African genetic diversity and adaptation inform a precision medicine agenda

Luisa Pereira[1,2], *Leon Mutesa*[3], *Paulina Tindana*[4] *and Michèle Ramsay*[5]✉

Abstract | The deep evolutionary history of African populations, since the emergence of modern humans more than 300,000 years ago, has resulted in high genetic diversity and considerable population structure. Selected genetic variants have increased in frequency due to environmental adaptation, but recent exposures to novel pathogens and changes in lifestyle render some of them with properties leading to present health liabilities. The unique discoverability potential from African genomic studies promises invaluable contributions to understanding the genomic and molecular basis of health and disease. Globally, African populations are understudied, and precision medicine approaches are largely based on data from European and Asian-ancestry populations, which limits the transferability of findings to the continent of Africa. Africa needs innovative precision medicine solutions based on African data that use knowledge and implementation strategies aligned to its climatic, cultural, economic and genomic diversity.

Advances in DNA sequencing technologies and digital innovations are primers for the introduction of precision medicine approaches and apply also to resource-limited settings in African countries[1,2]. They have led to affordable diagnostic assays for pathogens and inform testing for inherited diseases. Although infectious diseases such as malaria, tuberculosis and HIV continue to dominate morbidity and mortality in sub-Saharan Africa (SSA), non-communicable diseases (NCDs) are on the rise[3,4]. NCDs are characterized by genetic risk variants in the presence of environmental triggers and modulators that together lead to a physiological state exceeding a threshold towards developing disease.

In Africa, the most recent centuries have been characterized by extreme changes in the environment, including exposures, urbanization, diet and behaviour, and this has led to shifts in allele frequencies in many human populations, with important consequences for disease susceptibility. We believe that precision medicine approaches have transformative potential based on a clear understanding of genetic susceptibility and gene–environment interactions and can contribute towards developing more effective treatment approaches in African settings[5,6].

In this Review, we reflect on the emergence of modern humans in Africa and the impact of extensive genetic diversity on disease susceptibility and survival. Our focus is on human genetic diversity, and we do not cover pathogen genomic diversity or transcriptome-based and proteome-based diagnostic approaches. The description of key characteristics of African genome architecture provides a lead into a vision for precision medicine that will require deep knowledge of the causal mutations for monogenic diseases and an understanding of the heritability, genetic contributions and environmental risk factors for complex multifactorial diseases and traits. Despite lagging behind in terms of data, laboratory infrastructure and health system access, key enablers for precision medicine are being developed in Africa[1,7,8].

African genetic diversity

Archaeologists and geneticists are in agreement that the *Homo sapiens* lineage of anatomically modern humans (AMHs) originated in Africa between 500 and 300 thousand years ago (kya)[9–11]. The location of the AMH cradle within the continent[11–14] has been a topic of much debate, with contradicting evidence pointing independently towards all major African regions, from North to West, Central, East and South Africa (FIG. 1a). A 'Pan-African' or 'African multiregionalism' model is therefore emerging for the origin of our species[12]. This model proposes that subdivided populations living across Africa and connected by sporadic gene flow contributed to the origin of our species. The model also allows for the occurrence of hybridization between *H. sapiens* and various divergent hominins within Africa. Phantoms of these events are being explored in extant African genomes (reviewed in[15]), in a similar way to preliminary evidence of hybridization in European and East Asian genomes

[1]*i3S — Instituto de Investigação e Inovação em Saúde, Universidade do Porto, Porto, Portugal.*

[2]*Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP), Universidade do Porto, Porto, Portugal.*

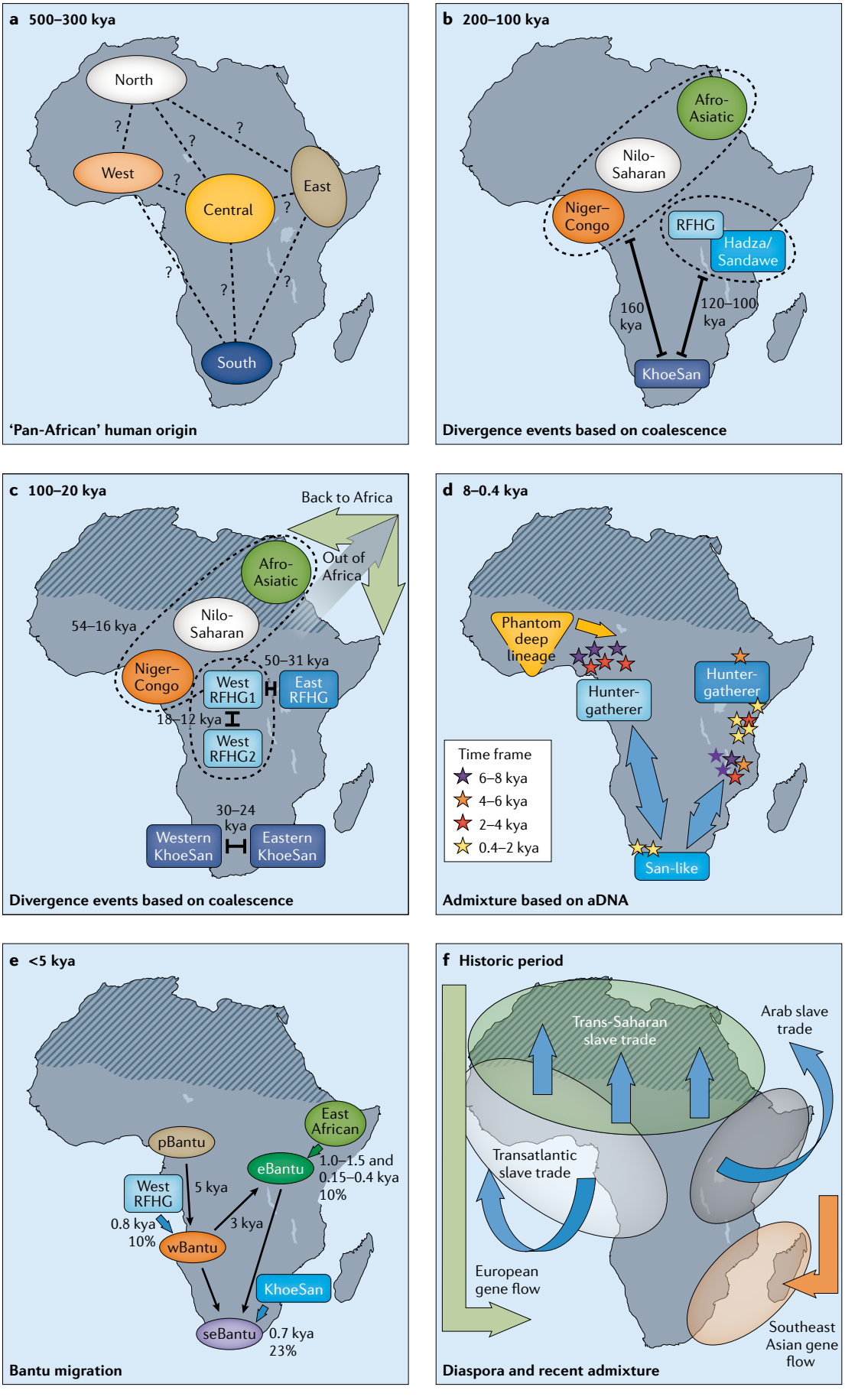
[3]*Center for Human Genetics, College of Medicine and Health Sciences, University of Rwanda, Kigali, Rwanda.*

[4]*School of Public Health, College of Health Sciences, University of Ghana, Accra, Ghana.*

[5]*Sydney Brenner Institute for Molecular Biosciencee, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.*

✉*e-mail: michele.ramsay@wits.ac.za*

https://doi.org/10.1038/s41576-020-00306-8



Precision medicine
An approach to clinical practice that includes information from state-of-the-art technologies to understand the underlying causes of a disease such that a patient can receive the most appropriate therapeutic intervention for the best possible health outcome.

Heritability
An estimation of the degree of variation in a phenotype that is due to genetic variation between individuals. Heritability can be estimated from general pedigrees using linear mixed models and from genomic relatedness estimated from genetic markers, but traditionally was based on twin studies.

Fig. 1 | **Important demographic events on the African continent. a** | Archaeological and whole-genome data[12] indicate that the origin of modern humans results from the contribution of heterogeneous groups inhabiting different parts of the continent and exchanging genes between themselves and with archaic groups, from 500 to 300 thousand years ago (kya), a model known as the 'Pan-African' human origin. **b**,**c** | Coalescence methods, based on a single-group human origin, point to several divergence events occurring between African groups (hunter-gatherers shown in rectangles and non-hunter-gatherers in circles) in the past 200 kya[21]. In the time interval 200–100 kya (part **b**), the following major events occurred: a split ~160 kya leading to KhoeSan, Niger-Congo, Afro-Asiatic and Niger–Congo Saharan lineages; and the later divergence between San and other hunter-gatherers (rainforest hunter-gatherers (RFHG), Hadza and Sandawe) by ~120–100 kya. In the time frame 100–20 kya (part **c**), Niger–Congo, Nilo-Saharan and Afro-Asiatic lineages diverged from each other in the 54–16 kya interval; within the Central African rainforest lineages, eastern and western groups diverged by ~50–31 kya and the two western groups separated by ~18–12 kya; within KhoeSan, western and eastern groups diverged between 30 and 24 kya; the out-of-Africa migration took place around 70 kya; and North African (under the striped shadow) population origins derive largely from a back-to-Africa migration around 45 kya, from the Near East/Arabia, following an interruption of settlement of the region by modern humans. **d** | Admixture based on evidence thus far collected from ancient DNA (aDNA) (oldest ages ~8 kya; stars indicate analysed individuals[29–31]) indicates contacts between the diverse hunter-gatherer groups in the west, east and south of the continent (blue arrows), as well as admixture with multiple deep lineages in the west (represented by the yellow triangle and arrow). **e** | Bantu migration beginning 5 kya consisted of substantial movements of people, from its origin (parental Bantu (pBantu)), across most of the sub-Saharan African geography (movement represented by black arrows), with episodes of local admixtures (indicated by wide arrows) with hunter-gatherers as west RFHG and KhoeSan (in rectangles) and with East African populations (in ovals). Western Bantu-speakers (wBantu) showed admixture of their parental population (~70% contribution; best proxy Yoruba) with RFHG at around 0.8 kya (~10% contribution), whereas eastern Bantu-speakers (eBantu) displayed two admixture events (at 1.0–1.5 kya and 0.15–0.4 kya) of their parental population (~75% contribution; best proxy Angola) with an Afro-Asiatic speaking population from Ethiopia (~10% contribution), and, in southeast Bantu-speakers (seBantu), the parental population (~70% contribution; best proxy Angola with some admixture from eBantu) mixed with the Ju|'hoansi San from Namibia (~23% contribution) at 0.7 kya[27]. **f** | Historic diaspora out of the continent (represented by blue arrows[42]) and admixture with non-African ancestries (the European admixture[40] in green; the admixture with Southeast Asians[37] in orange).

Admixture
Two populations come into contact with one another, often due to migration, and interbreeding occurs, generating a hybrid or admixed descendant population.

Coalescent models
Models that, assuming that variants sampled from a population may have originated from a common ancestor, reconstruct backwards in time how variants can be shown to originate from a single ancestor according to a random process in coalescence events. The model produces many theoretical genealogies that can be compared with the observed data to test assumptions about the demographic history of a population.

Population structure
Significant differences in allele frequencies between populations or between subpopulations in a population.

(reviewed in[16]). It has been estimated that people of West African descent have 2–19% genomic introgression from an archaic population (an older splitter in the human lineage than Neanderthals)[17], whereas the Neanderthal introgression in African populations is thought to have occurred mainly indirectly via European admixture and back migration[18]. This 'Pan-African' model for the origin of modern humans proposes that the roots of our African origins are based on a much higher level of diversity at earlier times than previously imagined.

Structure, migration and admixture. The analysis of increasingly available African genomes coupled with refined coalescent models is being used to infer ancient divergence events between African populations[19,20]. The results suggest that African populations began to display population structure ~200 kya. In a recent study using whole-genome sequences[21], African ancestry was confirmed to be largely partitioned by geography and language (FIG. 1b,c), as previously suggested[22]. North Africa has a probable discontinuity of AMH settlement, justifying its separation from SSA in genetic terms and warranting a separate review[23,24].

African populations display admixture signals in their genomes, indicating that populations are linked by gene flow due to both short-range and long-range migration and admixture events[25–28]. In Europe and America, ancient DNA (aDNA) is currently providing direct insights into Palaeolithic and Mesolithic movements[16], but in SSA the few available aDNA genomes date only from 0.4 to 8 kya[29–31] (FIG. 1d). Even so, the aDNA genomes from hunter-gatherers from current-day Malawi, Tanzania and South Africa[30] showed that in the past a population with a San-like genomic component had a wider distribution with its impact declining towards the north, and with a cline between East Africa and southern Africa. aDNA from four skeletal remains from Shum Laka (Cameroon) dated 3 and 8 kya[31] indirectly testified to widespread admixture in Africa between multiple deep lineages (one of which was an early-diverging ancestry source in West Africa), and that central African hunter-gatherers diverged from other African populations as early as the split of southern African hunter-gatherers.

Genomics on extant populations using admixture dating methods based on linkage disequilibrium (LD) decay provide a time window for admixture events taking place up to 4.5 kya[32]. The time frame matches the 2–5 kya interval of the agricultural Bantu migration (FIG. 1e), and has provided insights into admixture with local groups encountered along the migration path[27,33,34]. These studies hint at the extensive replacement and assimilation of populations living throughout most of SSA by Bantu lineages, with the few surviving hunter-gatherer communities such as the Hadza, Sandawe, Khoe and San providing a glimpse of the former greater African diversity.

The semi-arid Sahel Belt (including South Sudan, Ethiopia and Somalia) escaped a major Bantu influence. Genomic studies conducted there[26,35] indicate its main role as a corridor for bidirectional migrations within Africa, and Eurasian input from the eastern region and in the surviving nomadic pastoral communities such as the Tuareg, Fulani and Daza. The Eurasian ancestry is mainly Arabic/Near Eastern-related and major admixture events took place at least twice: 2.4–3.2 kya, possibly related to the spread of Afro-Asiatic languages[36]; and 0.15–1.5 kya[35], most probably related to the expansion of Islamism[37]. The Comoros[37], Madagascar[38] and the Swahili people (coastal populations from Kenya to north Mozambique)[39] make up the cosmopolitan cultural and trading zone known as the Swahili Corridor that established the Indian Ocean trading network at the dawn of the second millennium (FIG. 1f).

The colonial and postcolonial periods added a new layer of admixture with non-African ancestry, which is epitomized in the South African Coloured group (a recognized multi-ethnic population) with up to five-way admixtures including African (Khoe, San and Bantu-speakers), European (mainly central and northern) and South Asian ancestries[40,41]. People with African ancestry moved beyond the continent in the historic period via waves of forced enslavement. The transatlantic slave trade, extending from the sixteenth to the nineteenth centuries, was responsible for the movement of at least 12 million people of African descent to the New World[42]. They came mainly from the western African coast[27] and there is great inter-individual

