotypic population differentiation [1]. The selected alleles underlying these adaptations may be more beneficial in one environment than others, or may be beneficial in one environment while being neutral in others [2].

Local adaptation is particularly important in human evolution because modern humans have rapidly colonised a wide variety of novel environments, both within Africa and in other territories after the 'Out of Africa' migration (50 000–70 000 years ago) [3,4]. The settlement in each territory, and the development of local cultural practices, drove local adaptations. In fact, we know that local adaptation contributed significantly to genic population differentiation [5] and that humans have locally adapted to diverse diets [6–8], pathogens [9,10], altitudes [11,12], and ambient temperatures [13]. Many of these adaptations result in average phenotypic differences amongst people of different ancestries, leading to population differences in important phenotypes, including health-related **traits** (see Glossary), disease prevalence, and drug response. Thus, investigating local adaptation can reveal the origin of critical population phenotypic differences. Because the recent split among populations means that most genetic and phenotypic variation is found within (rather than between) populations, identifying the cases where population differences have an adaptive origin is particularly important; these are the cases where genetic changes affect function, phenotype, and fitness. In fact, local adaptation is responsible for some of the most striking differences among human groups (Table 1, Key Table).

More generally, a deep understanding of local adaptation (and conversely, of neutral population differentiation) is vital in our understanding of how modern patterns of genetic variation in humans arose. Finally, unravelling the evolution and function of locally adaptated alleles has the potential to help us determine the functional consequences of phenotypically relevant alleles and reveal the genetic basis of the biological response to critical environmental pressures (e.g., how changes in certain genes mediate tolerance to low temperature or hypoxia), contributing directly to our understanding of the biological basis of important phenotypes, including disease (Box 1).

[1]UCL Genetics Institute, Department of Genetics, Evolution and Environment, University College London, London, UK
[2]Genetics and Genomic Medicine Programme, Great Ormond Street Institute of Child Health, University College London, London, UK
[3]UCL Genomics, Great Ormond Street Institute of Child Health, University College London, London, UK

*Correspondence:
a.andres@ucl.ac.uk (A.M. Andrés).

Here, we provide a global perspective of, in our view, the main developments in human local adaptation over the last few years. Several aspects of this field have been recently covered in excellent reviews (e.g., phenotypes under selection [14], adaptive archaic introgression [15], and extreme environmental adaptation [16]) and local adaptation is a large research field. We have thus focused on the most recent advancements in our understanding of the origin of adaptive alleles and modes of local adaptation, as well as new types of data and strategies to identify the targets of adaptation that can help circumvent the main challenges in the field. Some of the most striking known adaptations are presented in Table 1 and discussed in Boxes 1–3.

Origin of Selected Alleles

A beneficial allele can arise from a *de novo* mutation, be previously polymorphic (standing variation) or be introduced by **admixture** with other populations [17,18] (Figure 1A). Understanding the origin of alleles is important, not only because it informs on how adaptations originate, but also because each mode generates different signatures and requires different identification methods.

Selection on *de novo* mutation (SDN) acts on a new allele that is advantageous upon emergence because the selective pressure predates the mutation. The new, advantageous mutation tends to rapidly increase in frequency in the population under selection. However, an adaptive mutation must appear at the right time and in the right population and even then, most new advantageous mutations are lost due to **random genetic drift**. The likelihood of loss is reduced if selection is strong or the new allele is over-dominant and most beneficial in heterozygosity (since alleles are mostly in heterozygotes while their frequency is very low) [19].

Selection on standing variation (SSV) is when previous polymorphisms (segregating alleles) become advantageous following changes in selective pressure, usually resulting from environmental change [17,18,20,21]. The selected allele is thus older than the selective pressure. SSV was likely important in human local adaptation: the recent and rapid nature of local selection in non-Africans means that there was little time for the emergence of *de novo* mutations, the small effective population size translates in low effective mutation rate, and population growth allows the maintenance of low-frequency polymorphisms that would otherwise be lost (but that can mediate SSV) [18,20]. Here, selection may favour alleles that were previously slightly advantageous or deleterious (due to a shift in selective pressure), or maintained by **balancing selection** [20]. Alleles under balancing selection have great potential to mediate quick genetic adaptations when the population encounters new environments, since they necessarily affect phenotype and fitness, and are already at intermediate frequencies at the onset of selection [19,20].

Admixture between genetically distinct populations may also introduce beneficial variants in a population, a process of gene flow that has been pervasive in humans [22–26]. For example, alleles conferring lactase persistence spread from eastern African pastoralists to southern African Khoisan populations [23] and from early European pastoralists into modern European populations [27,28]. Also, a Duffy blood group allele of African origin under strong **positive selection** has provided widespread malaria resistance in modern Malagasy populations [29].

The extensive admixture between modern and archaic humans such as Neanderthals and Denisovans (called introgression) has been a major recent surprise [30–32]. It is now clear that whilst most non-neutral introgressed alleles were deleterious and removed by purifying selection [33,34], some were then advantageous in modern humans (adaptive introgression). Because only some human populations experienced introgression (from Neanderthals only into non-Africans, from Denisovans only into Melanesians and some other Asians [30–32,35–37]), all adaptive introgression generated population differences. It is perhaps unsurprising that archaic

Glossary

Admixture: gene flow between populations that have undergone genetic differentiation.
Approximate Bayesian computation (ABC): a computational method to estimate the posterior distributions of model parameters using Bayesian statistics; it can be used to infer the most likely evolutionary model for a locus and to infer relevant parameters.
Background selection: the effects of purifying selection on linked variation. It reduces the local effective population size and increases the effects of random genetic drift and can sometimes generate genetic differentiation between populations.
Balancing selection: the maintenance by selection of polymorphisms at a given locus in a population.
Incomplete lineage sorting: the random segregation of alleles across populations by genetic drift that leads to the coalescence in an ancestral population rather than within the current population/species. This results in differences between the gene tree and species tree given that some variants are shared between more diverged populations, but not between more closely related populations.
Linkage disequilibrium (LD): the nonrandom association of alleles along a chromosome, broken by recombination and affected by demographic history and selection.
Positive selection: the process where beneficial genetic variants (those that increase the fitness of individuals) rise in frequency in a population faster than they would under random genetic drift.
Random genetic drift: the stochastic changes in allele frequency due to the random transmission of alleles over generations.
Site frequency spectrum: the distribution of allele frequencies across polymorphic sites in a population, skewed under positive selection.
Supervised machine learning (SML): artificial intelligence analysis that can use patterns in the data to classify loci into prespecified models of evolution. It can be used to infer the most likely evolutionary model for a locus.
Trait: a character of an organism, here with a focus on those with a genetic origin.