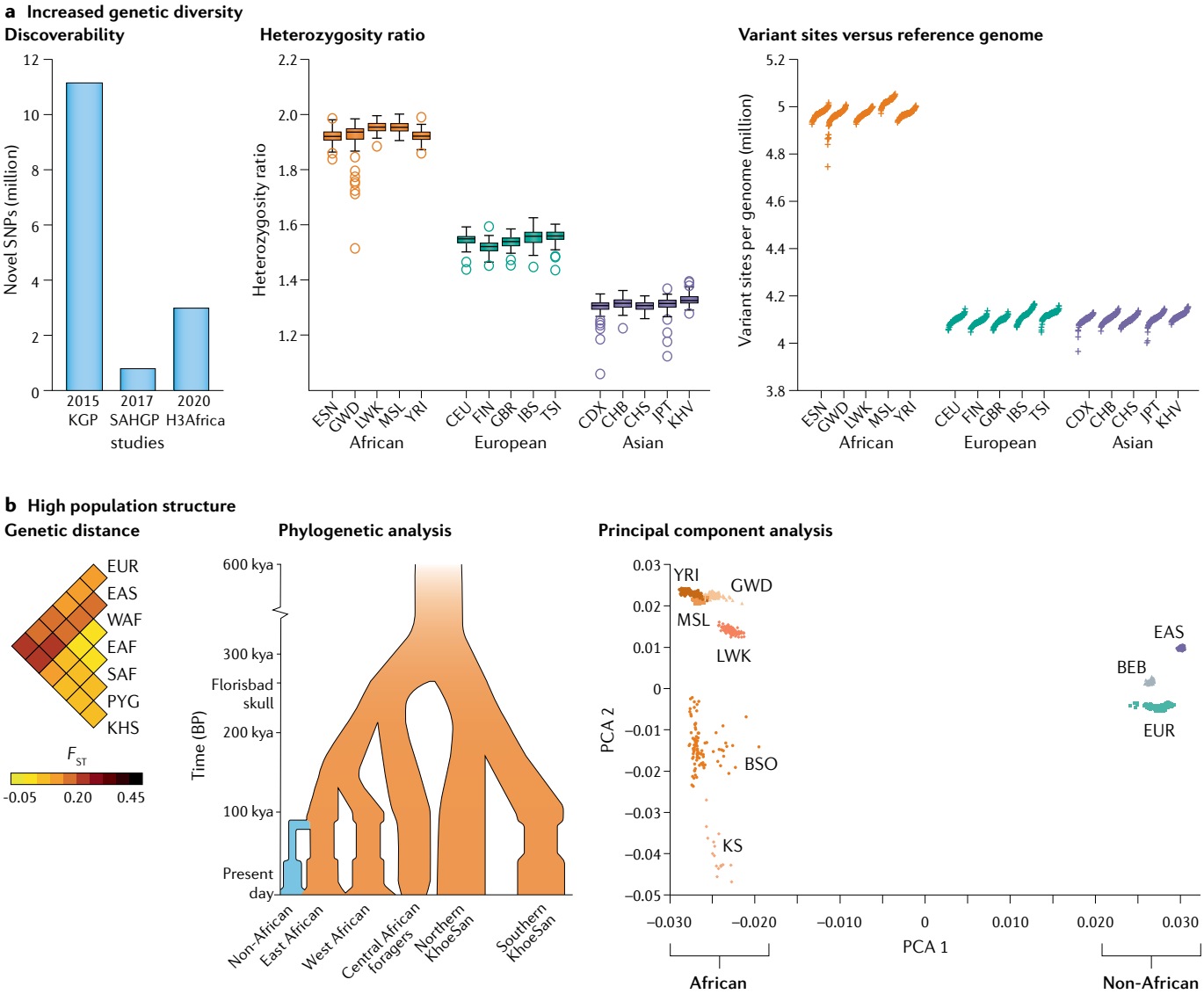

a Increased genetic diversity
Discoverability
Novel SNPs (million)
2015 KGP
2017 SAHGP
2020 H3Africa studies
Heterozygosity ratio
Heterozygosity ratio
ESN GWD LWK MSL YRI
African
CEU FIN GBR IBS TSI
European
CDX CHB CHS JPT KHV
Asian
Variant sites versus reference genome
Variant sites per genome (million)
b High population structure
Genetic distance
EUR
EAS
WAF
EAF
SAF
PYG
KHS
F_{ST}
-0.05 0.20 0.45
Phylogenetic analysis
600 kya
300 kya
Florisbad skull
200 kya
100 kya
Present day
Time (BP)
Non-African
East African
West African
Central African foragers
Northern KhoeSan
Southern KhoeSan
Principal component analysis
YRI
GWD
MSL
LWK
BSO
KS
EAS
BEB
EUR
PCA 2
PCA 1
African
Non-African
c Lower LD and impact on GWAS and PRS
LD decay
EAS
EUR
SSA
r^2
Distance (kb)
Distribution of allele frequencies
GWAS
KGP African aggregate
% of SNPs
Frequency
PRS based on European data
African
East Asian
European
R^2
Basophil dbp sbp mchc hb ht Neutrophil wbc bmi Lymphocyte Monocyte Eosinophil rbc mch Height mcv plt
Phenotype

Fig. 2 | **Features of African genome architecture.** The study of African genomes has important advantages. **a** | Increased genetic diversity leads to high discoverability of new variants (in 2020, the Human Heredity and Health in Africa (H3Africa) paper[34] from 426 whole-genome sequences revealed more than 3.4 million novel variants, despite the number of African variants already in databases from previous studies) (left panel). This high variation is also reflected in the heterozygosity ratio (number of heterozygous sites in an individual divided by the number of non-reference homozygous sites), which is highest in African populations, followed by European and Asian populations[192] (middle panel). Lastly, the number of variants that differ from the reference sequence is also much higher in African populations[45] (right panel). **b** | Population structure is another important feature of African populations. The considerable genetic differences (pairwise F_{ST}; that is, the proportion of genetic differentiation between populations) between African and European or East Asian populations and across African populations that give rise to population substructure enable studies on the origins of specific variants (left panel) (data from[193]). In a demographic model of African history based on coalescence analyses of ancient and modern complete genomes from sub-Saharan Africa (SSA), it was possible to infer that the emergence of *Homo sapiens* occurred between 350,000 and 260,000 years ago, highlighting the remarkable low population size (represented by the width of branches) of the non-African group in relation to all other SSA groups[10] (middle panel). Population structure is evident from the principal component analysis (PCA), which shows African populations and non-African populations separate in the first principal component and the African populations showing within-continental structure in the second principal component (right panel) (data from[41,45,193]). **c** | Due to the generally lower linkage disequilibrium (LD) (data inferred for 10,000 polymorphic sites from the 1000 Genomes Project (KGP)[45]) and smaller haplotype blocks in African populations (left panel), there is an opportunity to narrow down causal variants more effectively. In genome-wide association study (GWAS) projects, allele frequencies of disease-associated variants often differ between the GWAS Catalogue (predominantly European) and African populations (KGP AFR), showing that there is an abundance of low allele frequencies in the African populations, in part because the GWAS genotyping arrays are based on common variants in non-African populations (middle panel). Transferability of polygenic risk scores (PRSs) to African populations is also affected by the discovery populations used in the GWAS that was used to develop the PRS, with improved transferability to a target population if the discovery data set is from a similar population, in this case a European population[117]. Transferability was lowest to African populations (right panel). BEB, Bangladeshi; bmi, body mass index; BP, before present; BSO, Black South African; CDX, Dai Chinese; CEU, Utah residents with Northern and Western European Ancestry; CHB, Han Chinese; CHS, Southern Han Chinese; dbp, diastolic blood pressure; EAF, east sub-Saharan Africa; EAS, East Asian; ESN, Esan in Nigeria; EUR, European; FIN, Finnish; GBR, British; GWD, Gambian; hb, haemoglobin; ht, haematocrit; IBS, Iberian; JPT, Japanese; kb, kilobase; KHS, KhoeSan; KHV, Vietnamese; KS, KhoeSan; kya, thousand years ago; LWK, Luhya in Kenya; mch, mean corpuscular hemoglobin; mchc, mean corpuscular hemoglobin concentration; mcv, mean corpuscular volume; MSL, Mende in Sierra Leone; plt, platelet count; PYG, rainforest hunter-gatherers; rbc, red blood cell count; SAF, south sub-Saharan Africa; SAHGP, Southern African Human Genome Programme; sbp, systolic blood pressure; SNP, single-nucleotide polymorphism; TSI, Italian; WAF, west sub-Saharan Africa; wbc, white blood cell count; YRI, Yoruba in Nigeria. Part **a** (middle panel) adapted with permission of Genetics Society of America, from [Heterozygosity ratio, a robust global genomic measure of autozygosity and its association with height and disease risk., Samuels, D.C. et al., 204, 2020]; permission conveyed through Copyright Clearance Center, Inc. Part **a** (right panel) and part **c** (left panel) adapted from REF.[45], Springer Nature Limited. Part **b** (middle panel) adapted with permission from REF.[10], AAAS. Part **c** (right panel) adapted from REF.[117], Springer Nature Limited.

variation in the proportion of African ancestry among the large African diaspora in the USA, the Caribbean, Latin America, the UK and other countries and regions. The limited representation of African genomic diversity among the diaspora raises concerns about the extrapolations made from genomic studies in African Americans (AAs) to the entire African continent (FIG. 1f).

African genomic architecture and advantages for identifying disease-causing variants. Genome architecture has important health-related implications arising from the deep and highly structured African genetic diversity (FIG. 2). The extreme bottleneck, ~70 kya, of the out-of-Africa migration, with as few as 1,000 people of East African descent giving rise to all non-African populations (first estimates based on mitochondrial DNA diversity[43] and later supported by whole-genome data[44]), renders African populations the most diverse in the world.

The 1000 Genomes Project (KGP) sequenced five SSA populations (all from the Niger–Congo language family) and two admixed African-ancestry populations, showing that an African genome has, on average, 20% more non-reference single-nucleotide polymorphisms (SNPs) than Europeans and East Asians (4.31 million, 3.53 million and 3.55 million, respectively)[45]. The Southern African Human Genome Programme pilot study, which sequenced 16 black south-eastern Bantu-speakers, identified ~800,000 novel variants[40], and the Human Heredity and Health in Africa (H3Africa) population study of 426 whole-genome sequences from individuals representing 50 African ethnolinguistic groups detected >3 million novel variants[34]. Further efforts to capture more of the diversity across the continent, including whole-genome sequence data from previously understudied African populations, are urgently needed to enrich the catalogue of human variants.

The deep evolutionary roots of SSA populations have resulted in considerably lower LD, compared with European and East Asian populations[45], resulting in each common variant having typically more than 15–20 tagging variants ($r^2 > 0.8$) in non-African populations, but only about 8 in SSA populations. This renders the design of informative genome-wide genotyping arrays for African populations more challenging, as they require more variants and enrichment of African SNPs to adequately cover all of the haplotype blocks. For this reason, the H3Africa Genotyping Chip was designed as a customized genotyping array enriched for common variants in African populations for African studies. On the positive side, fine mapping and the identification of causative variants are more informative in African populations, as haplotype blocks are generally shorter[46,47]. According to the Genome-Wide Association Study (GWAS) Catalogue, only about 2.4% of participants in large studies are of African origin (predominantly AA), but they account for 7% of the disease-related genetic associations[48], illustrating the enhanced discovery potential when studying African populations.

Human populations had a recent explosive growth phase, leading to an excess of rare genetic variants[49]. Information gathered from 15,336 genes re-sequenced in 6,515 individuals of European American (EA) and AA ancestry allowed Fu et al.[49] to infer that 73% of all protein variants and 86% of predicted-deleterious variants amongst 1 million total variants arose in the past 5–10 kya. There are contradictory reports about the burden of deleterious variants in African populations compared with non-African populations and the impact this may have on detecting genetic susceptibility for disease[50].

When studying complex traits in African populations, it is useful to examine the role of environmental adaptation that occurred through a process of natural selection. There are several modes of selection (reviewed in[51]),

Linkage disequilibrium (LD). A measure of the non-random association of alleles at different loci in a given population. LD is lower in African populations compared with European and Asian populations, given the more ancient ancestry of African populations.

for example, a classic or hard sweep occurs due to strong selection on a new/rare mutation, rapidly increasing its frequency, whereas a soft sweep takes place when a previously neutral mutation becomes advantageous due to an environmental change. Polygenic selection acts simultaneously on multiple pre-existing variants, with changes in minor allele frequencies at multiple loci resulting in a more advantageous phenotype. Balancing selection, on the other hand, leads to the preservation of deleterious alleles in populations where the driving agent, for example malaria or trypanosomiasis, is acting due to heterozygote advantage (higher fitness of heterozygotes compared with homozygotes) or frequency-dependent selection (the fitness of a phenotype is dependent on the relative frequency of other phenotypes in the population). Selection can also occur in association with admixture, in which a neutral variant in group A is selected for in the recipient group B and increases in frequency. If group A is an archaic population, this is referred to as adaptive introgression, but if a modern human population, then as adaptive admixture.

Examples of adaptation in African genomes. Africa follows a north–south axis, spanning approximately 70° of latitude, in contrast to the west–east axis of Europe and continental Asia. This explains the enormous diversity of climates and natural environments in Africa and the heterogeneous environmental pressures, one of the strong drivers of selection, across the African continent. It has been demonstrated that infectious agents and food-related factors are the main drivers of selection in humans[52], with both having a direct impact on health and survival, and being instrumental in ongoing humanitarian crises across the African continent. Countries located across the equator display the highest general species and pathogen diversities[53], which matches the African malaria distribution[54], meningitis belt[55] and Ebola outbreaks[56]. In addition, environmental diversity is correlated with high diversification in available food sources, with some African communities still practising hunter-gatherer subsistence whereas others are farmers and pastoralists, with an increasing gap from traditional ways of life to westernized megacities. Recent increased urbanization is having a profound impact on the health transition, especially in more developed African countries.

Descriptions of classic sweeps for lactose tolerance, skin colour, short stature, arsenic tolerance and adaptation to arctic and high-altitude conditions abound elsewhere (for example, reviewed in[57]) and are not specifically included in this Review although they have important relevance to Africa. We focus on other examples of adaptation identified through GWAS approaches, admixture mapping, positive selection and adaptive introgression (FIG. 3). We also recognize that it is often difficult to accurately determine possible causes and related effects in these studies.

A good example to illustrate the complexity of natural selection is the protection against dengue fever in populations with African ancestry, through variants associated with the *OSBPL10* and *RXRA* genes that act in a pathway that controls lipid metabolism in hepatocytes and macrophages[58]. This association was found in the admixed Cuban population, as dengue outbreaks in Africa are either scarce (affecting mostly non-African residents) or misdiagnosed as other haemorrhagic fevers. Dengue viruses need a lipid raft to enter cells and depend on the availability of lipids to build new capsules. The authors presented functional evidence that the knockdown of *OSBPL10*, matching the naturally low expression of this gene in African populations, leads to decreased replication of the dengue virus. This association may have been driven by a related flavivirus, having a higher impact on fitness (such as the yellow fever virus) but also protecting against dengue and zika viruses. African populations have a distinctive lipid profile from European populations[59], usually considered healthier, with lower triglycerides (TGs) and higher high-density lipoprotein (HDL) cholesterol concentrations. A GWAS in AAs[60] identified loci significantly associated with HDL cholesterol (*CD36*, *PPP1R3B*, *LPL*, *CETP*, *LOC55908*, *APOA*/*APOC* gene cluster and intergenic locus on 21q22), low-density lipoprotein (LDL) cholesterol (*PCSK9*, *APOB*, *ABCG8*, *APOE* and *LDLR*) and TGs (*LPL*, *APOA*/*APOC*, *APOC1* and *GCKR*). Interestingly, an independent study[61] showed that AAs have common nonsense mutations in the *PCSK9* gene that are rare in EAs, and are associated with a 40% reduction in plasma levels of LDL, leading to a novel intervention to treat dyslipidaemia using *PCSK9* inhibitors.

Unexpectedly, given their healthier lipid profile, African populations have a high risk for cardiometabolic diseases, suggesting the need for caution in using the same risk cut-off values for serum lipids as biomarkers for these diseases across ethnicities[59]. Serum lipids are strongly correlated with fat distribution, particularly with visceral adipose tissue that contributes to increased liver fat content. African populations are less prone to deposit fat in the visceral adipose tissue, accumulating it in less metabolically active regions, such as the subcutaneous tissue and in the lower extremities. The fat deposition in the visceral adipose tissue is best correlated to the waist circumference, and several loci have been associated with the waist to hip ratio in African populations (*DCST2*, *PPARG*, *ADAMTS9*, *SNX10*, *KLF13*, *CMIP* and *PEMT*, after meta-analysis with EAs)[62]. The genetic architecture of obesity and type 2 diabetes in African populations was recently reviewed, emphasizing these differences among ancestries[63].

Higher prevalence of hypertension in African populations[7,64] may be a major contributor to risk for cardiometabolic diseases, and Young et al.[65] advanced the hypothesis that susceptibility to hypertension is ancestral and due to latitudinal heat adaptation. Supporting this claim, they found that latitude and a variant in *GNB3* account for a major portion of worldwide variation in blood pressure. According to this hypothesis, modern populations from hot environments are more susceptible to hypertension. A large GWAS meta-analysis of AAs on blood pressure, including 29,378 individuals from 19 discovery cohorts and subsequent replication in additional samples of AA ($n = 10,386$), EA ($n = 69,395$) and East Asian ($n = 19,601$) ancestry, identified association with five loci (*EVX1–HOXA*, *ULK4*, *RSPO3*, *PLEKHG1*

Bantu migration
A massive migration of Bantu-speaking peoples that began 5,000 years ago in the region of Cameroon/Nigeria towards the southern and eastern parts of the African continent, with genetic, linguistic and cultural impacts. Currently, Bantu-speakers make up ~30% of the African population of ~1.3 billion people.

African diaspora
People of African origin or ancestry resident in non-African countries.

Human Heredity and Health in Africa
(H3Africa). A pan-African consortium that aims to study the genomic and environmental determinants of common and rare diseases with the goal of improving the health of African populations.

Genome-wide association study
(GWAS). A study that aims to identify candidate genetic markers associated with diseases or traits when applied to case–control cohorts or to quantitative traits within a cohort. This is done by genotyping several million common single-nucleotide polymorphisms from across the genome and applying statistical analyses to determine the probability of association between individual genetic markers and the phenotype.

Adaptation
A response to an environmental challenge such that an advantageous phenotype is enriched by positive or balancing selection.

Balancing selection
Multiple alleles at a locus are maintained in the population gene pool at higher frequencies than expected from genetic drift. Two possible causes are heterozygote advantage (higher fitness of heterozygotes compared with homozygotes) and frequency-dependent selection (fitness of a phenotype depends on the relative frequency of other phenotypes in the population).