Key Table

Table 1. Overview of Key Known Genes under Local Adaptation in Human Populations

Category	Adaptation	Gene target(s)	Population with adaptations	Refs[a]
Diet	Lactase persistence	*LCT*	Eurasians and Africans	[S1–S4]
	Fatty diets	*FADS*	Greenland Inuit	[S5]
	High arsenic levels	*AS3MT*	Argentinians	[S6]
	Low selenium levels	*DI2*, *SelS*, *GPX1*, *GPX3*, *CELF1*, *SPS2*, *SEPSECS*	Chinese	[S7–S9]
	Low iron levels	*HFE*	Europeans	[S10,S11]
	Low iodine levels	*TRIP4*	Central African Pygmies	[S12]
	Low calcium levels	*TRPV6*	Non-Africans	[S13,S14]
	Zinc levels	*SLC30A9*, *SLC39A8*	East Asians and Africans	[S9,S15]
	Ergothioneine deficiency	*IBD5 (SLC22A4*, *SLC22A5)*	Europeans	[S16,S17]
	Frequent starvation	*CREBRF*	Samoans	[S18]
	Alcohol consumption	*ADH1B*	Asians	[S19–S21]
Pathogens	Malaria resistance	*HBB*, *HBA*, *HPA*, *GYPA*, *GYPB*, *GYPC*, *G6PD*, *FY*	Sub-Saharan Africans	[S22–S24]
	'African sleeping sickness'	*APOL1*	Western Africans	[S25]
	Hepatitis C	*IFNL4*, *IL28B*	Eurasians	[S26–S28]
	HIV	*CUL5*, *TRIM5*, *APOBEC3G*	Biaka	[S28,S29]
	General immune function	*ADAM17*, *ITGAL*, *LAG3*, *IL6*, *LRRC19*, *PON2*, *OAS1*[b], *OAS* group[c], *HLA* group[c], *STAT2*[b], *STAT2*[c], *TLR1-TLR6-TLR10*[c]	Across populations	[S16,S30–S35]
Oxidative stress	High altitude[d]	*EGLN1*, *EPAS1*[b], *PPARA*	Tibetans (EPAS1 in Han Chinese)	[S36–S38]
		EGLN1, *EDNRA*, *PRKAA1*, *NOS2A*, *BRINP3*	Andeans	[S39–S41]
		VAV3, *CABARA1*, *ARNT2*, *THRB*, *BHLHE41*	Ethiopians	[S42,S43]
	Breath-hold diving	*PDE10A*, *BDKRB2*	Bajau (Indonesia)	[S44]
Cold resistance	Cold perception	*TRPM8*	Eurasians	[S45]
	Energy regulation and metabolism	*CPT1A*, *LRP5*, *THADA*	Siberians	[S46]
	Cardiovascular function	*PRKG1*	Siberians	[S46]
	Differentiation of brown and brite adipocytes	*TBX15*	Greenlandic Inuit	[S5]
UV exposure	Pigmentation changes	*SLC24A5*, *SLC45A2*, *OCA1-4*, *TYRP1*, *DCT*, *TYR*, *MC1R*[c], *HYAL2*[c]	Across populations	[S47–S53]
	Low vitamin D levels	*DHCR7*, *NADSYN1*	Northern European populations	[S16]
Height	Undetermined function	*DOCK3*, *CISH*, *HESX1*, *POU1F1*	Central African rainforest hunter-gatherers	[S54–S56]
	Undetermined function	Highly polygenic	Europeans	[S16,S57–S62]
Hair thickness	Undetermined function	*EDAR*	East Asians	[S1,S63]
Diet	Starchy foods	*AMY1*[e]	Across populations	[S64]
Unknown	Undetermined function	17q.21.31 gene region[e]	Icelandic	[S65]

Box 1. Impact of Local Adaptation in the Health of Modern Populations

Past local adaptations can result in population differences in the genetic risk or prevalence of disease, contributing to population health disparities. These can be due to previously advantageous alleles being deleterious in modern societies. For example, Samoan populations have a higher prevalence of type 2 diabetes and related metabolic disorders that likely reflects a frequent starvation genotype [108]. A variant of the *CREBRF* gene, which allows rapid weight gain, was likely adaptive under highly variable dietary conditions, but is problematic now since calorific availability has dramatically increased [108]. Similar issues are seen in Canadian, Greenlandic Inuit, and Siberian populations with a particular *CPT1A* gene variant, which maintains energy and sugar homeostasis under low carbohydrate intake but is linked to hypoketotic hypoglycaemia and high infant mortality under modern diets [109].

In other cases, advantageous alleles may have deleterious pleiotropic effects. Adaptations to low amino acid levels, likely due to agricultural diets, may lead to higher risk of celiac disease and inflammatory bowel disease today [27,110]. Alleles mediating adaptations to malaria, African sleeping sickness, and cold ambient temperature are also associated with increased prevalence of sickle cell anaemia, chronic kidney disease, and migraine in various populations [13,111,112].

Modern populations may also be exposed to novel conditions due to recent migration or ecological change. Micronutrient levels in soils are highly variable and have resulted in genetic adaptations to at least toxic arsenic levels and low selenium levels [7,70]. Individuals from other populations that migrate to these environments (and that lack such local adaptations) face numerous complications if micronutrient levels are not managed or supplemented via other means.

Population-specific adaptation can result in differences in the outcome of treatment. The pseudogenising allele of *IFNL4*, adaptive in Eurasian populations, results in more rapid clearing of hepatitis C virus infection in these populations and a better health outcome following treatment [54]. There are also multiple alleles associated with protection against HIV that are fixed (*CUL5*) or at high frequency (*TRIM5, APOBEC3G*) in the Biaka, suggesting population differences in resistance to HIV [113].

These local differences ultimately contribute to population differences in the genetic basis of common disease and the response to treatment, reducing the transferability of GWAS and contributing to an ongoing bias where highly studied populations benefit most from genomic and personalised medicine. An improved understanding of the adaptations that populations have historically undergone will thus help develop targeted healthcare strategies and identify the populations most at risk.

alleles were beneficial in Eurasia because Neanderthals and Denisovans inhabited these territories since shortly after their divergence from modern humans (estimated over 800 000 years ago) and in theory had time to locally adapt; humans could rapidly acquire these adaptions through introgression, easing their transition into these novel environments [30,36,38].

The importance of adaptive introgression has been one of the main recent realisations in the field and it has already been shown to contribute to immune function, oxidative stress, and pigmentation (Figure 2) [39–43]. Introgressed alleles have been suggested to be particularly important in resistance against viruses, especially against RNA viruses in Europeans [44]. In rare instances, multiple archaic alleles are maintained, perhaps because genetic diversity itself was advantageous [43,45].

Signatures of Local Adaptation

Positive selection leaves characteristic signatures in genomes, which differ depending on the genetic basis of the selected phenotype(s), the origin of the allele, and the tempo and strength of selection.

Monogenic adaptation shows strong selection signatures in a single adaptive allele. These signatures have been studied for decades, so most known adaptations are monogenic (Table 1 and

Notes to Table 1:
[a]Selection noted as across populations indicates that selection is seen differentially across multiple populations according to the strength of the selective pressure. Table references are listed in the supplemental information online.
[b]Indicates gene variants that are a result of adaptive introgression with Denisovan populations (in *OAS1* it may be the result of an introgression event with an older archaic human or be the result of a double introgression event [S31,S66]).
[c]Indicates gene variants that are a result of adaptive introgression with Neanderthal populations.
[d]For a more comprehensive list of the genes identified as targets of selection under high altitude, see [S67].
[e]Selection acting on structural variations (deletions, insertions, inversions, duplicates, and copy number variations [S68]).

Box 2). Selection signatures were historically categorised into those from 'hard' or 'soft' sweeps [18,21,46,47], where hard sweeps are a result of strong SDN [18,21,46,47] and soft sweeps result from weak selection, SSV, or recurrent mutations [18,21,46,47] (although recurrent mutations are not particularly likely in humans, owing to our low effective mutation rate [18]).

The frequency of the selected allele, and hence the differentiation between populations, depends on the strength of selection and is largely independent of the allele origin. Thus, unusually large allele frequency differentiation in a SNP can be considered a nearly universal signature of strong local adaptation. The mode of adaptation impacts mostly the signatures of linked variation (Figure 1B), with alleles that rapidly increase in frequency having linked haplotypes of low diversity and many low-frequency variants, and their genealogies having a star-like shape (a long internal branch, short terminal tips) [48]. With SDN, classical signatures in the genomic region neighbouring the selected site include extended high population differentiation, skews in the **site frequency spectrum**, and extended haplotype homozygosity in long genomic regions (a selective sweep), with the strength of signatures largely dependent on the timing and strength of the selective sweep [12,48–51]. Under SSV the signatures of linked variation are usually highly reduced, largely due to the presence of the selected allele on multiple haplotype backgrounds and populations. Thus, linked signatures of SSV can be highly variable depending on the age and frequency of the segregating allele [17]. Nevertheless, signatures of SSV may resemble those of SDN if the beneficial allele only slightly precedes the onset of selection [18].