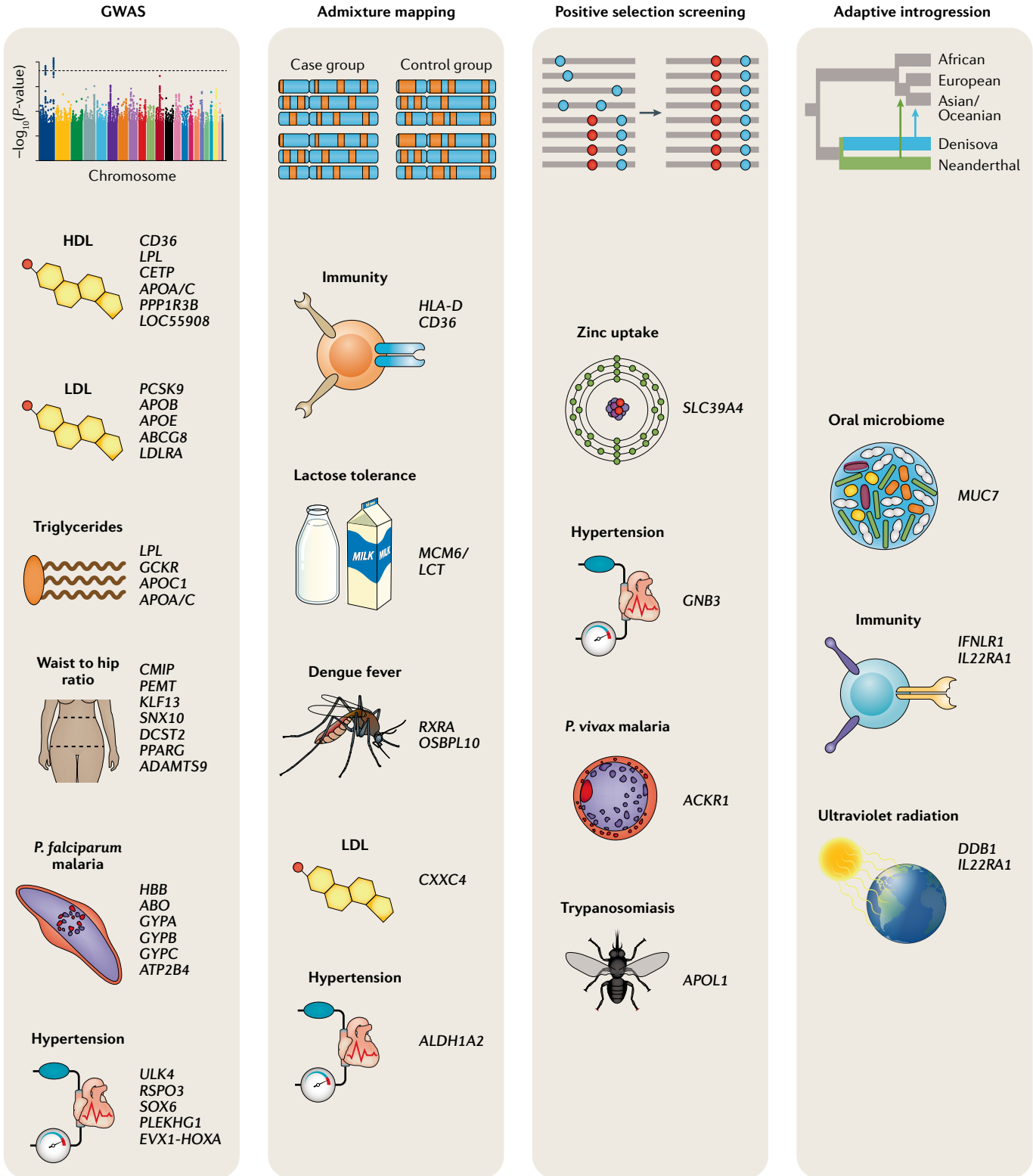


Fig. 3 | **Adaptive genetic variants in people with African ancestry have been discovered through several approaches.** Genome-wide association study (GWAS) investigations conducted in cohorts of African ancestry have identified several candidate genes associated with specific traits, such as lipid metabolism[60], waist to hip ratio[62], *Plasmodium falciparum* malaria infection[67] and hypertension[66]. Three other methods pinpointed gene variants that confer protection against various challenges, some of which evolved in localized geographic regions. They were acquired by admixture from neighbouring populations (immunity and lactose tolerance[27], dengue fever[58], low-density lipoprotein (LDL) levels and hypertension[80]), identified by screening for positive selection (such as zinc uptake[77], hypertension[65], *Plasmodium vivax* malaria[68] and trypanosomiasis[71]) or are proposed to have been derived through adaptive introgression from archaic humans (oral microbiome[82], immunity and ultraviolet radiation[18]). HDL, high-density lipoprotein.

and *SOX6*)[66] in this trans-ethnic meta-analysis after correction for multiple testing. Pathway analysis of these genes indicated relevant insights, including into nitric oxide signalling, which influences vasodilation, endothelial function and cardiac contraction. Interestingly, associated loci in EA individuals also showed significant effects in AA individuals, indicating that a largely common set of genes regulate blood pressure across human populations.

Some genetic variants conferring natural resistance to malaria are classic examples of balancing selection, where selected variants in the heterozygous state confer relative resistance to *Plasmodium* species, but in the homozygous state predispose individuals to specific severe genetic disorders, such as sickle cell disease. A GWAS[67] performed for life-threatening *Plasmodium falciparum* infection (severe malaria) in more than 11,000 African children, with a replication data set of 14,000 individuals, replicated signals at *HBB*, *ABO* and *ATP2B4*. The study also revealed a new candidate cluster of genes (*GYPE*, *GYPB* and *GYPA*) encoding glycophorins that are receptors for erythrocyte invasion by *P. falciparum*. A particular haplotype at this locus provided 33% protection against severe malaria and was linked to polymorphisms that have features of ancient balancing selection. In relation to malaria caused by *Plasmodium vivax* infection, the selective sweep on a variant (rs2814778) in the atypical chemokine receptor 1 (*ACKR1*) gene, previously known as *DARC* (Duffy antigen receptor for chemokines), is extreme[68]. The encoded antigen acts as a receptor for *P. vivax*, and red blood cells of homozygotes for the null allele (Duffy O blood group) are resistant to invasion. *P. vivax* malaria is rare in autochthonous African people, as this variant has gone to near fixation in almost all SSA populations (except in Khoe and San[69]), and even populations of Arab-descent in Sudan have enrichment of this variant through admixture with local populations[26], and the Malagasy through admixture with the SSA populations[70].

As another example, the deadly human African trypanosomiasis caused by two tsetse fly-transmitted African trypanosomes, *Trypanosoma brucei rhodesiense* (responsible for the acute East African form) and *Trypanosoma brucei gambiense* (causing the more chronic West African form), has been proposed as the motor of positive selective pressure for variants in the *APOL1* gene[71]. Two *APOL1* haplotypes, known as G1 and G2[71], show signs of selection within the last 10,000 years and vary in frequency across African populations. Individuals with the G2 haplotype are naturally protected against the East African parasite, with five times lower risk of infection due to the trypanolytic action triggered by the presence of this variant[72]. When the West African parasite infects individuals with the G1 and G2 haplotypes, there is a differential response such that individuals with the G1 haplotype have a lower parasite load and less severe disease. Interestingly, these variants have in recent years been associated with higher risk for kidney disease in African-ancestry populations, and especially with HIV-associated nephropathy (HIVAN)[73]. The *APOL1* variants therefore have either positive or negative consequences depending on environmental exposures.

Due to the high prevalence of several haematological variants under strong selection from exposure to malaria, it is possible to study their epistatic interactions in African populations. These interactions can have a positive outcome, as in the case of the co-occurrence of α-thalassaemia and β-thalassaemia. Alternatively, they can give rise to negative epistatic effects, such as the combination of α-thalassaemia and the sickle cell trait that together no longer confer protection against malaria[74]. Another negative epistatic effect in African populations is for increased haemolytic risk and sickle cell nephropathy in individuals with both the sickle cell mutation and *APOL1* alleles (G1 and G2)[75]. Furthermore, children with both sickle cell anaemia and *APOL1* alleles are more likely to develop albuminuria early during childhood as a result of glomerular injury[76]. On the other hand, individuals with both the sickle cell mutation and the *BCL11A* rs1427407 T variant that increases γ-globin expression have a positive epistatic effect of reduced sickle globin polymerization[75].

A less well known example of pathogen resistance is a non-synonymous substitution (rs1871534) in *SLC39A4*[77], characterized by extreme differences in allele frequencies between West African populations and Eurasian populations, with near fixation of the derived allele in people of West African descent. The gene encodes the human intestinal zinc uptake transporter ZIP4, and the authors demonstrated that the variant reduces zinc intracellular uptake, hypothesizing that this may starve certain pathogens of zinc. This illustrates potential 'nutritional immunity' by which the human host restricts access by pathogens to certain micronutrients, so that they become less virulent[78]. However, deficiencies in zinc and selenium are important risk factors in oesophageal cancer[79], but the regions with the highest frequency of the rs1871534-derived allele (around Nigeria, Cameroon and Central African Republic) do not match the eastern African oesophageal cancer corridor, suggesting alternative risk factors and the need for further investigation.

Adaptation can also occur by selection for advantageous variants introduced into a population through admixture. Examples have been found of both more recent (last 4,000 years) and ancient admixture events. As expected, the Bantu expansion was an excellent opportunity for the acquisition of local adaptations from resident groups by the newcomer farmers[27]. Western Bantu-speakers acquired immune response-related *HLA-D* and *P. falciparum* malaria-related *CD36* diversity from rainforest hunter-gatherers, whereas eastern Bantu-speakers acquired *LCT*-related selection on lactase persistence from people of Eurasian-admixed eastern African descent. In the reverse direction, Khoe and San acquired selected genomic regions from the pastoralists and farmers with whom they mixed[80]. Ju|'hoan acquired a region containing variation in the *CXXC4* gene that has been associated with LDL cholesterol levels from eastern African populations and Eurasian populations, whereas Nama have enrichment of Eurasian diversity of the *DCC* gene that has been associated with body weight changes and alcoholism, and of the *ALDH1A2*

Admixture mapping
A gene-mapping algorithm applied to case–control cohorts in a recently admixed population, where there are differences in the rates of the disease or trait between the two parental populations and those differences are partly due to differences in the frequencies of associated or causal variants. If the associated variant in one ancestry is protective, it will be enriched in the control group; if it is causative, it will be enriched in the case group.

Positive selection
A phenotype (and its associated alleles) confers an advantage to the individual in response to a challenge (for example, environmental), such that the variant alleles that confer the favourable phenotype rapidly increase in frequency in the population, sometimes attaining fixation.

Autochthonous African people
Indigenous or native African people.

Epistatic interactions
Interactions pertaining to epistasis, when specific combinations of multiple genetic variants at different loci have non-additive effects on a specific phenotype (for example, disease or trait).

gene that has been associated with hypertension and blood pressure in AAs[80].

Despite the absence of genomic data for archaic groups in Africa, screening for adaptive introgression in extant African genomes has been taking place. Some authors interpreted the pronounced depletion of archaic sequences in genic regions from African hunter-gatherers as suggesting that they were maintained under neutrality[81], although the failure of current methods to properly detect archaic admixture and also limited archaic admixture are possible explanations. Later studies[82] found that African diversity of the salivary expressed gene *MUC7* seemed to have been enriched by archaic introgression, which may have an impact on the oral microbiome. However, in a study published in 2020[17], *MUC7* was not amongst the archaic segments with selected elevated frequencies in people of West African descent, which included genes involved in tumour suppression (*NF1*), mitochondrial aerobic respiration in the testis (*MTFR2*), hormone regulation (*HSD17B2*), potassium channels (*KCNIP4*) and trichorhinophalangeal syndrome (*TRPS1*).

Surprisingly, the recent implementation of a new algorithm that does not assume absence of Neanderthal ancestry in African populations[18] revealed that African individuals carry a stronger signal of Neanderthal admixture (~17 Mb of Neanderthal sequence per African individual) than previously suggested. This input is thought to be due to genuine Neanderthal ancestry due to migrations of Eurasian populations back to Africa, and gene flow into Neanderthals from an early dispersing group of humans out of Africa. The authors found shared African–European high-frequency Neanderthal haplotypes, one of which contained the *TRIM55* gene, and 13 African-specific high-frequency Neanderthal haplotypes, enriched for genes involved in immunological function (such as *IL22RA1* and *IFNLR1*) and ultraviolet-radiation sensitivity (*DDB1* and *IL22RA1*).

Translational medicine in Africa

Globally, data used to inform precision medicine approaches are largely skewed towards European populations[83–86], making them the prime beneficiaries of more accurate diagnoses, improved genetic risk prediction for complex traits and targeted treatment approaches, leading to widening global health disparities[84,87]. Increasing evidence now supports the contribution of African genetic diversity to understanding differences in disease association and causation, opening up opportunities for the discovery of novel biological pathways to inform therapeutic interventions[87]. However, the relatively limited research in African populations highlights the urgent need for more data across the ethnically diverse populations on the continent.

Monogenic disorders and precision medicine. Over the past decade, next-generation sequencing (NGS) has exponentially increased our understanding of monogenic diseases by enabling the discovery of causal mutations in many families[88]. This technology is uncovering many variants of uncertain significance (VUS) and has raised ethical dilemmas in decision-making about feeding back both VUS and incidental findings that may, or may not, be clinically actionable, given our current understanding[89]. Decision algorithms have been developed for the classification of VUS, but rely heavily on data from non-African populations, and are frequently revised as more information and, consequently, better understanding comes to light. Assessing allele frequencies in African populations for variants that are indicated as pathogenic in ClinVar has led to the reclassification of some variants as non-pathogenic[90]. Nonetheless, African populations remain under-represented in population-based genomic studies.

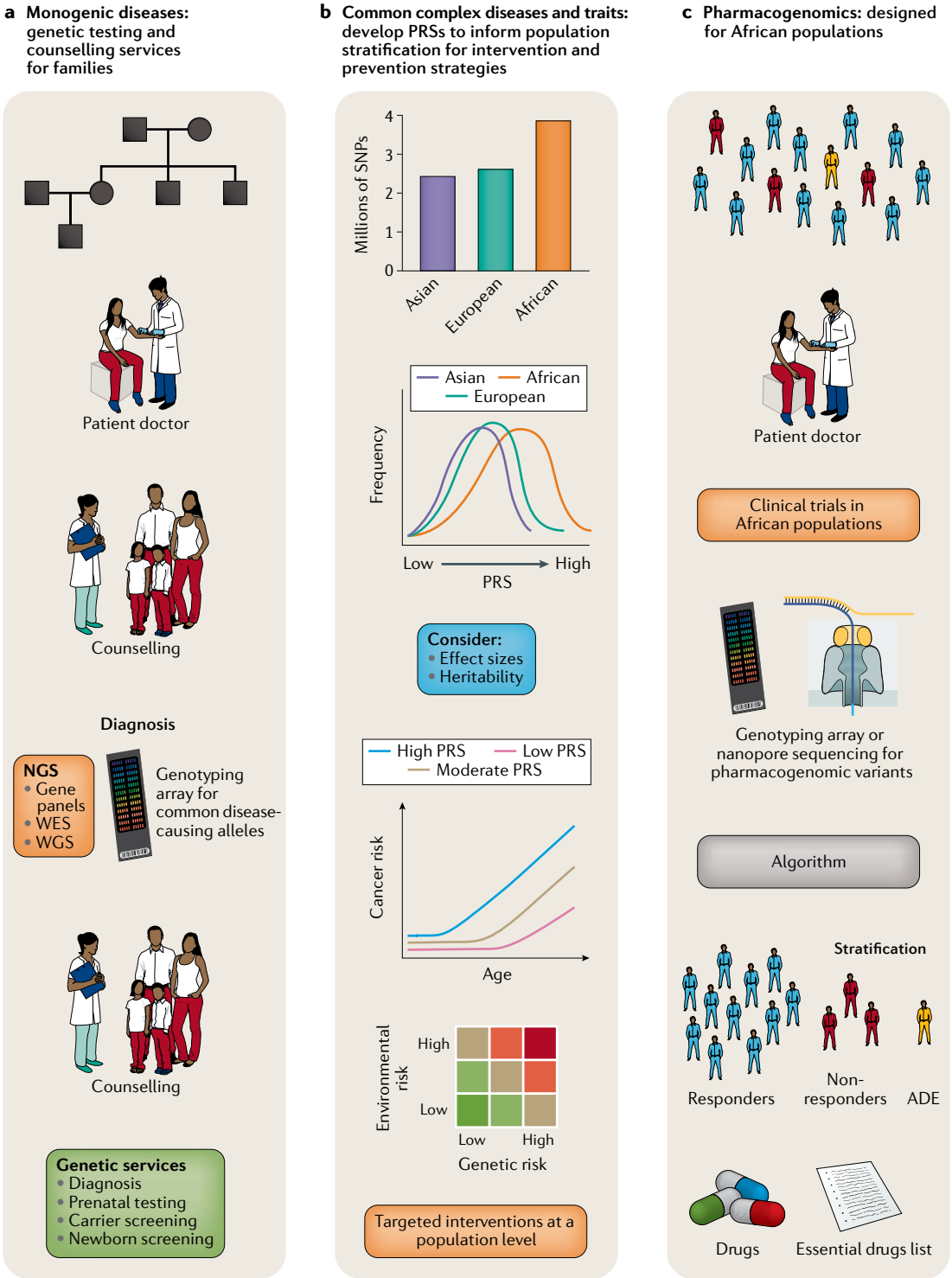
Currently, a limited number of African clinical facilities perform individual variant genotyping for well-characterized diseases such as sickle cell anaemia[91], and when NGS technologies are used, they mostly target specific genes or gene panels with genes previously found to harbour disease-causing mutations. These approaches limit diagnostic utility and the ability to detect novel causal mutations and genes in African families and populations. Below, we describe findings primarily from SSA populations, but we recognize that much work and clinical implementation have been done in North African countries with populations that have high rates of consanguinity (recently reviewed in[92]).

Allelic heterogeneity (that is, different mutations in the same gene) has been described for specific monogenic diseases among African populations, and mutation profiles differ from those in European populations, both in terms of the nature of the mutations and their frequency. Several monogenic disease-causing mutations of African origin have been discovered in African diaspora patients and explored further in Africa. Examples include the *CFTR* 3120+1G>A variant that is the most common causal variant in African cystic fibrosis patients, the *GALT* S135L (c.404C>T) variant for galactosaemia, the *GBA* p.T36del (c.222_224delTAC) Rec*Nci*I variant for Gaucher disease, the *GCDH* A293T (c.877G>A) variant for glutaric aciduria, the *FKBP10* c.831dupC variant for osteogenesis imperfecta type 3 and the *AABCC6* R134W (c.3940C>T) variant for pseudoxanthoma elasticum (reviewed in[93]).

Locus heterogeneity (that is, mutations in different genes that cause the same or a very similar diseases) have also been described in African populations (reviewed in[93]). There are four well-studied examples. More than 90% of Fanconi anaemia cases in Black South African populations are caused by a 7-bp deletion (c.637_643del) in the *FANCG* gene, whereas other pathogenic mutations in this gene are responsible for only ~10% of cases in European populations. In European populations, the majority of Fanconi anaemia mutations occur in the *FANCA* and *FANCC* genes. Huntington disease in African populations can be caused by triplet expansion mutations in either the *HTT* or the *JPH3* genes in roughly 67% and 33% of cases, respectively, whereas more than 99% of European patients with Huntington disease have the expansion mutation in the *HTT* gene. These differences have important implications for diagnosis and the identification of pre-symptomatic individuals in affected families, as well as prenatal testing. Monogenic non-syndromic hearing loss shows great

Variants of uncertain significance
(VUS). Genetic variants in coding regions or regulatory regions of known disease-associated genes with insufficient evidence to assess their potential functional or phenotypic impact.

Incidental findings
Genetic variants of potential disease relevance that are unrelated to the condition under investigation. For example, searching for a mutation responsible for developmental delay in a child and then detecting a *BRCA1* breast cancer susceptibility variant.



locus heterogeneity (>90 genes implicated) in all populations. Mutations in *GJB2* and *GJB6* predominate in European and Asian populations, but are rare or absent in African populations[94–96]. A notable exception is a founder mutation (*GJB2*-R143W) that is common in Ghana, accounting, in the homozygous state, for a quarter of familial cases and ~8% of cases with non-familial non-syndromic congenital hearing impairment in specific communities[97–99]. Lastly, the spectrum of loci and mutations that cause oculocutaneous albinism vary between African populations and European populations, with people of African descent predominantly having causal mutations in the *OCA2* gene, with only a few cases described with mutations in the *TYR* gene[93,100].

The abundance of uncharacterized monogenic disorders in Africa due to poor diagnostic and infrastructural capacity slows down the development of clinical genetic services on the continent (FIG. 4a). There are several reasons why it is important to know the causal variant for a monogenic disorder in a family: it makes definitive

Fig. 4 | **Precision public health strategies could benefit African populations.** The aim of precision public health is to use precision medicine informed approaches to diagnose, treat and monitor patients for the greatest benefit to the majority of the population. This can be applied to monogenic diseases and common complex diseases such as diabetes, hypertension and cancer. However, the most practical example for precision public health is in the field of pharmacogenomics, to understand the role of genetic variation in drug efficacy and the reduction of adverse effects. Three approaches are outlined. **a** | Families with monogenic diseases have distinct pedigrees characterized by the distribution of affected individuals. This can help to identify the genes and detect mutations that cause diseases. With sufficient knowledge of key genes and mutations that are more common in African populations, screening arrays could provide an affordable tool as the first line in making a definitive diagnosis and doing population carrier screening for autosomal and X-linked disorders. This array would need to consider the incidence of babies born with specific disorders in African populations, and allelic and locus heterogeneity in African families. **b** | The global debate on the clinical utility of polygenic risk scores (PRSs) in stratifying populations has led to translational scenarios in some countries (for example, breast screening in the UK). In an African setting, an affordable assay for population stratification could lead to the identification of high-risk individuals (by considering both genetic and environmental risk) and developing targeted interventions. This would be more affordable than targeting the entire population for intervention. **c** | Adverse drug events (ADEs) are responsible for considerable morbidity and mortality. Identifying common genetic variants in African populations could guide drug dose and the use of appropriate drugs to ensure the best outcome. Essential drug lists would contain the most effective drugs to benefit the majority of the population and would be available at all hospitals. NGS, next-generation sequencing; SNP, single-nucleotide polymorphism; WES, whole-exome sequencing; WGS, whole-genome sequencing.

diagnosis possible, it provides at-risk couples with reproductive choices through prenatal or pre-implantation diagnosis and it may inform prognosis and treatment options. The potential acceptability of the implementation of several of these options in African communities is poorly explored, and socio-economic studies would need to include community and individual engagement processes to understand the impact of cultural and religious differences.

Common complex traits and disorders, and transferability of polygenic risk scores. Common complex and multifactorial diseases and traits, such as type 2 diabetes, hypertension, cancer and psychological disorders, are influenced by a multitude of genetic and environmental factors. Due to behavioural choices and environmental exposures, in the context of genetic susceptibility, many of these diseases are increasing in prevalence across the world. Understanding these complex interactions could lead to important interventions to improve health through precision medicine informed approaches but would need to be tailored to specific communities and individuals (FIG. 4b). This would require knowledge of the genetic background and cultural and socio-economic environments.

Genetic susceptibility to most common diseases is the cumulative effect of many genetic variants, each with a relatively small effect, usually explaining only a modest portion of the heritability of the disease. What has been referred to as the 'missing heritability' could be accounted for by several factors, including an overestimation of the heritability, an inability to detect rare variants with relatively large effects, a variation in genomic architecture across populations (for example, patterns of LD and population substructure), cryptic relatedness, sampling biases, underdeveloped statistical methods to detect epistasis and complex gene–environment interactions[101,102].

The section above on adaptation presents descriptions of associations that have led to allele frequency increases for protective and risk-related associations with lipid traits, hypertension, kidney disease, reduced zinc uptake and resistance to *P. falciparum* and *P. vivax*-related malaria. GWAS investigations in African populations have also made important contributions to understanding other traits such as type 2 diabetes mellitus[46,103], primary open-angle glaucoma[104], rheumatoid arthritis[105] and schizophrenia[106]. A GWAS for primary open-angle glaucoma in individuals with African ancestry (The Genetics of Glaucoma in People of African Descent (GGLAD) Consortium), with a discovery cohort of 2,320 cases and 2,121 control participants, revealed novel associations with the amyloid-β A4 precursor protein-binding family B member 2 (*APBB2*) locus, which had not previously shown association in people of European or Asian ancestry[104]. GWAS investigations for rheumatoid arthritis have also highlighted important novel genetic associations, in addition to a strongly replicating HLA association, where the latter has different lead SNPs in African populations[105].

To discover variants with small effects requires very large participant cohorts and rich phenotype data sets to gain a better understanding of the genetic contribution to complex traits[107]. An analysis of large GWAS projects has revealed poor representation of individuals with African ancestry (often represented by AAs) but a proportionally large number of significant associations[48], arguing for greater representation of African populations and other under-represented groups in GWAS investigations. The contributing factors to the increased discovery potential of even modest African data sets are attributed to increased genetic diversity, improved fine-mapping potential as a result of generally lower LD in African genomes, improved imputation accuracy of rarer alleles through the availability of more African whole-genome reference sequences and clearer definitions of phenotypes and potential confounders. However, it has been suggested that the GWAS threshold for African studies should be more stringent ($<5\times10^{-9}$) due to the increased number of haplotype blocks, thereby increasing the number of independent tests[86].

The largest published African GWAS meta-analysis to date for 34 complex traits included the Uganda Genome Resource (UGR) ($n=6,407$), the African American Diabetes Mellitus (AADM) study ($n=5,231$), the Durban Diabetes Study (DDS) ($n=1,165$) and the Diabetes Case Control Study (DCC) (n = 1,542), with a total 14,345 participants from Ghana, Kenya, Nigeria, South Africa and Uganda[86]. They found that whereas most known associations with data available in more than one cohort were transferable, there were several known and functionally important loci (for example, the *DARC* locus associated with monocyte count) with limited transferability among the African cohorts[86]. This could reflect allele frequency differences among cohorts (sometimes due to positive selection), allelic heterogeneity (multiple distinct variants at specific loci) or gene–environment interactions, as well as novel

association signals. Differences in LD structure around causal or lead variants across populations or allelic heterogeneity led to some strong evidence of statistical heterogeneity in regions around several lead SNPs within known and biologically important regions (for example, *LDLR* and *PCSK9*). The authors also identified different SNPs at the same gene associated with a given trait among different populations (for example, different population-specific variants in the *GPT* locus drive the association with liver function-related alanine aminotransferase levels in African and European populations). Possible explanations include different tag SNPs associated with an unidentified causal variant or multiple signals that are driving the associations in the different populations.

The H3Africa Consortium has pooled resources from several projects and developed a harmonized data set of ~55,000 individuals from 13 SSA countries, referred to as the Cardiovascular H3Africa Innovation Resource (CHAIR)[7]. In an analysis of body mass index (BMI) and hypertension, the CHAIR resource has revealed extensive regional and sex-specific variation patterns, which may reflect differences in genetic susceptibility, behaviours, environmental exposures and gene–environment interactions. GWAS will be performed on this cohort using the H3Africa genotyping array of ~2.3 million SNPs, enriched for common variants in African populations[7].

Cancer research has revealed that people of African descent often have different subtypes, an earlier age at onset and more severe cancers, suggesting a higher genetic predisposition to some cancers[108]. In addition, viral-linked cancers are also common among African populations. The high prevalence of Burkitt lymphoma in people of African descent helped to make the connection with Epstein–Barr virus as the causal factor[109] and HIV as the driver of the increase in Kaposi sarcoma and non-Hodgkin lymphoma[110]. The different patterns of cancer distribution and manifestation require different approaches to healthcare in Africa. For example, the triple-negative breast cancer subtype is more common among African women, rendering their cancers highly aggressive and unresponsive to treatments commonly used in other settings[111]. The profiles of *BRCA1* and *BRCA2* mutations have also been shown to differ between Nigerian and South African women[112,113], with few data from other SSA countries. Knowledge of allele frequencies in African populations led to the reclassification of a variant reported to increase cancer risk, from pathogenic to benign, when it was discovered that it is common in African populations[114]. As African populations continue to age, NCDs including cancer will place an increased burden on already stretched health services[115] and there is an urgent need for more data to plan appropriate interventions and treatment approaches.

Standardized methods that sum up the genomic risk burden in individuals, such as polygenic risk scores (PRSs), have been developed, but are of limited predictive value for complex traits such as neuropsychiatric disorders, diabetes, cardiovascular diseases and inflammatory disorders[116–118]. Their clinical utility is a topic of much debate and although many have argued that their predictive ability is useful in stratifying populations for surveillance, for example for breast cancer, others have emphasized that they have important limitations[119–124]. As alleles that confer susceptibility or protection are neither necessary nor sufficient to cause or protect against disease, their outcome is modified by known and unknown exposures. Many common causative or protective alleles are shared between populations and therefore contribute to the disease on a trans-continental level, as described for a large set of genes regulating blood pressure in AAs, EAs and East Asians[66]. However, contributing alleles that are either limited to African populations or of high frequency among African populations could only be detected when including African data. GWAS projects in Africa therefore require better representation of common and rare African variants in genotyping arrays, such as has been developed by the H3Africa Consortium. More genome sequencing efforts in Africa are needed to uncover the high differentiation between African populations, and genetic associations may reveal that PRS algorithms would need to be tailored for different African regions. Several comparative PRS studies have concluded that trans-ethnic transferability is limited, especially when the discovery and target populations are distantly related[48,117,118,125,126]. The factors that influence PRSs and limit transferability are shown in BOX 1 and the complex scenarios highlight the importance of using genomic data from multiple populations to develop appropriate population-specific applications.

Pharmacogenomics. Although personalized medicine based on genomic data is still limited in a clinical setting, the anticipated benefit of pharmacogenomics includes the ability to tailor drug choices to the biology of the individual patient, with more effective treatment and a reduction in adverse drug events[127,128] (FIG. 4c). As only between 25% and 80%[129] of patients respond adequately to specific drug treatments, targeted therapy informed by knowledge of genomic variants that alter the effects of drugs could improve the rate of successful treatment[130]. There is high heterogeneity between human populations[131] for variants associated with the function of genes involved in absorption, distribution, metabolism and excretion of drugs (referred to as ADME genes). With an increase in the availability of African whole-genome sequences it is now possible to use a bioinformatics approach to systematically characterize variants across African populations. Below are several examples of how the application of pharmacogenomics has impacted healthcare in Africa.

Efavirenz (EFV) has been a first-line drug for treating HIV infection but is associated with adverse effects that include depression, rashes and suicidal tendencies, and these are more commonly observed in African patients on a standard dose of 600 mg/day. Many people in Zimbabwe and Botswana have a variant (*CYP2B6* 516G>T) that increases the metabolism of EFV and leads to toxic doses when the standard dose is administered[132]. A solution to these adverse outcomes was to start treatment at 200 mg/day and monitor efficacy. This approach led to increased compliance and better viral control[133]. In Botswana, genomic studies showed that about 13.5%

Polygenic risk scores (PRSs). Predictive scores made up from multiple genetic loci associated with a trait and weighted by the relative contribution of each marker/allele to the trait that can be used to stratify a population on a spectrum of high to low risk. The higher the heritability of the trait, the more predictive the PRS will be; however, clinical use is still debatable.

Box 1 | **Transferability of polygenic risk scores to an African setting**

Polygenic risk scores (PRSs), also referred to as genetic risk scores (GRSs), have been developed to stratify individuals from a given population into low, moderate and high genetic risk categories[124,194]. The predictive value or clinical utility of a PRS[120–123,195] is dependent on the heritability of the trait (percentage of the causality due to genetic factors), the content of the genotyping array (number and choice of single-nucleotide polymorphisms (SNPs)) used in the genome-wide association study (GWAS) to provide the data for the PRS, the contribution of each associated genetic variant to the trait (effect size) and the joint contribution of many variants.

Current models for PRSs are relatively simple, taking into consideration the additive effects of individual genetic variants ($i = 1$–n SNPs), weighted according to the size of their effect on the phenotype (w) and according to the number of high-risk alleles per locus ($X = 0$, 1 or 2):

$$\mathrm{PRS} = \sum_{i=1}^{n} w_i . X_i$$

This model does not include the challenging considerations of epistasis (gene–gene interactions), pleiotropy (the effect of one variant on many different phenotypes) and gene–environment interactions on a trait.

Over a 10-year period (2008–2017), only 15 out of 733 (2%) PRS studies included individuals of African origin[118]. Several publications have addressed the potential trans-ethnic transferability of PRSs, with evidence suggesting that the predictive utility of a PRS developed in one population will diminish proportionally to the genetic distance from the intended target population[117,119,196–198]. Although some argue that the verdict is still out because we do not have enough data from African populations to test the hypothesis[118], there are important differences in the genomic architecture of diseases in people of African descent that challenge the transferability of PRSs developed in European populations to an African setting.

Genetic diversity and effect size

African populations have increased genetic diversity[34,40,45]. Not only are there more variants relative to the reference sequence but the allele frequencies differ, even among different African populations[199]. This affects both the effect sizes of variants in a population and their overall contribution to PRSs[196,198].

Genomic architecture

The lower average linkage disequilibrium (LD) across chromosomal regions of African genomes increases their use for fine mapping and discoverability of causal variants[62]. GWAS genotyping arrays, however, need to include more SNPs to cover all haplotype blocks, and LD-pruned content needs to be carefully chosen and based on data from African populations[66,187]. Sample sizes also need to be powered to detect small-effect variants that together could contribute to a large proportion of the heritability of a trait. A lack of transferability may be due to associated variants in LD with the causal variant in the discovery population, but being on a smaller untagged LD block in a target African population[47,63].

Novel associations

Although disease associations of many genetic loci replicate at the level of the genomic region, the lead associated SNPs for a particular trait often differ in African populations[103]. If a PRS is based only on the lead associated SNPs, they will be poor proxies for transferability[118]. There is evidence that African populations have unique and novel genetic associations for complex diseases, which would not be included in the calculation of a PRS based on European data, weakening their transferability[117].

Heritability estimates and regional differences across ethnicities

The narrow-sense or SNP heritability of a trait can be linked to specific gene–environment interactions, where, for example, a population with high food insecurity such as in Uganda would not reach its genetic potential for height, affecting the estimated heritability of height (70–80% heritability in European populations and <50% heritability in Uganda)[86]. Regional heterogeneity across African populations would exacerbate differences, potentially requiring different algorithms for PRSs across regions or along ethnolinguistic–cultural divides.

Considerations for interventional strategies in an African setting

The availability of infrastructure and resources will have a profound effect on the clinical utility of PRSs in low-resource settings, and leapfrog opportunities need to carefully consider national health service platforms[156,169,186].

of the population would be unable to effectively utilize EFV-based therapies, and this led to a change in the HIV management policy where they opted out of EFV-based therapies in favour of dolutegravir[134].

Tamoxifen has been used for the treatment of breast cancer in patients with oestrogen receptor-positive breast cancers and to prevent recurrent breast cancer[135]. *CYP2D6* encodes an enzyme responsible for metabolizing tamoxifen to endoxifen, the active agent. Genetic variants in the gene, including the African variant *CYP2D6*17*, reduce enzyme activity and lead to poor clinical outcome[136].

Multidrug-resistant tuberculosis also requires a precision medicine approach to ensure the most effective treatments. It is therefore important for research to be performed to characterize the effects of different variants and allele frequencies to assess pharmacogenomics applications[137] in different African populations[107,108].

Warfarin-dosing has been associated with genetic variants in people of African descent with a high rate of thromboembolic episodes, with one study identifying four SNPs on chromosome 6 that were associated with increased risk for major bleeding[138]. This led to the development of adapted pharmacogenetic warfarin dosing algorithms in Black African populations that take into consideration genetic ancestry and population diversity.

Epigenetics, adversity and health. Epigenetic modifications are known as somatically heritable changes to the genome that do not alter the DNA sequence but influence gene expression levels, and in some cases can be transmitted across generations. Although epigenetic changes are responsible for normal development and health, they can also cause disease. This emerging field may contribute to disease treatment and prevention through modulation of gene expression by changes in epigenetic signals, and although research in African populations remains scarce, there are some good examples[139].