Determining the relative importance of SSV versus SDN in human evolution has been a recent focus of investigation and debate [52,53]. Whilst distinguishing between the two modes of

Box 2. Locally Adapted Genes and Traits in Humans

A selection of genes identified as playing a role in human local adaptation are presented in Table 1. Key observations are:

Genetic changes mediate adaptations to local cultural pressures. These can allow unique cultural practices, such as breath-hold diving in the Bajau [114].

Adaptive responses to agricultural practices are widespread. Lactase persistence is easily the best-known example of local adaptation (reviewed by [115]) and emerged with agriculture [22,23,115]. Additional agricultural adaptations to diet (micronutrient or amino acid levels) and to pathogens (facilitated by densely packed living conditions and exposure to zoonotic diseases [22,23,116]) accompanied the switch to an agricultural lifestyle.

Micronutrients are typical targets of dietary local adaptation. Micronutrients are necessary and irreplaceable in the diet, but toxic in excessive amounts. Local adaptation has been identified in multiple micronutrients [7,70], particularly in the regulatory variants maintaining micronutrient homeostasis [117].

There is strong and diverse evidence for selection in response to pathogen pressure. This highlights disease as a major driver of adaptation [10,27,44]. There is also evidence for the role of adaptive introgression in pathogen resistance [30,41,44].

The same adaptive phenotype may be achieved by different genes in different populations. This is particularly clear for altitude adaptation in Ethiopian, Andean, and Tibetan populations (the latter population acquiring an adaptive gene variant from archaic humans) [12,82,100]. Further, multiple different variants upstream of the *LCT* gene region evolved independently in African and European populations to allow people to consume milk past weaning [22].

Adaptive introgression from archaic humans has played a fundamental role in adaptation to multiple environmental factors. This includes resistance to pathogens, UV exposure, and high altitude [29,39,41–43,82].

The adaptive function of some population-specific phenotypes remains unclear. For example, the short stature of multiple rainforest hunter-gatherers (found in Central Africa, South America, and Southeast Asia) appears to be selected by a shared environmental pressure that is currently unknown [118]. A selective origin of the cline in height in Europeans has also been suggested, although the statistical validity of this result is currently under debate [48,96,97].

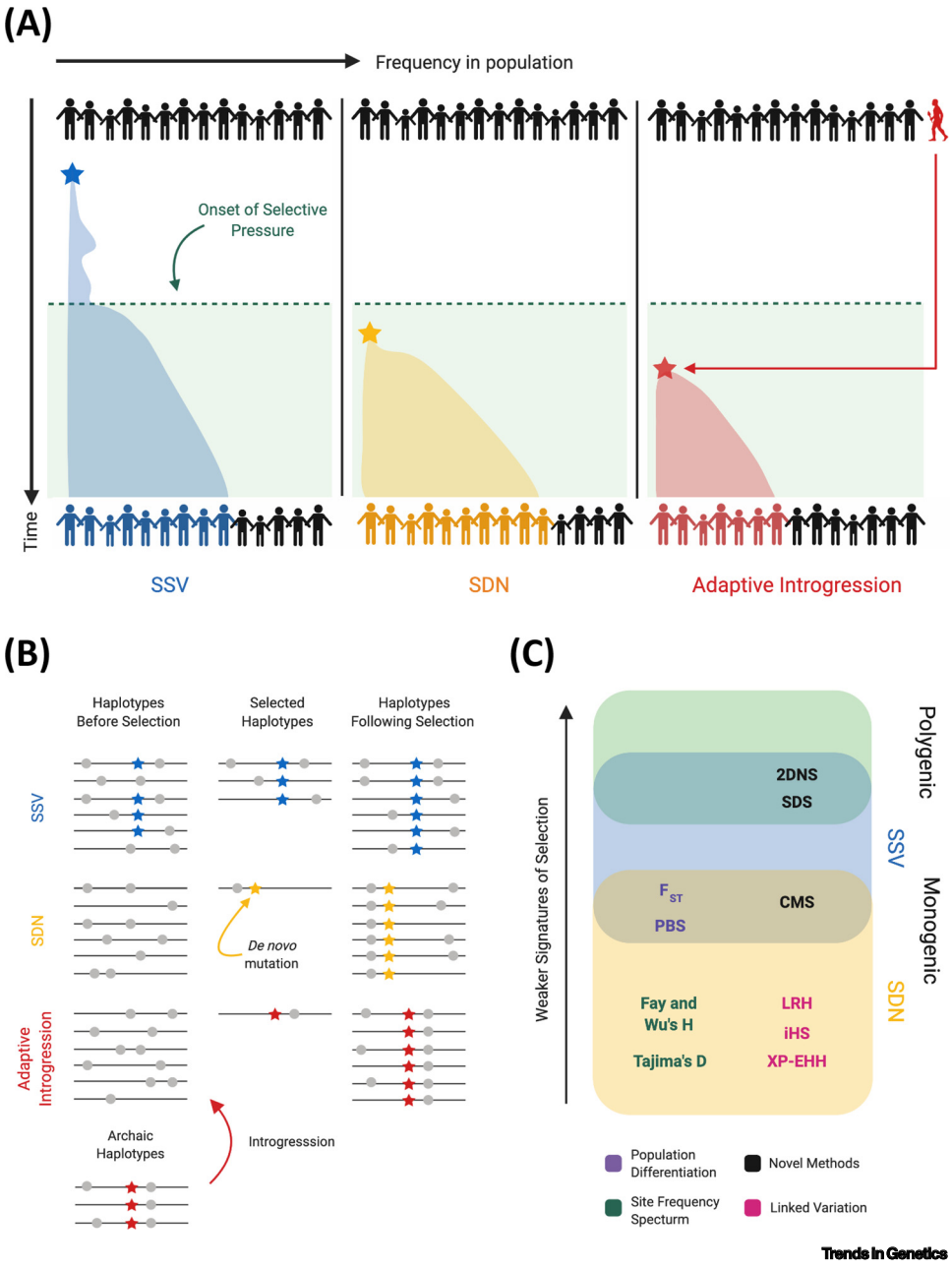

Figure 1. Characterisation of Selection on Standing Variation, Selection on *de Novo* Mutation and Adaptive Introgression. (A) Cartoon showing the rise in frequency of an allele through a population according to its origin. The stars represent mutations (blue: mutation present in the population prior to selection; yellow: mutation occurring after the onset of selection; and red: mutation present in an archaic population which spreads through a receiving population following an admixture event). The frequency of each mutation in a population following a selection event is represented by the number of people icons of their respective colours under each panel. The red walking person icon in the top right represents an archaic human. (B) Cartoon depicting the haplotypes arising from selection on standing variation (SSV), selection on *de novo* mutation (SDN), and adaptive introgression. Stars represent the beneficial allele and grey circles represent neutral linked polymorphic alleles. (C) Some of the most common statistics used to identify local adaptation (not a comprehensive list) [12,48,49,51,121–124]). Figure created using Biorender.com. Abbreviations: iHS, integrated haplotype score; LRH, long range haplotype; PBS, population branch statistic; SDS, Singleton density score; XP-EHH, cross-population extended haplotype homozygosity.