Evidence shows that adversity in early life increases the risk of epigenetic changes associated with childhood mental health and lifelong risks of chronic disorders of health and well-being[140]. The highly divergent DNA methylation patterns between African and European populations have been shown to be due largely to a combination of differences in allele frequencies and complex epistasis or gene–environment interactions[141]. Recently, GWAS projects have identified methylation quantitative trait loci (meQTLs; that is, genetic variants associated with different levels of methylation at specific loci) both in *cis* and in *trans* that are linked to disease-associated genetic variants[142]. Interestingly, the application of meQTL GWAS data on SNPs and array-based gene expression has been demonstrated to detect common causal variants and shared biological mechanisms in the human genome[143]. Fagny et al.[144] generated genome-wide genotype and DNA methylation profiles for 362 rainforest hunter-gatherers and sedentary farmers revealing methylation patterns associated with recent changes in habitat, mostly related to immune and cellular functions. This methylation variation was strongly associated with

Pleiotropy
One variant that exerts an effect or has an association with multiple different (but sometimes related) phenotypes.

nearby genetic variants that are themselves enriched in signals of natural selection.

Gene–environment interactions cause risk of post-traumatic stress disorder (PTSD) in the post-conflict zones across the world. DNA methylation has shown differential patterns in people with PTSD, and research has revealed increased DNA methylation at the NGFI-A binding site of the *NR3C1* promoter that is associated with less intrusive memory of the traumatic event and reduced PTSD risk in men, but not women, in an African population[145]. More recent work showed a positive correlation between epigenetic signals in mothers exposed to the Rwandan genocide and their offspring, illustrating transgenerational effects in this African population[146,147].

DNA methylation is also influenced by nutritional status, and African studies have described epigenetic mechanisms that are associated with clinical outcomes related to malnutrition[148,149]. A study from The Gambia demonstrated persistent epigenetic changes (metastable epialleles) in the offspring of mothers with specific seasonally associated nutritional deficiencies during early pregnancy, examined in DNA extracted from lymphocytes and hair follicles[148,150]. Another study showed that oedematous severe acute malnutrition (ESAM), which affects many children under 5 years of age, is associated with DNA hypomethylation at several loci, as assessed in DNA from saliva tissue, in Jamaicans and Malawians, but that these alterations were not observed in adults who had recovered from severe acute malnutrition[149,151,152]. These findings highlight the potential role of methyl-group supplementation during early pregnancy and in children, to prevent poor developmental outcomes.

The disease status of the mother can also influence the epigenome of her offspring. A study in South Africa showed significant epigenetic changes in women with gestational diabetes and their offspring at genes involved with the pentose phosphate pathway, by examining gene expression levels in blood and placental tissue linked to altered DNA methylation patterns[153].

Epigenetic modifications have emerged as a potential treatment that can lead to biological regulation of gene expression in chronic diseases such as cancer, cardiovascular diseases, metabolic diseases and neuropsychiatric disorders[154], and merit further investigation in African populations.

Enablers of precision medicine in Africa

In Africa, clinical genetic services have improved over the past two decades in selected countries, with one catalyst being the hosting of an African Society of Human Genetics Congress and the founding of a local society. This has also been a trigger for strengthening human capacity-building and the development of genomics research to support a precision medicine approach[155–157].

Although precision medicine is likely to be expensive and targeted to small patient populations, leapfrog technology could provide affordable solutions based on population characteristics related to drug responses, common inherited diseases and cancer screening. There is growing interest in the potential of genomics to transform healthcare service delivery, including through precision medicine approaches that use genomic and other types of information in scientific research, and in disease prevention, diagnosis and treatment.

For the past decade there has been growing recognition that the ethical, legal and sociocultural implications of genomics and precision medicine ought to be given greater attention if Africa is to harness the potential benefits of these advances in science. Issues such as informed consent, privacy and confidentiality, benefit sharing, feedback of findings and equitable sharing of samples and data within the context of international collaborations in Africa have been well documented[158–161]. There are also growing efforts to build the right evidence base to support best practices through consultations with relevant stakeholders and communities, and conducting empirical ethics projects that seek to explore preferences and perspectives. Research consortia such as H3Africa[1], the Plasmodium Diversity Network in Africa (PDNA[162]) and a host of others are developing specific guidelines to support the implementation of human and pathogen genomics research in Africa. These guidelines have all highlighted the limitations of existing institutional, national and regional ethical and legal frameworks to govern genomics research in Africa and the need for a governance framework that not only promotes research collaborations but also ensures these are conducted with regard to fair and equitable sharing of samples and data[163]. The ultimate aim is to translate these resources and knowledge into improving the health of individuals and populations within Africa and the global community through precision medicine and public health interventions.

There is growing consensus that precision medicine has the potential to deliver health benefits to individuals, by improving the state of health of individuals and populations (FIGS 4 and 5). However, there are still a myriad of ethical and governance issues that need to be addressed, including: whether and how data should be shared with commercial entities; which consent model should govern sample and data sharing; which benefits should be shared, how and by whom; and how to ensure privacy and confidentiality of data. Furthermore, best practices for community engagement, genuine informed consent, return of results, ownership of genetic information and who should be involved in decisions about data and sample sharing remain subjects of much debate.

Collaboration, public and community engagement. Community engagement has gained recognition as an important requirement for genomics research, and both its intrinsic and instrumental values have been highlighted as key to the success of genomics research[164,165]. The development of precision medicine will require the collaboration of key stakeholders including health authorities, patients and their families. The need to strengthen patient involvement in decisions in clinical practice is a recurring theme in precision medicine[166] and would require innovative ways of strengthening scientific literacy to ensure that patients are able to make informed decisions on how their personal data should be handled and who gets access to them.

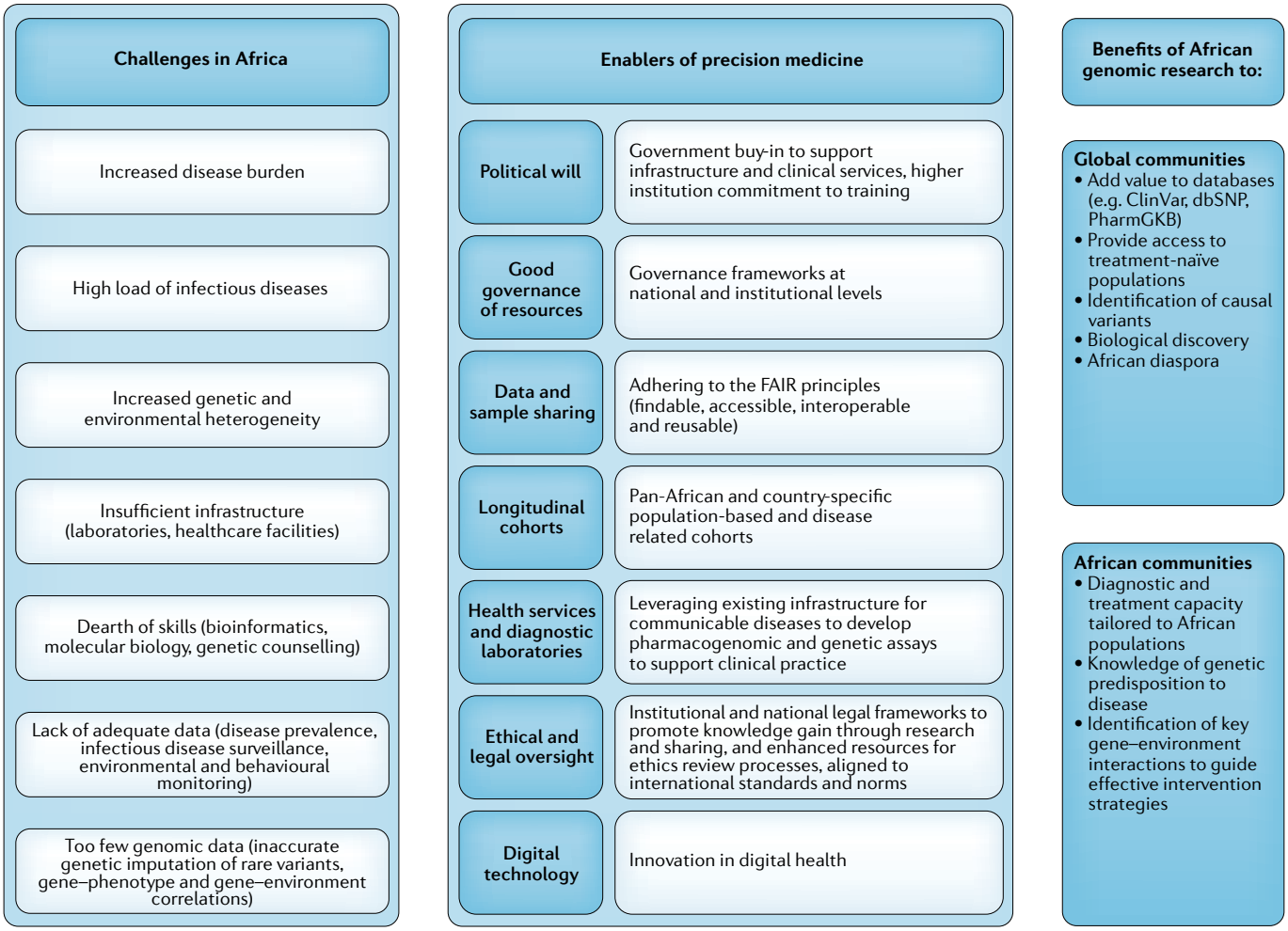


Fig. 5 | **Challenges in Africa, key enablers and potential benefits of African genome research.** Africa faces challenges for developing and implementing precision medicine due to increased disease prevalence (communicable and non-communicable diseases) and lack of infrastructure (including laboratory equipment). Key enablers for improving precision medicine and African genome research include digital technology, good governance with political will for improving health services and diagnostics, and implementation of guidelines for ethical, legal and data sharing oversight. ClinVar, public database of variant interpretations; dbSNP, public-domain archive for human single-nucleotide variants; PharmGKB, Pharmacogenomics Knowledge Base.

In the context of genomics research, several approaches to community and public engagement have been utilized. These include traditional community platforms (meetings), media (including social media) and the use of science cafes, as well as advocacy through community advisory boards and patient groups[164]. Effective science communication has been key in ensuring that participants and communities understand the scientific rationale for conducting genomic studies and increasing the utility of samples and data by sharing. The use of information leaflets explaining scientific terminology such as genes, DNA and genetic risk will be essential enablers. Empirical studies have suggested gradual shifts in perspectives when key stakeholders are actively engaged in genomic studies and when there is a platform to address concerns that are pertinent to the conduct of these studies including data sharing. For example, Bull et al.[167] reported that acceptance of data sharing practices, across five research sites in low and middle-income countries, improved as participants became more familiar with the concepts involved, the potential advantages of sharing and safeguards that could be implemented to address concerns. Deliberative methods of engagement will therefore be key in ensuring that concerns about data sharing, consent and feedback of findings are adequately addressed.

The success of precision medicine in Africa will rest on several factors. In addition to demonstrating clinical utility in a local context, it will require building trusting relationships with communities to ensure equitable partnerships, especially in underserved communities. This would include an understanding that translational science depends largely on the availability of quality data and the cooperation of individuals and their communities in granting access for the sharing and use of their data and samples[168–170]. Another focus of future work should be to increase awareness of the ability of precision medicine to benefit not only individuals but also populations, where research and knowledge could inform policy decisions in a 'precision public health' scenario.

Precision public health
Using information about genomic variation in a population to guide practices and develop policies that will benefit the majority of individuals in a given population. For example, identifying a common pharmacogenomic variant in a population that predicts an adverse drug effect, such that it leads to a policy advising against using that drug as a first-line treatment in that population.

Box 2 | **Governance models for genomics research**

Genomic research generates large data sets and leads to the collection of biospecimens on large numbers of research participants. Stored data and biospecimens can be shared between research groups in a fair and equitable manner, with appropriate attribution, if the processes are guided by sound governance structures. These include high-level ethical and legal oversight and nationally recognized regulatory infrastructure with bodies to approve and monitor research processes. In the first instance, the principal investigators are held accountable by their institutions, research ethics committees and national laws and regulations, then to those of their collaborators in other countries and, finally, also to their funders.

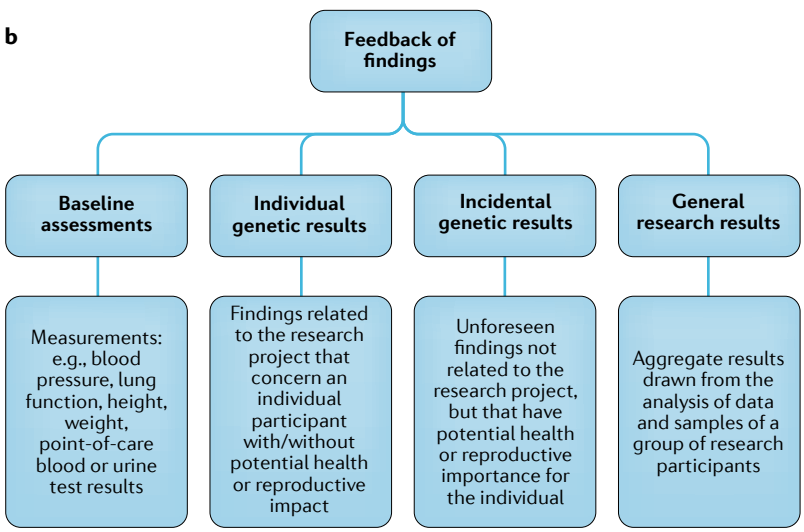
Governance frameworks need to ensure that data and biospecimens can be shared in a responsible manner without harm to participants (for example, stigmatization) and that there is a reasonable expectation of benefit sharing (see the figure, part **a**). In the first instance, data and samples can only be shared in accordance with the specific consent provided by the participant. Research on biobank resources require broad, tiered (provided it includes an appropriate option) or dynamic consent. There is general consensus that blanket consent is not a responsible principle for human research. In most African contexts, broad consent and tiered consent that includes the option of broad consent are considered the most appropriate models, usually with ethics approval required for new projects for secondary use of samples.

One of the elements of good governance concerns feedback of findings (see the figure, part **b**). Empirical research is underway to better understand the barriers and preferred methods for feedback of genetic/genomic results in different African settings (culture, context and availability of genetic counselling and services).

a

Specific consent	Consent for specific research purposes without permissions for secondary use
Broad consent	Consent that allows the use of human biological samples and associated data for future research, supported by a governance framework regulating secondary use
Tiered consent	Presents potential participants with a menu of research categories to which they may consent
Dynamic consent	A personalized, digital communications interface that enables two-way communication between participants and researchers, and puts participants at the centre of decision-making
Blanket consent	Consent for unspecified use of samples and data without restrictions

b



Consent, feedback of findings and trust-building. There is intense debate centred on the most appropriate model (or models) for consent that is appropriate for future reuse of human biological samples and associated data, a key feature of genomics and biobanking research. Informed consent is widely recognized as an important pillar for the ethical conduct of research with humans, whereas, at the same time, the challenges of obtaining valid consent are acknowledged. These challenges will apply equally to precision medicine.

Several models of informed consent have been proposed[171–174], including specific consent (consent only for the specified research), broad consent (consent obtained at the time of sample collection for future research purposes), tiered consent (allowing participants to choose to consent to one or more aspects of future research) and dynamic consent (being re-consented for every new study that would use their data and samples) (BOX 2). Despite the ongoing debates, broad consent is increasingly being accepted as the best 'compromise' given the limitations of specific consent in its traditional sense[175], and is the model proposed in some recent revised international ethics guidelines[176]. In addition, there is growing recognition that broad consent ought to be supported by appropriate governance structures that allow important scientific projects to be undertaken while respecting the wishes of research participants and their communities[171]. Questions remain, however, about what form these governance mechanisms should take, what interests and whose interests need to be protected and what key elements should inform the development of a robust framework that can support the review and conduct of genomics and biobanking research in Africa.

In a review[171] it was suggested that when participants give consent, this is equivalent to consenting to governance, thereby allowing decisions for future use of samples to be made by a body such as a data and biorepository access committee. The authors recommended that five key elements (respect; authentic community engagement and trust-building; the preservation of privacy and confidentiality; feedback of results; and capacity strengthening) should be taken into consideration when establishing a robust governance framework for genomics and biobanking research that also seeks to promote global health and research equity. Nembaware et al.[177] propose a framework for tiered consent that follows a consent process of providing information to participants and giving them a choice on a range of possible uses of their samples and data. One of the options includes agreement to "general secondary data/biospecimen use in other unrelated studies". As precision medicine develops in Africa, there will be the need for consensus-building on appropriate models of consent that will allow the sharing of data while protecting the interests of participants and patients. This process will need to be empirically informed, incorporating the views and preferences of key stakeholders, including participants and their communities.

A related recurring theme in the discourse around consent is the importance of feeding back research findings or results to participants, including findings that

Broad consent
Consent provided by a study participant for the use of their data and biospecimens for future research that could not be defined at the time the consent was obtained. The governance process of the resources are explained to participants and, usually, the future use will be approved by an ethics committee.

Tiered consent
A consent process that has several independent requests for specific uses of data and biospecimens. For example, research participants can consent to one or more of the following: consent for use in specific research; consent for data sharing; and consent for data and biospecimen sharing, with information on how the process will be governed.

were not anticipated (that is, incidental findings[178–180]). There are different types of results that can be returned and this could inform strategies for feedback of results to individuals and communities in the context of research and precision medicine[178]. Empirical studies have suggested that research participants expect some feedback on analysis conducted on their samples and data, and there is growing consensus that researchers have an ethical obligation to return research results to individuals and communities. However, questions still remain on what type of results should be returned, who should return the results in the context of the limited genetic counselling capacity in most African countries[180] and what the implications would be of such feedback for individuals, their families and communities. For example, the H3Africa Guideline for the Return of Individual Genetic Research Findings recommends that researchers should proceed with caution in making decisions on feedback of findings. The policy includes a decision tree to guide researchers to navigate the complex process of determining whether the findings are medically actionable and hold value to the individual and whether prior consent was obtained to return these findings. There are currently several ongoing empirical studies on the continent to explore participants' preferences on what, how and when findings from genetics and genomics studies should be returned.

Finally, although genetic services have improved in some African countries, there are still limitations for implementing specialized services due to cultural beliefs, taboo or religious considerations. These include various reproductive options, such as prenatal diagnosis and selective termination of pregnancy in the case of affected fetuses.

Data sharing and governance. Sharing genomics and health-related data is essential to advancing an understanding of diseases in different contexts and to translating this knowledge into tangible benefits for individual patients and populations. The responsible governance of data and biological samples in Africa should include the key principles of promoting health and well-being, fostering trust, integrity and reciprocity, and respecting individuals, families and communities (BOX 2). Furthermore, the development of national ethics and regulatory frameworks needs to address capacity-strengthening, public science education, innovative consent processes, community engagement, equitable research collaborations, feedback of research results, research integrity and improving links between research and healthcare. Three examples of governance frameworks for genomic research are presented in TABLE 1 with a summary of the key principles and elements to promote good governance.

Table 1 | **Governance frameworks for responsible data and sample sharing from large research studies**

Governance framework	Key principles	Key elements
H3Africa: Ethics and Governance Framework for Best Practices in Genomic Research and Biobanking in Africa (2017)[a]	Sensitivity of and respect for African values and cultures Designed primarily to benefit African people, but acknowledging potential benefit to the global population Genuine and active intellectual participation of African investigators and other stakeholders in research and in dissemination of findings Promote relationships characterized by respect, fairness, equity and reciprocity	African intellectual leadership and good governance Ethics review Consent Capacity-building Community engagement and benefit sharing Feedback of findings Export of data and samples Avoidance of group harm or stigma
GA4GH: Framework for the Responsible Sharing of Genomic and Health-Related Data[b]	Respect individuals, families and communities Advance research and scientific knowledge Promote health, well-being and fair distribution of benefits Foster trust, integrity and reciprocity	Transparency and accountability Engagement and risk–benefit analysis Data quality and security Privacy, data protection and confidentiality Recognition and attribution Education and training, and sustainability Accessibility and dissemination
European Commission: Global Governance of Science (2012)[c]	Openness and accessible communication with the public Participation by citizens as much as possible in all policy formation Accountability clearly apportioned among European Union institutions Effectiveness in achieving goals and objectives Coherence among institutions and policies	Citizen participation Coherent policy frameworks

GA4GH, Global Alliance for Genomics and Health; H3Africa, Human Heredity and Health in Africa. [a]https://h3africa.org/wp-content/uploads/2018/05/Final-Framework-for-African-genomics-and-biobanking_SC-.pdf. [b]https://www.ga4gh.org/genomic-data-toolkit/regulatory-ethics-toolkit/. [c]https://ec.europa.eu/research/science-society/document_library/pdf_06/global-governance-020609_en.pdf.

Unfortunately, most guidelines and regulations within Africa do not provide clear guidance on governance and how data and samples ought to be shared[181–183]. Data sharing using internationally recognized repositories is often a condition imposed by funding agencies, but is also a requirement for publication in many high-impact journals and will be a key prerequisite for the development of precision medicine in Africa. This principle is widely recognized globally to lead to better science and the more efficient use of research resources for the greater good[166], in addition to providing opportunities for verifying, replicating and facilitating future research studies.

Many challenges remain, however, and although empirical studies have suggested that despite a growing acceptance of the need for data to be shared, there is still a lack of consensus on how this ought to be done fairly to promote global justice. A systematic review[184] highlighted the barriers to data sharing to include the lack of reciprocity, suggesting that "data sharing practices have not always been fair, and data producers have often felt exploited in transactions where they receive little credit or benefit from their work, while data users that can rapidly analyse data and publish results benefit from academic credit and career advancement". Bull et al.[167] identified several concerns about data sharing, including the "potential for data sharing to harm or to fail to respect the interests of participants" and the possibility of being able to identify participants when data sets are combined. Jao et al.[175] also reported a general support for data sharing but highlighted key concerns of stakeholders and the need for a participatory process for building trust and close partnerships involving government health authorities who would be in a position to promote translational benefits.

In Africa, advancing the introduction of precision medicine approaches could complement existing healthcare systems by implementing more targeted and novel approaches to diagnosing and treating diseases. This could, in effect, help developing countries to bypass ineffective legacy systems commonly found in developed countries. For example, the World Economic Forum Precision Medicine Programme has as its anchor project a case study to develop diagnostic capability for cancer treatment in Rwanda.

Capacity development to support genomics and precision medicine. There has been an increase in the number of initiatives aimed at strengthening the capacity of African researchers and institutions to conduct genetics and genomics research in Africa. Initiatives such as the H3Africa Consortium and the African Academy of Science's Developing Excellence in Leadership, Training and Science (DELTAS) programme are developing genomics capacity in several research institutions across the continent. These initiatives have led to a shift from the practice of African researchers being relegated to sample and data collectors to being active leaders in driving science. The ultimate aim is that these developments will contribute to the implementation of precision medicine on the continent. Despite these achievements, gaps remain in translating these gains into improving medical genetics services within healthcare systems in Africa. As discussed earlier, most African countries are limited in services such as newborn screening, genetic tests and genetic risk-factor tests. Although the number of trained genomics researchers including bioinformaticians have increased over the years, most are based in research institutions. Most African countries lack genetic counsellors and the capacity for prenatal testing, particularly at the primary healthcare level. To promote equity in access to medical genetics services, capacity development should not only be concentrated at specialist hospitals but should also extended to primary care facilities that are more accessible to the majority of the population[185,186].

Another challenge that most African countries are likely to face will be related to healthcare and health systems priority-setting. With so many unmet health needs and existing challenges to addressing infectious diseases on the continent, governments would have to make difficult decisions on how to invest their limited healthcare budgets.

The future of precision medicine in Africa

Medical practice has always been personal, but precision medicine would make it possible to consider the complex nature of individual variation and population-level knowledge when examining states of health and disease. Our aim in this Review was to demonstrate that Africa, despite considerable challenges, has an enormous contribution to make and much to gain from the implementation of precision medicine approaches (FIG. 5).

Firstly, there is a need to develop, expand and scale basic and specialized skills and to build a sustainable health infrastructure. This should stretch from community workers to nurses to people with scientific and implementation skills, leading to a clinical workforce to diagnose genetic risk for disease, to counsel families and to treat and monitor health. African countries with stronger economies and better political stability are emerging as leaders for genomic research on the continent and could contribute to establishing continental harmonization and increasing the utility and visibility of genetic services. Good examples include the country-specific societies of the African Society of Human Genetics (AfSHG), the H3Africa Consortium suite of projects and the West African Genetic Medicine Centre (WAGMC).

Secondly, African research needs to focus on generating more high-quality data to better understand the functional contribution and biological impact of sets of variants on the phenotype[187]. We have provided many good examples of the contribution of African research, and although genomics currently dominates the field, other omics approaches including metabolomics, transcriptomics, proteomics, methylomics and microbiomics must be vigorously pursued in Africa[188,189]. This will require well-equipped laboratories, tissue culture facilities, animal houses, data repositories with skilled computational biologists and data managers, scientists with bioinformatics analysis and interpretation skills, biobanks that follow international good practice guidelines and governance frameworks for the management and sharing of data and research samples.

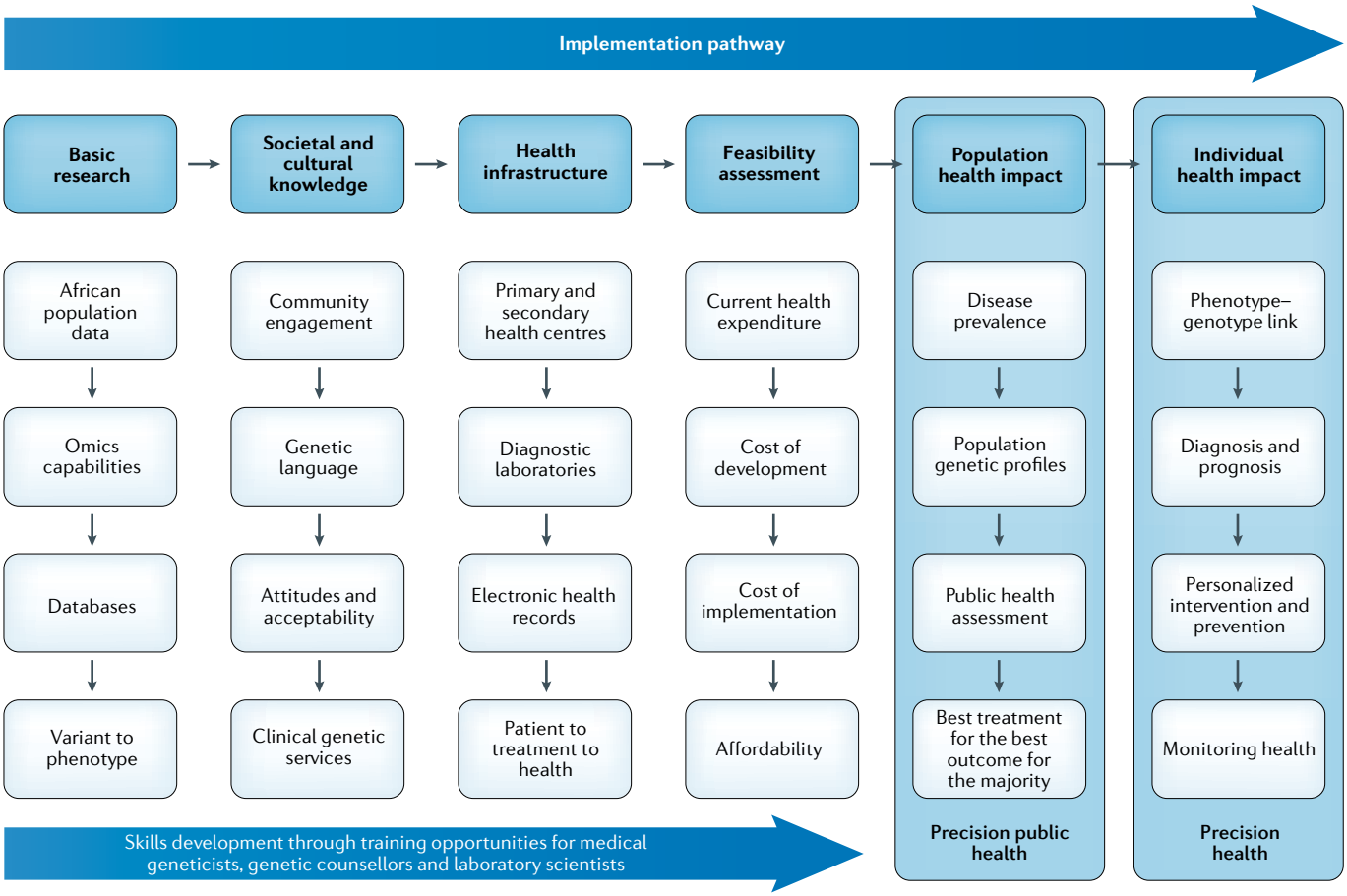


Fig. 6 | **A complex implementation pathway towards precision health.** In an African setting there is a need for much more basic research to underpin evidence-based interventions. A clinical service will need to be sensitive to country-specific and ethnic-specific societal and cultural norms to ensure that services will benefit populations. Implementation of clinical genetic services will require a health infrastructure in terms of laboratories, but also, more importantly, skilled personnel such as medical geneticists, genetic counsellors and laboratory scientists. Once these are in place, the affordability and logistic feasibility need to be assessed for both precision public health and personalized health approaches.

Thirdly, the knowledge needs to be translated into clinical practice through a combination of skills development and relevant African data to design and construct tools for precise diagnosis and targeted treatment. The global market has developed genotyping solutions that are affordable, automated, fast and easy to interpret. For example, the recently implemented CRISPR–Cas13a-based genotyping assay[190] is being tested in Africa for the diagnosis of several tropical diseases, such as yellow fever, Lassa fever, dengue fever and Zika, and can be easily applied to any known and emergent pathogens. The development of treatments mimicking the natural protection of African populations against pathogens or based on adapted metabolism is a promising avenue to explore. African governments, philanthropists, non-governmental organizations and venture capitalists are key players in funding start-up companies to introduce precision medicine solutions in Africa. A newcomer to the field, based in Nigeria, is 54gene, an African genomics research, services and development company that recently launched the African Centre for Translational Genomics (ACTG). Big pharma is also beginning to show interest in African genomics, possibly triggered by the African diaspora and the projected population growth on the continent. Africa is home to ~1.6 billion people and its current annual growth rate is 2.3%, with more than 40% of the population younger than 15 years (see the United Nations World Population Prospects).

Lastly, there is a need to develop accessible, relevant, integrated and affordable healthcare systems to serve the population. Medical doctors and health professionals need to be supported by diagnostic centres and treatment facilities. Hospitals and clinics need good government support and sound governance structures that can evaluate the potential gains from precision medicine approaches. For example, pharmacogenomic and population-based approaches could be leveraged to ensure that African government-implemented essential medicine lists have the best drugs at the right doses to benefit the majority of the population and to minimize adverse events. Initial investment in research can be capitalized by lower wastage and more effective treatment through precision medicine. Governments should be more hesitant to approve the implementation of expensive treatments that would be effective in <10% of patients and focus on stratifying patients using affordable diagnostic assays into groups for targeted affordable interventions.

Genomic medicine
The use of genetic information (genomic variants) to guide diagnosis and clinical interventions or to be used in weighing reproductive options.

We propose that successful implementation of precision medicine in Africa could be achieved through three overarching strategies. First is building capabilities for omics-data generation and electronic health records in Africa and strengthening the skills base for molecular biology, genomics and bioinformatics through infrastructural and digital innovation. Second is promoting data sharing and data integration in Africa and ensuring that data are standardized and comparable, while adhering to the international FAIR principles (findable, accessible, interoperable and reusable[191]), to support wide secondary use of data and knowledge generation. Last is establishing strong data-science platforms in Africa that have the capacity to link electronic health records captured as part of clinical care to omics-data and to make these data available to clinicians and health professionals. The elements of an implementation pathway, from basic research to understanding societal context and health infrastructure to assessing feasibility and potential impact, are outlined in FIG. 6.

Conclusions

African populations and their genomes present a treasure trove for discovery to support global health. Well-characterized examples, such as PCSK9 inhibitors for treating hypercholesterolaemia, *CYP2B6* genotyping to assist with the EFV dosage for African HIV-positive patients, and the link between *APOL1* risk variants and kidney disease, illustrate the role of African diversity on treatment approaches. Africa is still grappling with many infectious diseases but has contributed to understanding and coping with the burden of new viral infections. The African Union's newly established Africa CDC, together with the World Health Organization, has an important continent-wide role in preventing pandemics and anticipating future risk, while supporting national initiatives. However, as NCDs are on the rise in Africa, there is a need to develop precision medicine leapfrog technology for prevention and delayed onset of disease to minimize the harmful effects of obesity, cardiovascular disease, diabetes, and cancer. Leadership and political will are the key ingredients to promote health across Africa's permeable borders to meet the United Nations Sustainable Development goals. Africa needs a complex multilayered model of innovative solutions across the socio-economic spectrum to reap the benefits of both genomic medicine and a precision medicine approach.