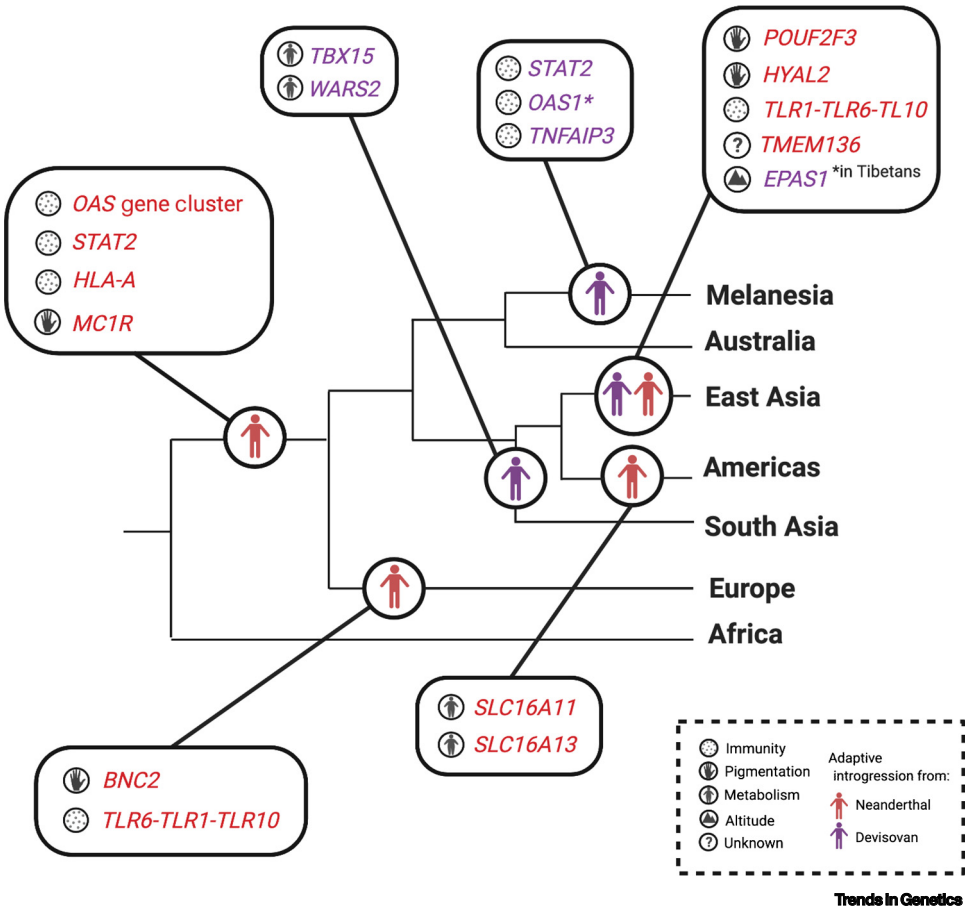

Figure 2. Visualisation of the Events of Adaptive Introgression from Archaic Human Populations into Modern Humans. The red and purple persons indicate human populations having undergone adaptive introgression with Neanderthals and Denisovans, respectively [9,31,38,39,41–43,63,125–131]. Strongly or moderately supported examples of adaptive introgression from archaic populations are shown across a simplified modern human lineage (phylogeny edited from [132]). Adaptive introgression events are shown on the lineages on which they happened. The gene(s) believed to be under selection in each case is noted in the linked boxes. *STAT2* shows two haplotypes arising from introgression, haplotype N from Neanderthals in non-Africans and haplotype D from Denisovans in Melanesians [38,130]. *OAS1* is highlighted as, whilst it shares much variance with the Denisovan gene region, the direct source is suggested as an unknown archaic population [40]; its presence in Melanesians could possibly be the result of a double introgression event [133]. Figure created using Biorender.com.

adaptation is challenging, approaches that integrate multiple aspects of the signature, for example, **Approximate Bayesian Computation (ABC)** or **Supervised Machine Learning (SML)** [17,47], allow us to start addressing this question. ABC has been used to determine the mode of selection in multiple studies [13,54–57], but only very recently has it been applied genome-wide, due to computational requirements [58]. Applications of SML have suggested that these methods perform relatively well when demographic history is unknown and that SSV and polygenic selection are dominant in human adaptation [46,47,59], although this has been contested [60]. Because selection signatures depend on the age, strength, and frequency of the selected allele, none of which follow discrete distributions, a binary categorisation (hard versus soft) does not necessarily represent selection signatures [59] and more work is needed to integrate all different models of selection.

Adaptive admixture and introgression pose a double challenge: distinguishing genome segments shared among populations due to admixture/introgression from those shared due to shared ancestry or **incomplete lineage sorting** [39] and distinguishing admixed/introgressed segments that are adaptive from those that are neutral. Introgressed segments are long, accompanied by a high level of observed **linkage disequilibrium (LD)**, young, present only in some human populations, and unusually similar to the archaic genomes [61,62]. These characteristics allow us to identify introgressed segments [38,39,43,63] but they mean that classical selection signatures cannot be used to identify adaptive introgression as we expect, for example, long-range LD under neutral introgression. Given the challenges associated with using signatures of linked variation, unusually high frequency of an introgressed segment, when compared with the empirical distribution of other introgressed segments, is probably the strongest evidence for adaptive introgression [41–43].

Polygenic adaptation refers to cases when multiple adaptive alleles are under selection in a population, resulting in concerted shifts in allele frequencies that shift the phenotype (even when no single allele experiences strong allele frequency change). Polygenic adaptation is probably common, since most traits are likely polygenic, and it has been proposed to explain adaptations related to diet, metabolism, pathogen resistance, or altitude [64–71].

However, the signatures of polygenic selection are often weak: the multiple small frequency shifts, which may also occur at different time-points, mean that polygenic adaptation is often hard to distinguish from neutral drift [48,72]. Further, the signatures are also highly variable, dependent on the number of loci under polygenic adaptation, their effect sizes, and the origin of these alleles [64,65,67–69,71]. Importantly, polygenic adaptation does not necessarily preclude the presence of strong signatures of selection and, conversely, strong signatures of selection do not negate polygenic adaptation. If many variants controlling a trait have negative pleiotropic effects, a single allele may be selected [73]. Further, the alleles with the largest effect sizes may sweep to fixation, accompanied by additional small increases in frequency of other alleles [73].

Identifying Local Adaptation

Identifying instances of local adaptation remains difficult due to the complexity of signatures and the potential misidentification of other processes as positive selection. Alleles can increase in frequency neutrally (due to random genetic drift, often aided by demographic events), sometimes quickly enough that their patterns mimic those of positive selection. Therefore, when there is uncertainty surrounding the demographic model, an unfortunate commonality, it can be difficult to tease apart demographic processes from selection [13,17,46]. Further, purifying selection must be accounted for alongside demography. For example, **background selection** can increase the effect of random genetic drift and cause strong differentiation between populations, as does local adaptation [74]. Recessive deleterious introgressed alleles can generate heterosis and signatures that can potentially mimic, in some instances, those of adaptive introgression [75,76].

Therefore, any methods to identify positive selection must account for neutral genetic change, demographic factors, and purifying selection. Distinguishing between these processes remains among the greatest challenges in population genetics, but is also one of the topics where recent advances are most exciting. Besides identifying the genetic signatures described earlier, correlations between genetics and the environment represent a promising avenue to identify targets of local adaptation. Naturally, both are limited in the identification of weak signals of selection [64,77].

Identifying Genetic Signatures

A common method to identify signatures of local adaptation is sampling many loci throughout the genome (e.g., all SNPs or genomic windows of a certain length), calculating summary statistics that capture selection signatures at each locus, and comparing these statistics with expectations under neutrality to identify loci with signatures that are unexpected under neutrality. While arguably the best approach, this relies on an accurate neutral expectation based on neutral simulations (which in turn depends on the accuracy of demographic inferences) [78]. The most common approach, thus, is to compare each locus with the empirical genomic distribution to identify outliers, which we cannot categorically state harbour signatures unexpected under neutrality, but that are strong candidate targets of selection [27]. A limitation is that most classic summary statistics were designed to identify strong SDN and are often poorly equipped at detecting weaker signatures of selection (Figure 1C) [12,49,51], with a few notable exceptions [79,80]. For this reason, many new methods focus on the allele frequency change on putatively selected alleles themselves. Including or conceptually rooted in F_{ST}, such methods include levels of exclusively shared differences (LSD) [81], population branch statistic (*PBS*; comparing three pairwise F_{ST} values between three populations [12]), and several derivatives: PBS_{n1} [82], PBS_{nj} [83], and population branch excess (*PBE*) [84]. To better account for complex population structure (when multiple populations are analysed), methods such as BayEnv can correct for the covariance among populations due to shared ancestry [77,85].