Published online: 11 January 2021

1. Rotimi, C. et al. Research capacity. Enabling the genomic revolution in Africa. *Science* **344**, 1346–1348 (2014).
2. Cohoon, T. J. & Bhavnani, S. P. Toward precision health: applying artificial intelligence analytics to digital health biometric datasets. *Per. Med.* **17**, 307–316 (2020).
3. Mapesi, H. & Paris, D. H. Non-communicable diseases on the rise in sub-Saharan Africa, the underappreciated threat of a dual disease burden. *Praxis* **108**, 997–1005 (2019).
4. Mudie, K. et al. Non-communicable diseases in sub-Saharan Africa: a scoping review of large cohort studies. *J. Glob. Health* **9**, 020409 (2019).
5. Tekola-Ayele, F. & Rotimi, C. N. Translational genomics in low- and middle-income countries: opportunities and challenges. *Public Health Genomics* **18**, 242–247 (2015).
6. Mulder, N. et al. H3Africa: current perspectives. *Pharmgenomics Pers. Med.* **11**, 59–66 (2018).
7. Owolabi, M. O. et al. Data resource profile: Cardiovascular H3Africa Innovation Resource (CHAIR). *Int. J. Epidemiol.* **48**, 366–367g (2019).
8. Adedokun, B. O., Olopade, C. O. & Olopade, O. I. Building local capacity for genomics research in Africa: recommendations from analysis of publications in sub-Saharan Africa from 2004 to 2013. *Glob. Health Action.* **9**, 31026 (2016).
9. Stringer, C. The origin and evolution of *Homo sapiens*. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **371**, 20150237 (2016).
10. Schlebusch, C. M. et al. Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science* **358**, 652–655 (2017).
11. Hublin, J. J. et al. New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature* **546**, 289–292 (2017).
12. Scerri, E. M. L. et al. Did our species evolve in subdivided populations across Africa, and why does it matter? *Trends Ecol. Evol.* **33**, 582–594 (2018).
13. Cruciani, F. et al. A revised root for the human Y chromosomal phylogenetic tree: the origin of patrilineal diversity in Africa. *Am. J. Hum. Genet.* **88**, 814–818 (2011).
14. Rito, T. et al. The first modern human dispersals across Africa. *PLoS ONE* **8**, e80031 (2013).
15. Santander, C., Montinaro, F. & Capelli, C. Searching for archaic contribution in Africa. *Ann. Hum. Biol.* **46**, 129–139 (2019).
16. Skoglund, P. & Mathieson, I. Ancient genomics of modern humans: the first decade. *Annu. Rev. Genomics Hum. Genet.* **19**, 381–404 (2018).
17. Durvasula, A. & Sankararaman, S. Recovering signals of ghost archaic introgression in African populations. *Sci. Adv.* **6**, eaax5097 (2020).
18. Chen, L., Wolf, A. B., Fu, Q., Li, L. & Akey, J. M. Identifying and interpreting apparent neanderthal ancestry in African individuals. *Cell* **180**, 1–11 (2020).
19. Veeramah, K. R. et al. An early divergence of KhoeSan ancestors from those of other modern humans is supported by an ABC-based analysis of autosomal resequencing data. *Mol. Biol. Evol.* **29**, 617–630 (2012).
20. Verdu, P. et al. Origins and genetic diversity of pygmy hunter-gatherers from Western Central Africa. *Curr. Biol.* **19**, 312–318 (2009).
21. Fan, S. et al. African evolutionary history inferred from whole genome sequence data of 44 indigenous African populations. *Genome Biol.* **20**, 82 (2019).
22. Tishkoff, S. A. et al. The genetic structure and history of Africans and African Americans. *Science* **324**, 1035–1044 (2009).
23. Hernández, C. L. et al. Human genomic diversity where the Mediterranean joins the Atlantic. *Mol. Biol. Evol.* **37**, 1041–1055 (2020).
24. van de Loosdrecht, M. et al. Pleistocene North African genomes link Near Eastern and sub-Saharan African human populations. *Science* **360**, 548–552 (2018).
25. Pickrell, J. K. et al. The genetic prehistory of southern Africa. *Nat. Commun.* **3**, 1143 (2012).
26. Triska, P. et al. Extensive admixture and selective pressure across the Sahel Belt. *Genome Biol. Evol.* **7**, 3484–3495 (2015).
27. Patin, E. et al. Dispersals and genetic adaptation of Bantu-speaking populations in Africa and North America. *Science* **356**, 543–546 (2017).
This paper examines the genetic diversity of Bantu-speakers, modelling the timing of migration and admixture during the Bantu expansion. Interestingly, the paper identifies adaptive introgression of genes from local populations along the Bantu waves, including specific immune-related genes.
28. Lopez, M. et al. The demographic history and mutational load of African hunter-gatherers and farmers. *Nat. Ecol. Evol.* **2**, 721–730 (2018).
29. Gallego Llorente, M. et al. Ancient Ethiopian genome reveals extensive Eurasian admixture throughout the African continent. *Science* **350**, 820–822 (2015).
30. Skoglund, P. et al. Reconstructing prehistoric African population structure. *Cell* **171**, 59–71.e21 (2017).
31. Lipson, M. et al. Ancient West African foragers in the context of African population history. *Nature* **577**, 665–670 (2020).
32. Hellenthal, G. et al. A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).
33. Semo, A. et al. Along the Indian Ocean coast: genomic variation in Mozambique provides new insights into the Bantu expansion. *Mol. Biol. Evol.* **37**, 406–416 (2019).
34. Choudhury, A. et al. High-depth African genomes inform human migration and health. *Nature* **586**, 741–748 (2020).
This paper presents the largest whole-genome sequence study of African populations describing >3 million novel variants and highlighting disease-associated alleles.
35. Gurdasani, D. et al. The African genome variation project shapes medical genetics in Africa. *Nature* **517**, 327–332 (2015).
This paper introduces The African Genome Variation Project, with dense genotypes from 1481 individuals and whole-genome sequences from 320 individuals across SSA. Besides presenting novel evidence of complex population history and loci under selection, the paper demonstrates further improvements in imputation accuracy, strengthening the case for large-scale sequencing efforts of diverse African haplotypes.
36. Pickrell, J. K. et al. Ancient west Eurasian ancestry in southern and eastern Africa. *Proc. Natl Acad. Sci. USA* **111**, 2632–2637 (2014).
37. Brucato, N. et al. The Comoros show the earliest Austronesian gene flow into the Swahili Corridor. *Am. J. Hum. Genet.* **102**, 58–68 (2018).
38. Pierron, D. et al. Genomic landscape of human diversity across madagascar. *Proc. Natl Acad. Sci. USA* **114**, E6498–E6506 (2017).
39. Brucato, N. et al. Evidence of austronesian genetic lineages in East Africa and South Arabia: complex dispersal from madagascar and Southeast Asia. *Genome Biol. Evol.* **11**, 748–758 (2019).
40. Choudhury, A. et al. Whole-genome sequencing for an enhanced understanding of genetic variation among South Africans. *Nat. Commun.* **8**, 2062 (2017).
41. May, A. et al. Genetic diversity in black South Africans from Soweto. *BMC Genomics* **14**, 644 (2013).
42. Lovejoy, P. E. The impact of the Atlantic slave trade on Africa: a review of the literature. *J. Afr. Hist.* **30**, 365–394 (1989).
43. Macaulay, V. et al. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* **308**, 1034–1036 (2005).
44. Gravel, S. et al. Demographic history and rare allele sharing among human populations. *Proc. Natl Acad. Sci. USA* **108**, 11983–11988 (2011).

45. The 1000 Genomes Project Consortium, et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
46. Chen, J. et al. Genome-wide association study of type 2 diabetes in Africa. *Diabetologia* **62**, 1204–1211 (2019).
47. Martin, A. R., Teferra, S., Möller, M., Hoal, E. G. & Daly, M. J. The critical needs and challenges for genetic architecture studies in Africa. *Curr. Opin. Genet. Dev.* **53**, 113–120 (2018).
48. Gurdasani, D., Barroso, I., Zeggini, E. & Sandhu, M. S. Genomics of disease risk in globally diverse populations. *Nat. Rev. Genet.* **20**, 520–535 (2019).
49. Fu, W. et al. Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**, 216–220 (2013).
50. Simons, Y. B., Turchin, M. C., Pritchard, J. K. & Sella, G. The deleterious mutation load is insensitive to recent population history. *Nat. Genet.* **46**, 220–224 (2014).
51. Quintana-Murci, L. Human immunology through the lens of evolutionary genetics. *Cell* **177**, 184–199 (2019).
52. Grossman, S. R. et al. Identifying recent adaptations in large-scale genomic data. *Cell* **152**, 703–713 (2013).
53. Guernier, V., Hochberg, M. E. & Guegan, J. F. Ecology drives the worldwide distribution of human diseases. *PLoS Biol.* **2**, e141 (2004).
54. Cairns, M. E. et al. Seasonality in malaria transmission: implications for case-management with long-acting artemisinin combination therapy in sub-Saharan Africa. *Malar. J.* **14**, 321 (2015).
55. Zhao, S., Lin, Q., He, D. & Stone, L. Meningitis epidemics shift in sub-Saharan belt. *Int. J. Infect. Dis.* **68**, 79–82 (2018).
56. Alexander, K. A. et al. What factors might have led to the emergence of Ebola in West Africa? *PLoS Negl. Trop. Dis.* **9**, e0003652 (2015).
57. Fan, S., Hansen, M. E., Lo, Y. & Tishkoff, S. A. Going global by adapting local: a review of recent human adaptation. *Science* **354**, 54–59 (2016).
58. Sierra, B. et al. OSBPL10, RXRA and lipid metabolism confer African-ancestry protection against dengue haemorrhagic fever in admixed Cubans. *PLoS Pathog.* **13**, e1006220 (2017).
59. Bentley, A. R. & Rotimi, C. N. Interethnic variation in lipid profiles: implications for underidentification of African-Americans at risk for metabolic disorders. *Expert Rev. Endocrinol. Metab.* **7**, 659–667 (2012).
60. Coram, M. A. et al. Genome-wide characterization of shared and distinct genetic components that influence blood lipid levels in ethnically diverse human populations. *Am. J. Hum. Genet.* **92**, 904–916 (2013).
61. Cohen, J. et al. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat. Genet.* **37**, 161–165 (2005).
This paper reports that loss-of-function mutations in the *PCSK9* gene, preventing hypercholesterolaemia, were common in AAs (combined frequency, 2%) but rare in EAs (<0.1%) and were associated with a 40% reduction in plasma levels of LDL cholesterol.
62. Ng, M. C. Y. et al. Discovery and fine-mapping of adiposity loci using high density imputation of genome-wide association studies in individuals of African ancestry: African Ancestry Anthropometry Genetics Consortium. *PLoS Genet.* **13**, e1006719 (2017).
63. Doumatey, A. P., Ekoru, K., Adeyemo, A. & Rotimi, C. N. Genetic basis of obesity and type 2 diabetes in Africans: impact on precision medicine. *Curr. Diab Rep.* **19**, 105 (2019).
64. Gomez-Olive, F. X. et al. Regional and sex differences in the prevalence and awareness of hypertension: an H3Africa AWI-Gen study across 6 sites in sub-Saharan Africa. *Glob. Heart* **12**, 81–90 (2017).
65. Young, J. H. et al. Differential susceptibility to hypertension is due to selection during the out-of-Africa expansion. *PLoS Genet.* **1**, e82 (2005).
66. Franceschini, N. et al. Genome-wide association analysis of blood-pressure traits in African-ancestry individuals reveals common associated genes in African and non-African populations. *Am. J. Hum. Genet.* **93**, 545–554 (2013).
67. Band, G., Rockett, K. A., Spencer, C. C. & Kwiatkowski, D. P. A novel locus of resistance to severe malaria in a region of ancient balancing selection. *Nature* **526**, 253–257 (2015).
68. Hamblin, M. T., Thompson, E. E. & Di Rienzo, A. Complex signatures of natural selection at the Duffy blood group locus. *Am. J. Hum. Genet.* **70**, 369–383 (2002).
69. McManus, K. F. et al. Population genetic analysis of the DARC locus (Duffy) reveals adaptation from standing variation associated with malaria resistance in humans. *PLoS Genet.* **13**, e1006560 (2017).
70. Pierron, D. et al. Strong selection during the last millennium for African ancestry in the admixed population of Madagascar. *Nat. Commun.* **9**, 932 (2018).
71. Genovese, G. et al. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* **329**, 841–845 (2010).
72. Cooper, A. et al. APOL1 renal risk variants have contrasting resistance and susceptibility associations with African trypanosomiasis. *eLife* **6**, e25461 (2017).
73. Ko, W. Y. et al. Identifying Darwinian selection acting on different human APOL1 variants among diverse African populations. *Am. J. Hum. Genet.* **93**, 54–66 (2013).
74. Williams, T. N. et al. Negative epistasis between the malaria-protective effects of α^+-thalassemia and the sickle cell trait. *Nat. Genet.* **37**, 1253–1257 (2005).
75. Saraf, S. L. et al. APOL1, α-thalassemia, and BCL11A variants as a genetic risk profile for progression of chronic kidney disease in sickle cell anemia. *Haematologica* **102**, e1–e6 (2017).
76. Zahr, R. S. et al. Children with sickle cell anemia and APOL1 genetic variants develop albuminuria early in life. *Haematologica* **104**, e385–e387 (2019).
77. Engelken, J. et al. Extreme population differences in the human zinc transporter ZIP4 (SLC39A4) are explained by positive selection in sub-Saharan Africa. *PLoS Genet.* **10**, e1004128 (2014).
78. Hood, M. I. & Skaar, E. P. Nutritional immunity: transition metals at the pathogen–host interface. *Nat. Rev. Microbiol.* **10**, 525–537 (2012).
79. Schaafsma, T. et al. Africa's oesophageal cancer corridor: geographic variations in incidence correlate with certain micronutrient deficiencies. *PLoS ONE* **10**, e0140107 (2015).
80. Vicente, M., Jakobsson, M., Ebbesen, P. & Schlebusch, C. M. Genetic affinities among southern Africa hunter-gatherers and the impact of admixing farmer and herder populations. *Mol. Biol. Evol.* **36**, 1849–1861 (2019).
81. Lachance, J. et al. Evolutionary history and adaptation from high-coverage whole-genome sequences of diverse African hunter-gatherers. *Cell* **150**, 457–469 (2012).
82. Xu, D. et al. Archaic hominin introgression in Africa contributes to functional salivary MUC7 genetic variation. *Mol. Biol. Evol.* **34**, 2704–2715 (2017).
83. Bustamante, C. D., Burchard, E. G. & De la Vega, F. M. Genomics for the world. *Nature* **475**, 163–165 (2011).
84. Petrovski, S. & Goldstein, D. B. Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol.* **17**, 157 (2016).
85. Popejoy, A. B. & Fullerton, S. M. Genomics is failing on diversity. *Nature* **538**, 161–164 (2016).
86. Gurdasani, D. et al. Uganda genome resource enables insights into population history and genomic discovery in Africa. *Cell* **179**, 984–1002.e36 (2019).
By performing a multi-trait pan-African GWAS of up to 14,126 individuals, this paper identifies novel loci associated with anthropometric, haematological, lipid and glycaemic traits. Several functionally important signals are driven by Africa-specific variants, highlighting the value of studying diverse populations across the region.
87. Rotimi, C. N. et al. The genomic landscape of African populations in health and disease. *Hum. Mol. Genet.* **26**, R225–R236 (2017).
88. Koboldt, D. C., Steinberg, K. M., Larson, D. E., Wilson, R. K. & Mardis, E. R. The next-generation sequencing revolution and its impact on genomics. *Cell* **155**, 27–38 (2013).
89. Richards, S. et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* **17**, 405–424 (2015).
90. Amr, S. S. et al. Using large sequencing data sets to refine intragenic disease regions and prioritize clinical variant interpretation. *Genet. Med.* **19**, 496–504 (2017).
91. Mitropoulos, K. et al. Success stories in genomic medicine from resource-limited countries. *Hum. Genomics* **9**, 11 (2015).
92. Romdhane, L. et al. Consanguinity and inbreeding in health and disease in North African populations. *Annu. Rev. Genomics Hum. Genet.* **20**, 155–179 (2019).
93. Krause, A., Seymour, H. & Ramsay, M. Common and founder mutations for monogenic traits in sub-Saharan African populations. *Annu. Rev. Genomics Hum. Genet.* **19**, 149–175 (2018).
94. Kabahuma, R. I. et al. Absence of GJB2 gene mutations, the GJB6 deletion (GJB6-D13S1830) and four common mitochondrial mutations in nonsyndromic genetic hearing loss in a South African population. *Int. J. Pediatr. Otorhinolaryngol.* **75**, 611–617 (2011).
95. Wonkam, A. et al. No evidence for clinical utility in investigating the connexin genes GJB2, GJB6 and GJA1 in non-syndromic hearing loss in Black Africans. *S Afr. Med. J.* **105**, 23–26 (2015).
96. Lebeko, K. et al. Targeted genomic enrichment and massively parallel sequencing identifies novel nonsyndromic hearing impairment pathogenic variants in Cameroonian families. *Clin. Genet.* **90**, 288–290 (2016).
97. Brobby, G. W., Müller-Myhsok, B. & Horstmann, R. D. Connexin 26 R143W mutation associated with recessive nonsyndromic sensorineural deafness in Africa. *N. Engl. J. Med.* **338**, 548–550 (1998).
98. Adadey, S. M. et al. GJB2 and GJB6 mutations in non-syndromic childhood hearing impairment in Ghana. *Front. Genet.* **10**, 841 (2019).
99. Hamelmann, C. et al. Pattern of connexin 26 (GJB2) mutations causing sensorineural hearing impairment in Ghana. *Hum. Mutat.* **18**, 84–85 (2001).
100. Manga, P., Kerr, R., Ramsay, M. & Kromberg, J. G. Biology and genetics of oculocutaneous albinism and vitiligo — common pigmentation disorders in southern Africa. *S Afr. Med. J.* **103**, 984–988 (2013).
101. Manolio, T. A. et al. Finding the missing heritability of complex diseases. *Nature* **461**, 747–753 (2009).
102. Yang, J. et al. Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* **42**, 565–569 (2010).
103. Adeyemo, A. A. et al. Evaluation of genome wide association study associated type 2 diabetes susceptibility loci in sub Saharan Africans. *Front. Genet.* **6**, 335 (2015).
104. Hauser, M. A. et al. Association of genetic variants with primary open-angle glaucoma among individuals with African ancestry. *JAMA* **322**, 1682–1691 (2019).
105. Govind, N. et al. HLA-DRB1 amino acid positions and residues associated with antibody-positive rheumatoid arthritis in Black South Africans. *J. Rheumatol.* **46**, 138–144 (2019).
106. Gulsuner, S. et al. Genetics of schizophrenia in the South African Xhosa. *Science* **367**, 569–573 (2020).
107. Manolio, T. A. In retrospect: a decade of shared genomic associations. *Nature* **546**, 360–361 (2017).
108. Drake, T. M., Knight, S. R., Harrison, E. M. & Søreide, K. Global inequities in precision medicine and molecular cancer research. *Front. Oncol.* **8**, 346 (2018).
109. Burkitt, D. A sarcoma involving the jaws in African children. *Br. J. Surg.* **46**, 218–223 (1958).
110. Spano, J. P., Atlan, D., Breau, J. L. & Farge, D. AIDS and non-AIDS-related malignancies: a new vexing challenge in HIV-positive patients. Part I: Kaposi's sarcoma, non-Hodgkin's lymphoma, and Hodgkin's lymphoma. *Eur. J. Intern. Med.* **13**, 170–179 (2002).
111. Pitt, J. J. et al. Characterization of Nigerian breast cancer reveals prevalent homologous recombination deficiency and aggressive molecular features. *Nat. Commun.* **9**, 4181 (2018).
112. Zheng, Y. et al. Inherited breast cancer in Nigerian women. *J. Clin. Oncol.* **36**, 2820–2825 (2018).
113. Rebbeck, T. R. et al. Mutational spectrum in a worldwide study of 29,700 families with BRCA1 or BRCA2 mutations. *Hum. Mutat.* **39**, 593–620 (2018).
114. Manrai, A. K. et al. Genetic misdiagnoses and the potential for health disparities. *N. Engl. J. Med.* **375**, 655–665 (2016).
115. Fitzmaurice, C. et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 29 cancer groups, 1990 to 2017: a systematic analysis for the Global Burden of Disease Study. *JAMA Oncol.* **5**, 1749–1768 (2019).
116. Reay, W. R. et al. Polygenic disruption of retinoid signalling in schizophrenia and a severe cognitive deficit subtype. *Mol. Psychiatry* **25**, 719–731 (2018).
117. Martin, A. R. et al. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat. Genet.* **51**, 584–591 (2019).