Ancient DNA (aDNA) provides snap-shots of allele frequencies in the past and as such it is an invaluable resource to identify past rapid allele frequency changes [5], to quantify the genome-wide influence of local adaptation in population differentiation [5], to confirm previous evidence of selection in individual sites [27,56,86], and to identify targets of local adaptation [5,27,56]. Due to the growing availability of historical and prehistorical samples, most studies have been conducted in Europe, showing, for example, low frequency of alleles conferring lactase persistence in Neolithic populations (as well as inferring the age and geographical origin of the selected European allele) [27,86] and questioning the link between selection on *FADS* and *AMY1* and the development of agriculture [56]. In cases of admixture, where identification of genetic adaptation is particularly challenging, ancient genomes can help identify the population of origin of alleles, and detect selected sites as those with unusually high contribution of one particular ancestral population (the one carrying the beneficial allele) at a particular locus [27,87]. This has helped identify loci mediating adaptation postadmixture in Europeans [27] and Americans [87] and will definitively improve our understanding in other populations as aDNA becomes available.

aDNA is also integral to identifying the alleles that arise from introgression [39]. As mentioned earlier, identifying adaptive admixture and introgression involves identifying admixture events (often using aDNA) [30,35–37] and then distinguishing neutral from adaptive admixed/introgressed alleles, usually via the higher frequency of the latter. This is typically done still in a series of steps [41–43], rather than using one unified method, with few exceptions [63,88].

Promising new approaches to identify selection are based in SML [46,47,59,89,90] or purely genealogical methods: if gene genealogies can be fully reconstructed, then selection can be more directly inferred from the genealogical tree (reviewed in [91]). This powerful technique to infer selection has been implemented in programmes such as ARGweaver, which is able to detect selection using Ancestral Recombination Graphs [92,93]. While inferring genealogies for large samples is computationally intensive, two recent methods (Relate, tsinfer) are able to efficiently do so. Tsinfer especially boasts computational efficiency (being able to reconstruct relatedness for up to millions of samples) whilst Relate is able to estimate branch lengths and allele ages, allowing its application to identify both monogenic and polygenic adaptation [94,95]. These methods, perhaps as a hybrid

model or in new flavours, promise to have higher power than existing methods to identify genetic adaptation, as well as potentially rewriting how we represent and store genetic data.

Still, identifying polygenic adaptation remains challenging as it involves moderate, concerted allele frequency change of multiple small-effect loci. Approaches to identify such changes include combining genome-wide association studies (GWAS) with population genetic modelling (searching for positive covariance between alleles of similar effects [64,65]). Recent studies suggest that GWAS-based methods can overestimate polygenic adaptation if the trait differs amongst populations and there is hidden population stratification in the GWAS sample [96,97]. Nevertheless, methods to identify polygenic adaptation are not necessarily limited to allele frequency change. Additional statistics include 2DNS, a McDonald-Kreitman-based test designed to consider multiple gene variants [66], and the Singleton Density Score (SDS), a summary statistic that considers the distance between the nearest singletons to a focal SNP [48]. The latter can identify extremely recent selection [48] but may still be influenced by population stratification if GWAS SNPs are used [96]. Nevertheless, methods to identify polygenic adaptation are not limited to GWAS hits: some focus on gene sets [67,98], gene networks [68,69], and admixture graphs [99], which are potentially more robust to these issues.

Environmental Correlations

Correlations between environmental factors and allele frequency, when they are more substantial than expected under population history and gene flow, provide among the strongest evidences of local adaptation [77]. For example, unexpected correlations of allele frequency with latitude and temperature, coupled with independent signatures of positive selection, provide evidence of local adaptation of the cold receptor *TRPM8* to cool ambient temperature [13]. BayEnv is probably the most widely used method to investigate such correlations at the genome scale [77]. Further, populations living in extreme environments can help reveal adaptations to harsh selective pressures, such the arsenic-rich soil in regions of Argentina [7]. When several populations live under similar environmental pressures, comparing them can reveal the same gene(s) conferring shared adaptations across populations or different genes/gene variants mediating independent adaptations (as in the case of high altitude [82,100,101] or lactose in the diet [22]).

The Causes and Consequences of Local Adaptation

Ultimately, we aim to identify both the selective force responsible for local adaptation and the consequences that these events have in current population differences. Both are challenging but possible. To identify the selective force, two main avenues are promising. First, determining the timing of the adaptations. Here aDNA is highly valuable, as it can show the frequency of the selected allele through time and estimate when the selective pressure was first exerted [27,56,86]. Second, establishing the function of the genes or genomic elements under selection may help identify the selective force. As usual in genetics, model organisms can be useful [102,103]. For example, the signature of positive selection on *SLC24A5* in European populations could be ascribed to pigmentation because the gene affects pigmentation in zebrafish [102]. Also, the insertion of the human form of the *EDAR* gene (under strong selection in Asian populations and associated with hair thickness and dental morphology) in mice revealed a role of the gene in the mammary and eccrine glands [103], suggesting a plausible selective force.

Still, experiments in model organisms can hardly inform on the consequences of human-specific genetic variants. Doing so is challenging because determining the consequences of human variants on the phenotype often relies on large association studies, while determining their consequences at the molecular level often relies on analyses of transcriptomics, metabolomics, and epigenomics datasets, at least outside of protein-coding regions (Box 3). High-throughput assays can categorise SNP function, including, but not limited to, exploring the effect of individual

Box 3. Epigenetics and Local Adaptation

Due to the links between epigenetic response and environmental change, recent studies have begun to question the importance of epigenetics in local adaptation. Epigenetic response, the somatically heritable change of chemical modification without changes in the DNA sequence, can occur under local changes in environment, especially during development [119]. Such chemical modifications, the most well-studied being DNA methylation, can take place over a much smaller time scale than genetic adaptation and may allow populations to adapt over a few generations, or even within a lifetime [119].

However, it is difficult to show that epigenetic changes are adaptive, rather than a response to change/stress. Unless there are convergent epigenetic changes over populations experiencing the same selective pressures (where an adaptive value may be easier to infer), it is difficult to determine the nature of epigenetic change. Nonetheless, a recent study has suggested a link between positive selection and epigenetic change [120]. The genetic variants associated with the methylation variation between populations of Central Africa, particularly between those with historically different lifestyles, show signatures of positive selection [118,120]. Variants highlighted as under selection for pygmy height (*DOCK3*, *MAPKAPK3*, and *CISH*) are also differentially methylated across populations according to their lifestyle [118,120]. Even if today the role of epigenetic change in local adaptation remains unclear, epigenetic change has great potential to mediate quick, plastic adaptations to the environment and will undoubtedly become a focus in studies of local adaptation.

genetic variants on methylation, transcription, and protein expression [104]. Experimentally, induced pluripotent stem cells and stem cell-derived organoids represent a new and exciting way to test the phenotypic consequences of mutations in human cells [105–107]. The development of technologies such as genome editing may further validate and possibly improve on the functional knowledge of variants, both in human cell lines and in other organisms.

Concluding Remarks

Investigations into human local adaptation have benefited enormously from recent developments both in datasets and methods, integrating exciting developments in aDNA, bioinformatics, and our conceptual understanding of how selection may be etched in the genome.

Still, important challenges remain in identifying weaker signatures of selection. This is especially relevant when considering polygenic adaptation and its likely prevalence throughout human adaptation. There has been much progress in identifying these selection signatures in recent years and we predict that continued fervour in tackling this topic will result in many more refined and even novel techniques that will be able to identify previously elusive signatures.

It is also clear that some populations are relatively understudied, limiting our understanding of the effects of local adaptation and biasing our evolutionary knowledge (e.g., in terms of the origin of population differences for health-relevant traits) to well-studied populations. In addition, data limitations can bias our understanding of local adaptation towards previously identified loci and well-known selective pressures. Future studies should be prioritised to populations that have been previously neglected, especially if these populations have complex demographic histories that are at present unresolved. The increasing availability of aDNA will undoubtedly continue to improve our inferences on demographic histories, including classifying complex admixture events, and selective histories of populations.

Advances in studying the genomics of human local adaptation will ultimately improve our insight into how multiple factors have influenced the evolutionary dynamics of different populations, as well as highlighting the complexity of adaptation in human history (including the influence of archaic populations into modern human genetic variance). Further, studies in this field inform on the phenotypically relevant genetic differences among populations, the origin of these differences, and how they can help address health inequalities among human groups (see Outstanding Questions).

Outstanding Questions

How much has local adaptation contributed to the (average) phenotypic differentiation that exists among human populations? Which selective pressures drove those adaptations? Which loci mediated those adaptations?

How, and how much, has local adaptation generated differences in health-related phenotypes among human populations? How can evolutionary studies help us identify such cases and improve healthcare in modern societies?

How can we identify weak signatures of selection, which likely represent most signatures of local adaptation? Should we aim to differentiate weak signatures from neutral patterns, or rather refine what constitutes a signature of selection? How should investigation into polygenic selection be best performed, especially for traits that we know differ among populations and can confound GWAS?

What is the level of detail of demographic inferences necessary to minimise false positives when identifying signatures of selection? How much does weak differentiation and re-admixture (which has been pervasive throughout human history) influence our inferences of local adaptation?

How are contemporaneous and ancient DNA samples best used to identify the targets of selection and to inform selective dynamics over time? How will ancient DNA change our understanding of the demographic and adaptive history of populations in geographic regions where ancient DNA is still scarce?

How big a role has admixture played in human local adaptation? What about adaptive introgression with archaic humans? Will studies of ancient DNA reveal older and currently uncharacterised introgression events? Has adaptive introgression influenced adaptation in Africa?

If a selection signature has been identified, how should the selective force be determined? What are the best methods to characterise, functionally and phenotypically, the genes that mediated local selection?

Acknowledgements

We are grateful to Joshua Schmidt, Hernán Burbano, and Mark Stoneking for comments on the manuscript and the Andrés group for discussions on the effects of local adaptation. A.A. is funded by UCL's Wellcome Institutional Strategic Support Fund 3 (grant reference 204841/Z/16/Z). S.C. and J.R. are funded by NIHR GOSH BRC. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health.

Supplemental Information

Supplemental information associated with this article can be found online at https://doi.org/10.1016/j.tig.2020.03.006.

References

1. Savolainen, O. *et al.* (2013) Ecological genomics of local adaptation. *Nat. Rev. Genet.* 14, 807–820
2. Tiffin, P. and Ross-Ibarra, J. (2014) Advances and limits of using population genetics to understand local adaptation. *Trends Ecol. Evol.* 29, 673–680
3. Haber, M. *et al.* (2019) A rare deep-rooting D0 African Y-chromosomal haplogroup and its implications for the expansion of modern humans out of Africa. *Genetics* 212, 1421–1428
4. Soares, P. *et al.* (2011) The expansion of mtDNA haplogroup L3 within and out of Africa. *Mol. Biol. Evol.* 29, 915–927
5. Key, F.M. *et al.* (2016) Human adaptation and population differentiation in the light of ancient genomes. *Nat. Commun.* 7, 10775
6. Davy, T. and Castellano, S. (2018) The genomics of selenium: its past, present and future. *BBA Gen. Subjects* 1862, 2427–2432
7. Schlebusch, C.M. *et al.* (2015) Human adaptation to arsenic-rich environments. *Mol. Biol. Evol.* 32, 1544–1555
8. Fumagalli, M. *et al.* (2015) Greenlandic Inuit show genetic signatures of diet and climate adaptation. *Science* 349, 1343–1347
9. Nédélec, Y. *et al.* (2016) Genetic ancestry and natural selection drive population differences in immune responses to pathogens. *Cell* 167, 657–669
10. Fumagalli, M. *et al.* (2011) Signatures of environmental genetic adaptation pinpoint pathogens as the main selective pressure through human evolution. *PLoS Genet.* 7, 1002355
11. Bigham, A.W. and Lee, F.S. (2014) Human high-altitude adaptation: forward genetics meets the HIF pathway. *Genes Dev.* 28, 2189–2204
12. Yi, X. *et al.* (2010) Sequencing of fifty human exomes reveals adaptation to high altitude. *Science* 329, 75–78
13. Key, F.M. *et al.* (2018) Human local adaptation of the TRPM8 cold receptor along a latitudinal cline. *PLoS Genet.* 14, 1–22
14. Fan, S. *et al.* (2016) Going global by adapting local: a review of recent human adaptation. *Science* 354, 54–59
15. Racimo, F. *et al.* (2015) Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet.* 16, 359–371
16. Ilardo, M. and Nielsen, R. (2018) Human adaptation to extreme environmental conditions. *Curr. Opin. Genet. Dev.* 53, 77–82
17. Peter, B.M. *et al.* (2012) Distinguishing between selective sweeps from standing variation and from a de novo mutation. *PLoS Genet.* 8, e1003011
18. Hermisson, J. and Pennings, P.S. (2017) Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods Ecol. Evol.* 8, 700–716
19. DeGiorgio, M. *et al.* (2014) A model-based approach for identifying signatures of ancient balancing selection in genetic data. *PLoS Genet.* 10, e1004561
20. De Filippo, C. *et al.* (2016) Recent selection changes in human genes under long-term balancing selection. *Mol. Biol. Evol.* 33, 1435–1447
21. Hermisson, J. and Pennings, P.S. (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169, 2335–2352
22. Tishkoff, S.A. *et al.* (2007) Convergent adaptation of human lactase persistence in Africa and Europe. *Nat. Genet.* 39, 31–40
23. Macholdt, E. *et al.* (2014) Tracing pastoralist migrations to southern Africa with lactase persistence alleles. *Curr. Biol.* 24, 875–879
24. Jeong, C. *et al.* (2019) The genetic history of admixture across inner Eurasia. *Nat. Ecol. Evol.* 3, 966–976
25. Brace, S. *et al.* (2019) Ancient genomes indicate population replacement in early Neolithic Britain. *Nat. Ecol. Evol.* 3, 765–771
26. Posth, C. *et al.* (2018) Reconstructing the deep population history of Central and South America. *Cell* 175, 1185–1197
27. Mathieson, I. *et al.* (2015) Genome-wide patterns of selection in 230 ancient Eurasians. *Nature* 528, 499–503
28. Allentoft, M.E. *et al.* (2015) Population genomics of Bronze Age Eurasia. *Nature* 522, 167–172
29. Pierron, D. *et al.* (2018) Strong selection during the last millennium for African ancestry in the admixed population of Madagascar. *Nat. Commun.* 9, 932
30. Prüfer, K. *et al.* (2014) The complete genome sequence of a Neanderthal from the Altai Mountains. *Nature* 505, 43–49
31. Jacobs, G.S. *et al.* (2019) Multiple deeply divergent Denisovan ancestries in Papuans. *Cell* 177, 1010–1021
32. Villanea, F.A. and Schraiber, J.G. (2019) Multiple episodes of interbreeding between Neanderthal and modern humans. *Nat. Ecol. Evol.* 3, 39–44
33. Juric, I. *et al.* (2016) The strength of selection against Neanderthal introgression. *PLoS Genet.* 12, 1006340
34. Harris, K. and Nielsen, R. (2016) The genetic cost of Neanderthal introgression. *Genetics* 203, 881–891
35. Green, R.E. *et al.* (2010) A draft sequence of the Neandertal genome. *Science* 328, 710–722
36. Reich, D. *et al.* (2010) Genetic history of an archaic hominin group from Denisova Cave in Siberia. *Nature* 468, 1053–1060
37. Meyer, M. *et al.* (2012) A high-coverage genome sequence from an archaic Denisovan individual. *Science* 338, 222–226
38. Mendez, F.L. *et al.* (2012) Haplotype at STAT2 introgressed from Neanderthals and serves as a candidate of positive selection in Papua New Guinea. *Am. J. Hum. Genet.* 91, 265–274
39. Huerta-Sánchez, E. *et al.* (2014) Altitude adaptation in Tibetans caused by introgression of Denisovan-like DNA. *Nature* 512, 194–197
40. Mendez, F.L. *et al.* (2012) Global genetic variation at OAS1 provides evidence of archaic admixture in Melanesian populations. *Mol. Biol. Evol.* 29, 1513–1520
41. Sankararaman, S. *et al.* (2014) The genomic landscape of Neanderthal ancestry in present-day humans. *Nature* 507, 354–357
42. Vernot, B. and Akey, J.M. (2014) Resurrecting surviving Neandertal lineages from modern human genomes. *Science* 343, 1017–1021
43. Dannemann, M. *et al.* (2016) Introgression of Neandertal- and Denisovan-like haplotypes contributes to adaptive variation in human Toll-like receptors. *Am. J. Hum. Genet.* 98, 22–33
44. Enard, D. and Petrov, D.A. (2018) Evidence that RNA viruses drove adaptive introgression between Neanderthals and modern humans. *Cell* 175, 360–371
45. Key, F.M. *et al.* (2014) Advantageous diversity maintained by balancing selection in humans. *Curr. Opin. Genet. Dev.* 29, 45–51
46. Schrider, D.R. and Kern, A.D. (2017) Soft sweeps are the dominant mode of adaptation in the human genome. *Mol. Biol. Evol.* 34, 1863–1877
47. Schrider, D.R. and Kern, A.D. (2016) S/HIC: robust identification of soft and hard sweeps using machine learning. *PLoS Genet.* 12, e1005928

48. Field, Y. *et al.* (2016) Detection of human adaptation during the past 2000 years. *Science* 354, 760–764
49. Sabeti, P.C. *et al.* (2007) Genome-wide detection and characterization of positive selection in human populations. *Nature* 449, 913–918
50. Li, H. *et al.* (2008) Ethnic related selection for an ADH class I variant within East Asia. *PLoS One* 3, e1881
51. Voight, B.F. *et al.* (2006) A map of recent positive selection in the human genome. *PLoS Biol.* 4, e072
52. Hernandez, R.D. *et al.* (2011) Classic selective sweeps were rare in recent human evolution. *Science* 331, 920–924
53. Enard, D. *et al.* (2014) Genome-wide signals of positive selection in human evolution. *Genome Res.* 24, 885–895
54. Key, F.M. *et al.* (2014) Selection on a variant associated with improved viral clearance drives local, adaptive pseudogenization of interferon lambda 4 (IFNL4). *PLoS Genet.* 10, 1004681
55. McManus, K.F. *et al.* (2017) Population genetic analysis of the DARC locus (Duffy) reveals adaptation from standing variation associated with malaria resistance in humans. *PLoS Genet.* 13, 48–65
56. Mathieson, S. and Mathieson, I. (2018) FADS1 and the timing of human adaptation to agriculture. *Mol. Biol. Evol.* 35, 2957–2970
57. Adhikari, K. *et al.* (2019) A GWAS in Latin Americans highlights the convergent evolution of lighter skin pigmentation in Eurasia. *Nat. Commun.* 10, 358
58. Laval, G. *et al.* (2019) A genome-wide approximate Bayesian computation approach suggests only limited numbers of soft sweeps in humans over the last 100,000 years. *bioRxiv* Published online December 23, 2019. https://doi.org/10.1101/2019.12.22.886234
59. Pybus, M. *et al.* (2015) Hierarchical boosting: a machine-learning framework to detect and classify hard selective sweeps in human populations. *Bioinformatics* 31, 3946–3952
60. Jensen, J.D. (2014) On the unfounded enthusiasm for soft selective sweeps. *Nat. Commun.* 5, 1–10
61. Liang, M. and Nielsen, R. (2014) The lengths of admixture tracts. *Genetics* 197, 953–967
62. Yang, M.A. *et al.* (2012) Ancient structure in Africa unlikely to explain Neanderthal and non-African genetic similarity. *Mol. Biol. Evol.* 29, 2075–2081
63. Racimo, F. *et al.* (2017) Archaic adaptive introgression in TBX15/WARS2. *Mol. Biol. Evol.* 34, 509–524
64. Berg, J.J. and Coop, G. (2014) A population genetic signal of polygenic adaptation. *PLoS Genet.* 10, e1004412
65. Berg, J.J. *et al.* Polygenic adaptation has impacted multiple anthropometric traits. bioRxiv. Published online November 6, 2017. https://doi.org/10.1101/167551
66. Daub, J.T. *et al.* (2015) Inference of evolutionary forces acting on human biological pathways. *Genome Biol. Evol.* 7, 1546–1558
67. Daub, J.T. *et al.* (2013) Evidence for polygenic adaptation to pathogens in the human genome. *Mol. Biol. Evol.* 30, 1544–1558
68. Gouy, A. *et al.* (2017) Detecting gene subnetworks under selection in biological pathways. *Nucleic Acids Res.* 45, e149
69. Gouy, A. and Excoffier, L. (2020) Polygenic patterns of adaptive introgression in modern humans are mainly shaped by response to pathogens. *Mol. Biol. Evol.* Published online January 14, 2020. https://doi.org/10.1101/732958
70. White, L. *et al.* (2015) Genetic adaptation to levels of dietary selenium in recent human history. *Mol. Biol. Evol.* 32, 1507–1518
71. Racimo, F. *et al.* (2018) Detecting polygenic adaptation in admixture graphs. *Genetics* 208, 1565–1584
72. Le Corre, V. and Kremer, A. (2012) The genetic differentiation at quantitative trait loci under local adaptation. *Mol. Ecol.* 21, 1548–1566
73. Chevin, L.-M. and Hospital, F. (2008) Selective sweep at a quantitative trait locus in the presence of background genetic variation. *Genetics* 180, 1645–1660
74. Cruickshank, T.E. and Hahn, M.W. (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol. Ecol.* 23, 3133–3157
75. Kim, B.Y. *et al.* (2018) Deleterious variation shapes the genomic landscape of introgression. *PLoS Genet.* 14, e1007741
76. Zhang, X. *et al.* (2020) Recessive deleterious variation has a limited impact on signals of adaptive 1 introgression in human populations. *bioRxiv* Published online January 14, 2020. https://doi.org/10.1101/2020.01.13.905174
77. Günther, T. and Coop, G. (2013) Robust identification of local adaptation from allele frequencies. *Genetics* 195, 205–220
78. Yu, F. *et al.* (2009) Detecting natural selection by empirical comparison to random regions of the genome. *Hum. Mol. Genet.* 18, 4853–4867
79. Garud, N.R. *et al.* (2015) Recent selective sweeps in North American *Drosophila melanogaster* show signatures of soft sweeps. *PLoS Genet.* 11, e1005004
80. Ferrer-Admetlla, A. *et al.* (2014) On detecting incomplete soft or hard selective sweeps using haplotype structure. *Mol. Biol. Evol.* 31, 1275–1291
81. Librado, P. and Orlando, L. (2018) Detecting signatures of positive selection along defined branches of a population tree using LSD. *Mol. Biol. Evol.* 35, 1520–1535
82. Crawford, J.E. *et al.* (2017) Natural selection on genes related to cardiovascular health in high-altitude adapted Andeans. *Am. J. Hum. Genet.* 101, 752–767
83. Schmidt, J.M. *et al.* (2019) Evidence that viruses, particularly SIV, drove genetic adaptation in natural populations of eastern chimpanzees. *bioRxiv* Published online March 21, 2019. https://doi.org/10.1101/582411
84. Yassin, A. *et al.* (2016) Recurrent specialization on a toxic fruit in an island *Drosophila* population. *Proc. Natl. Acad. Sci. U. S. A.* 113, 4771–4776
85. Duforet-Frebourg, N. *et al.* (2014) Genome scans for detecting footprints of local adaptation using a Bayesian factor model. *Mol. Biol. Evol.* 31, 2483–2495
86. Sverrisdóttir, O.Ó. *et al.* (2014) Direct estimates of natural selection in Iberia indicate calcium absorption was not the only driver of lactase persistence in Europe. *Mol. Biol. Evol.* 31, 975–983
87. Lindo, J. *et al.* (2016) A time transect of exomes from a Native American population before and after European contact. *Nat. Commun.* 7, 14
88. Racimo, F. *et al.* (2017) Signatures of archaic adaptive introgression in present-day human populations. *Mol. Biol. Evol.* 34, 296–317
89. Torada, L. *et al.* (2019) ImaGene: a convolutional neural network to quantify natural selection from genomic data. *BMC Bioinformatics* 20, 337
90. Sugden, L.A. *et al.* (2018) Localization of adaptive variants in human genomes using averaged one-dependence estimation. *Nat. Commun.* 9, 703
91. Hejase, H.A. *et al.* (2019) From summary statistics to gene trees: methods for inferring positive selection. *Trends Genet.* 4, 243–258
92. Hubisz, M. and Siepel, A. (2020) Inference of ancestral recombination graphs using ARGweaver. *Methods Mol. Biol.* 2090, 231–266
93. Rasmussen, M.D. *et al.* (2014) Genome-wide inference of ancestral recombination graphs. *PLoS Genet.* 10, e1004342
94. Kelleher, J. *et al.* (2019) Inferring whole-genome histories in large population datasets. *Nat. Genet.* 51, 1330–1338
95. Speidel, L. *et al.* (2019) A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet.* 51, 1321–1329
96. Berg, J.J. *et al.* (2019) Reduced signal for polygenic adaptation of height in UK Biobank. *eLife* 8, e39725
97. Sohail, M. *et al.* (2019) Polygenic adaptation on height is overestimated due to uncorrected stratification in genome-wide association studies. *eLife* 8, e39702
98. Daub, J.T. *et al.* (2017) Detection of pathways affected by positive selection in primate lineages ancestral to humans. *Mol. Biol. Evol.* 34, 1391–1402
99. Refoyo-Martínez, A. *et al.* (2019) Identifying loci under positive selection in complex population histories. *Genome Res.* 29, 1506–1520
100. Simonson, T.S. *et al.* (2010) Genetic evidence for high-altitude adaptation in Tibet. *Science* 329, 72–75
101. Foll, M. *et al.* (2014) Widespread signals of convergent adaptation to high altitude in Asia and America. *Am. J. Hum. Genet.* 95, 394–407

102. Lamason, R.L. *et al.* (2005) Genetics: SLC24A5, a putative cation exchanger, affects pigmentation in zebrafish and humans. *Science* 310, 1782–1786
103. Kamberov, Y.G. *et al.* (2013) Modeling recent human evolution in mice by expression of a selected EDAR variant. *Cell* 152, 691–702
104. Downes, D.J. *et al.* (2019) An integrated platform to systematically identify causal variants and genes for polygenic human traits. *bioRxiv* Published online October 24, 2019. https://doi.org/10.1101/813618
105. Kilpinen, H. *et al.* (2017) Common genetic variation drives molecular heterogeneity in human iPSCs. *Nature* 546, 370–375
106. Hwang, J.W. *et al.* (2019) iPSC-derived cancer organoids recapitulate genomic and phenotypic alterations of c-met-mutated hereditary kidney cancer. *bioRxiv* Published online January 11, 2019. https://doi.org/10.1101/518456
107. Kronenberg, Z.N. *et al.* (2018) High-resolution comparative analysis of great ape genomes. *Science* 360, eaar6343
108. Minster, R.L. *et al.* (2016) A thrifty variant in CREBRF strongly influences body mass index in Samoans. *Nat. Genet.* 48, 1049–1054
109. Clemente, F.J. *et al.* (2014) A selective sweep on a deleterious mutation in CPT1A in Arctic populations. *Am. J. Hum. Genet.* 95, 584–589
110. Wang, M.H. *et al.* (2012) Contribution of higher risk genes and European admixture to Crohn's disease in African Americans. *Inflamm. Bowel Dis.* 18, 2277–2287
111. Kwiatkowski, D.P. (2005) How malaria has affected the human genome and what human genetics can teach us about malaria. *Am. J. Hum. Genet.* 77, 171–192
112. Genovese, G. *et al.* (2010) Association of trypanolytic ApoL1 variants with kidney disease in African-Americans. 329, 841–845
113. Zhao, K. *et al.* (2012) Evidence for selection at HIV host susceptibility genes in a West Central African human population. *BMC Evol. Biol.* 12, 237
114. Ilardo, M.A. *et al.* (2018) Physiological and genetic adaptations to diving in sea nomads. *Cell* 173, 569–580
115. Ségurel, L. and Bon, C. (2017) On the evolution of lactase persistence in humans. *Annu. Rev. Genomics Hum. Genet.* 8, 297–319
116. Wolfe, N.D. *et al.* (2007) Origins of major human infectious diseases. *Nature* 447, 279–283
117. Engelken, J. *et al.* (2016) Signatures of evolutionary adaptation in quantitative trait loci influencing trace element homeostasis in liver. *Mol. Biol. Evol.* 33, 738–754
118. Jarvis, J.P. *et al.* (2012) Patterns of ancestry, signatures of natural selection, and genetic association with stature in Western African pygmies. *PLoS Genet.* 8, e1002641
119. Gokhman, D. *et al.* (2017) Inferring past environments from ancient epigenomes. *Mol. Biol. Evol.* 34, 2429–2438
120. Fagny, M. *et al.* (2015) The epigenomic landscape of African rainforest hunter-gatherers and farmers. *Nat. Commun.* 6, 10047
121. Fay, J.C. and Wu, C.-I. (2000) Hitchhiking under positive Darwinian selection. *Genetics* 155, 1405–1413
122. Grossman, S.R. *et al.* (2010) A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science* 327, 883–886
123. Sabeti, P.C. *et al.* (2002) Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419, 832–837
124. Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123, 585–595
125. Temme, S. *et al.* (2013) A novel family of human leukocyte antigen class II receptors may have its origin in archaic human species. *J. Biol. Chem.* 289, 639–653
126. Ding, Q. *et al.* (2014) Neanderthal introgression at chromosome 3p21.31 was under positive natural selection in east Asians. *Mol. Biol. Evol.* 31, 683–695
127. Ding, Q. *et al.* (2014) Neanderthal origin of the haplotypes carrying the functional variant Val92Met in the MC1R in modern humans. *Mol. Biol. Evol.* 31, 1994–2003
128. SIGMA Type, 2 Diabetes Consortium (2014) Sequence variants in SLC16A11 are a common risk factor for type 2 diabetes in Mexico. *Nature* 506, 97–101
129. Deschamps, M. *et al.* (2016) Genomic signatures of selective pressures and introgression from archaic hominins at human innate immunity genes. *Am. J. Hum. Genet.* 98, 5–21
130. Mendez, F.L. *et al.* (2013) Neandertal origin of genetic variation at the cluster of OAS immunity genes. *Mol. Biol. Evol.* 30, 798–801
131. Abi-Rached, L. *et al.* (2011) The shaping of modern human immune systems by multiregional admixture with archaic humans. *Science* 334, 89–94
132. Teixeira, J.C. and Cooper, A. (2019) Using hominin introgression to trace modern human dispersals. *Proc. Natl. Acad. Sci. U. S. A.* 116, 15327–15332
133. Hubisz, M.J. *et al.* (2019) Mapping gene flow between ancient hominins through demography-aware inference of the ancestral recombination graph. *bioRxiv* Published online June 30, 2019. https://doi.org/10.1101/687368