118. Duncan, L. et al. Analysis of polygenic risk score usage and performance in diverse human populations. *Nat. Commun.* **10**, 3328 (2019).
This paper presents a reanalysis of 10 years of polygenic scoring studies, showing that 67% involve European ancestry participants, 19% East Asian populations and 3.8% African, Hispanic or Indigenous people, revealing that the predictive performance of current polygenic scores is lower in non-European ancestry samples. There is a clear need for large-scale GWAS in diverse human populations.
119. Torkamani, A., Wineinger, N. E. & Topol, E. J. The personal and clinical utility of polygenic risk scores. *Nat. Rev. Genet.* **19**, 581–590 (2018).
120. Mosley, J. D. et al. Predictive accuracy of a polygenic risk score compared with a clinical risk score for incident coronary heart disease. *JAMA* **323**, 627–635 (2020).
121. Tada, H. et al. Risk prediction by genetic risk scores for coronary heart disease is independent of self-reported family history. *Eur. Heart J.* **37**, 561–567 (2016).
122. Rao, A. S. & Knowles, J. W. Polygenic risk scores in coronary artery disease. *Curr. Opin. Cardiol.* **34**, 435–440 (2019).
123. Elliott, J. et al. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA* **323**, 636–645 (2020).
124. Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
125. Kuchenbaecker, K. et al. The transferability of lipid loci across African, Asian and European cohorts. *Nat. Commun.* **10**, 4330 (2019).
This paper assesses whether the genetic determinants of blood lipids are shared across populations, finding evidence of reproducibility for ~75% of the major lipid loci from European discovery studies, except TG loci in the Ugandan samples (10% of loci). The authors hypothesize that the non-transferable loci might affect lipids by modifying food intake in environments rich in certain nutrients.
126. Wojcik, G. L. et al. Genetic analyses of diverse populations improves discovery for complex traits. *Nature* **570**, 514–518 (2019). **The Population Architecture using Genomics and Epidemiology (PAGE) study conducts a GWAS of 26 clinical and behavioural phenotypes in 49,839 non-European individuals, and the results show evidence of effect-size heterogeneity across ancestries for published GWAS associations, substantial benefits for fine-mapping using diverse cohorts and insights into clinical implications**.
127. Wei, C. Y., Lee, M. T. & Chen, Y. T. Pharmacogenomics of adverse drug reactions: implementing personalized medicine. *Hum. Mol. Genet.* **21**, R58–R65 (2012).
128. Loscalzo, J. Precision medicine. *Circ. Res.* **124**, 987–989 (2019).
129. Spear, B. B., Heath-Chiozzi, M. & Huff, J. Clinical application of pharmacogenetics. *Trends Mol. Med.* **7**, 201–204 (2001).
130. Urban, M. F. Genomics in medicine: from promise to practice. *S Afr. Med. J.* **105**, 545–547 (2015).
131. Hovelson, D. H. et al. Characterization of ADME gene variation in 21 populations by exome sequencing. *Pharmacogenet. Genomics* **27**, 89–100 (2017).
132. Nordling, L. How the genomics revolution could finally help Africa. *Nature* **544**, 20–22 (2017).
133. Masimirembwa, C., Dandara, C. & Leutscher, P. D. Rolling out efavirenz for HIV precision medicine in Africa: are we ready for pharmacovigilance and tackling neuropsychiatric adverse effects? *OMICS* **20**, 575–580 (2016).
134. Gross, R. et al. Slow efavirenz metabolism genotype is common in Botswana. *J. Acquir. Immune Defic. Syndr.* **49**, 336–337 (2008).
135. Zembutsu, H. Pharmacogenomics toward personalized tamoxifen therapy for breast cancer. *Pharmacogenomics* **16**, 287–296 (2015).
136. Walko, C. M. & McLeod, H. Use of CYP2D6 genotyping in practice: tamoxifen dose adjustment. *Pharmacogenomics* **13**, 691–697 (2012).
137. Popejoy, A. B. Diversity in precision medicine and pharmacogenetics: methodological and conceptual considerations for broadening participation. *Pharmgenomics Pers. Med.* **12**, 257–271 (2019).
138. De, T. et al. Association of genetic variants with warfarin-associated bleeding among patients of African descent. *JAMA* **320**, 1670–1677 (2018).
139. Hobbs, A. & Ramsay, M. Epigenetics and the burden of noncommunicable disease: a paucity of research in Africa. *Epigenomics* **7**, 627–639 (2015).
140. Sokolowski, H. M. et al. The *Drosophila* foraging gene human orthologue PRKG1 predicts individual differences in the effects of early adversity on maternal sensitivity. *Cogn. Dev.* **42**, 62–73 (2017).
141. Fraser, H. B., Lam, L. L., Neumann, S. M. & Kobor, M. S. Population-specificity of human DNA methylation. *Genome Biol.* **13**, R8 (2012).
142. Huan, T. et al. Genome-wide identification of DNA methylation QTLs in whole blood highlights pathways for cardiovascular disease. *Nat. Commun.* **10**, 4267 (2019).
143. Pierce, B. L. et al. Co-occurring expression and methylation QTLs allow detection of common causal variants and shared biological mechanisms. *Nat. Commun.* **9**, 804 (2018).
144. Fagny, M. et al. The epigenomic landscape of African rainforest hunter-gatherers and farmers. *Nat. Commun.* **6**, 10047 (2015).
145. Vukojevic, V. et al. Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *J. Neurosci.* **34**, 10274–10284 (2014).
146. Perroud, N. et al. The Tutsi genocide and transgenerational transmission of maternal stress: epigenetics and biology of the HPA axis. *World J. Biol. Psychiatry* **15**, 334–345 (2014).
147. Rudahindwa, S. et al. Transgenerational effects of the genocide against the Tutsi in Rwanda: a post-traumatic stress disorder symptom domain analysis. *AAS Open Res.* https://doi.org/10.12688/aasopenres.12848.1 (2020).
148. Dominguez-Salas, P. et al. Maternal nutrition at conception modulates DNA methylation of human metastable epialleles. *Nat. Commun.* **5**, 3746 (2014).
149. Schulze, K. V. et al. Edematous severe acute malnutrition is characterized by hypomethylation of DNA. *Nat. Commun.* **10**, 5791 (2019).
150. Waterland, R. A. et al. Season of conception in rural gambia affects DNA methylation at putative human metastable epialleles. *PLoS Genet.* **6**, e1001252 (2010).
151. Black, R. E. et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. *Lancet* **382**, 427–451 (2013).
152. Bhutta, Z. A. et al. Severe childhood malnutrition. *Nat. Rev. Dis. Prim.* **3**, 17067 (2017).
153. Steyn, A. et al. Epigenetic modification of the pentose phosphate pathway and the IGF-axis in women with gestational diabetes mellitus. *Epigenomics* **11**, 1371–1385 (2019).
154. Costantino, S. et al. Epigenetics and precision medicine in cardiovascular patients: from basic concepts to the clinical arena. *Eur. Heart J.* **39**, 4150–4158 (2018).
155. Musanabaganwa, C. et al. Building skills and resources for genomics, epigenetics, and bioinformatics research for Africa: report of the joint 11th Conference of the African Society of Human Genetics and 12th H3Africa Consortium, 2018. *Am. J. Trop. Med. Hyg.* **102**, 1417–1424 (2020).
Local African genetics societies are paramount in encouraging research in Africa. This paper is the meeting report for the 11th Congress of the African Society of Human Genetics (AfSHG) in conjunction with the 12th H3Africa Consortium meeting in 2018.
156. El-Kamah, G. Y. et al. Developing a road map to spread genomic knowledge in Africa: 10th Conference of the African Society of Human Genetics, Cairo, Egypt. *Am. J. Trop. Med. Hyg.* **102**, 719–723 (2020).
157. Ndiaye Diallo, R. et al. Strengthening human genetics research in Africa: report of the 9th meeting of the African Society of Human Genetics in Dakar in May 2016. *Glob. Health Epidemiol. Genom.* **2**, e10 (2017).
158. de Vries, J. et al. Ethical issues in human genomics research in developing countries. *BMC Med. Ethics* **12**, 5 (2011).
159. Nyika, A. Ethical and practical challenges surrounding genetic and genomic research in developing countries. *Acta Trop.* **112** (Suppl. 1), S21–S31 (2009).
160. Marshall, P. A. et al. Voluntary participation and informed consent to international genetic research. *Am. J. Public Health* **96**, 1989–1995 (2006).
161. Tindana, P. et al. Seeking consent to genetic and genomic research in a rural Ghanaian setting: a qualitative study of the MalariaGEN experience. *BMC Med. Ethics* **13**, 15 (2012).
162. Ghansah, A. et al. Monitoring parasite diversity for malaria elimination in sub-Saharan Africa. *Science* **345**, 1297–1298 (2014).
163. Yakubu, A. et al. Model framework for governance of genomic research and biobanking in Africa — a content description. *AAS Open Res.* **1**, 13 (2018).
164. Rotimi, C. et al. Community engagement and informed consent in the International HapMap project. *Community Genet.* **10**, 186–198 (2007).
165. Tindana, P. et al. Community engagement strategies for genomic studies in Africa: a review of the literature. *BMC Med. Ethics* **16**, 24 (2015).
166. Sankoh, O. & Ijsselmuiden, C. Sharing research data to improve public health: a perspective from the Global South. *Lancet* **378**, 401–402 (2011).
167. Bull, S. et al. Best practices for ethical sharing of individual-level health research data from low- and middle-income settings. *J. Empir. Res. Hum. Res Ethics* **10**, 302–313 (2015).
168. Kraft, S. A. et al. Beyond consent: building trusting relationships with diverse populations in precision medicine research. *Am. J. Bioeth.* **18**, 3–20 (2018).
169. Sabatello, M., Callier, S., Garrison, N. A. & Cohn, E. G. Trust, precision medicine research, and equitable participation of underserved populations. *Am. J. Bioeth.* **18**, 34–36 (2018).
170. Magnus, D. & Batten, J. N. Building a trustworthy precision health research enterprise. *Am. J. Bioeth.* **18**, 1–2 (2018).
171. Tindana, P. & de Vries, J. Broad consent for genomic research and biobanking: perspectives from low- and middle-income countries. *Annu. Rev. Genomics Hum. Genet.* **17**, 375–393 (2016).
172. Stein, D. T. & Terry, S. F. Reforming biobank consent policy: a necessary move away from broad consent toward dynamic consent. *Genet. Test. Mol. Biomarkers* **17**, 855–856 (2013).
173. Kaye, J. et al. Dynamic consent: a patient interface for twenty-first century research networks. *Eur. J. Hum. Genet.* **23**, 141–146 (2015).
174. Munung, N. S. et al. Obtaining informed consent for genomics research in Africa: analysis of H3Africa consent documents. *J. Med. Ethics* **42**, 132–137 (2016).
175. Jao, I. et al. Involving research stakeholders in developing policy on sharing public health research data in Kenya: views on fair process for informed consent, access oversight, and community engagement. *J. Empir. Res. Hum. Res Ethics* **10**, 264–277 (2015).
176. van Delden, J. J. & van der Graaf, R. Revised CIOMS International Ethical Guidelines for Health-Related Research Involving Humans. *JAMA* **317**, 135–136 (2017).
177. Nembaware, V. et al. A framework for tiered informed consent for health genomic research in Africa. *Nat. Genet.* **51**, 1566–1571 (2019).
This paper proposes a framework for the conduct of ethically sound, tiered informed-consent processes in Africa. This framework guarantees the autonomy and individual choices of African research participants but at the same time enables global health benefits gathered from sharing and meta-analysis of African genomic data.
178. Knoppers, B. M., Deschenes, M., Zawati, M. H. & Tasse, A. M. Population studies: return of research results and incidental findings policy statement. *Eur. J. Hum. Genet.* **21**, 245–247 (2013).
179. Fernandez, C. V., Kodish, E. & Weijer, C. Informing study participants of research results: an ethical imperative. *IRB* **25**, 12–19 (2003).
180. Wonkam, A. & de Vries, J. Returning incidental findings in African genomics research. *Nat. Genet.* **52**, 17–20 (2020).
181. de Vries, J. et al. Regulation of genomic and biobanking research in Africa: a content analysis of ethics guidelines, policies and procedures from 22 African countries. *BMC Med. Ethics* **18**, 8 (2017).
This comprehensive analysis of 30 ethics regulatory documents from 22 African countries concludes that there is a need for ethics guidelines in African countries to be adapted to the changing science policy landscape. Most pertinent are the principles of openness, storage, sharing and secondary use.
182. Barchi, F. & Little, M. T. National ethics guidance in sub-Saharan Africa on the collection and use of human biological specimens: a systematic review. *BMC Med. Ethics* **17**, 64 (2016).
183. Ramsay, M., de Vries, J., Soodyall, H., Norris, S. A. & Sankoh, O. Ethical issues in genomic research on the African continent: experiences and challenges to ethics review committees. *Hum. Genomics* **8**, 15 (2014).

184. van Panhuis, W. G. et al. A systematic review of barriers to data sharing in public health. *BMC Public Health* **14**, 1144 (2014).
185. Qureshi, N., Modell, B. & Modell, M. Timeline: raising the profile of genetics in primary care. *Nat. Rev. Genet.* **5**, 783–790 (2004).
186. Alwan, A. & Modell, B. Recommendations for introducing genetics services in developing countries. *Nat. Rev. Genet.* **4**, 61–68 (2003).
187. Bentley, A. R., Callier, S. L. & Rotimi, C. N. Evaluating the promise of inclusion of African ancestry populations in genomics. *NPJ Genom. Med.* **5**, 5 (2020).
188. Ramsay, M. Africa: continent of genome contrasts with implications for biomedical research and health. *FEBS Lett.* **586**, 2813–2819 (2012).
189. Mulder, N. J. et al. Development of bioinformatics infrastructure for genomics research. *Glob. Heart* **12**, 91–98 (2017).
190. Gootenberg, J. S. et al. Nucleic acid detection with CRISPR–Cas13a/C2c2. *Science* **356**, 438–442 (2017).
191. Wilkinson, M. D. et al. The FAIR guiding principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
192. Samuels, D. C. et al. Heterozygosity ratio, a robust global genomic measure of autozygosity and its association with height and disease risk. *Genetics* **204**, 893–904 (2016).
193. Schlebusch, C. M. et al. Genomic variation in seven Khoe-San groups reveals adaptation and complex African history. *Science* **338**, 374–379 (2012).
194. Chatterjee, N., Shi, J. & Garcia-Closas, M. Developing and evaluating polygenic risk prediction models for stratified disease prevention. *Nat. Rev. Genet.* **17**, 392–406 (2016).
195. Krapohl, E. et al. Multi-polygenic score approach to trait prediction. *Mol. Psychiatry* **23**, 1368–1374 (2018).
196. Martin, A. R. et al. Human demographic history impacts genetic risk prediction across diverse populations. *Am. J. Hum. Genet.* **100**, 635–649 (2017).
197. De La Vega, F. M. & Bustamante, C. D. Polygenic risk scores: a biased prediction? *Genome Med.* **10**, 100 (2018).
198. Reisberg, S., Iljasenko, T., Lall, K., Fischer, K. & Vilo, J. Comparing distributions of polygenic risk scores of type 2 diabetes and coronary heart disease within different populations. *PLoS ONE* **12**, e0179238 (2017).
199. Retshabile, G. et al. Whole-exome sequencing reveals uncaptured variation and distinct ancestry in the southern African population of Botswana. *Am. J. Hum. Genet.* **102**, 731–743 (2018).

Acknowledgements
L.P. is a principal researcher at i3S, which is financed by Fundo Europeu de Desenvolvimento Regional (FEDER) funds through COMPETE 2020, Portugal 2020 and by Portuguese funds through Fundação para a Ciência e a Tecnologia/Ministério da Ciência, Tecnologia e Inovação (POCI-01-0145-FEDER-007274). L.M. is a Professor of Human Genetics at the University of Rwanda and principal investigator of a National Institutes of Health (NIH)-funded Human Heredity and Health in Africa (H3Africa) grant (U01MH115485) (transgenerational epigenomics of trauma and post-traumatic stress disorder in Rwanda). P.T. is a senior lecturer at the University of Ghana School of Public Health and the co-principal investigator of an NIH-funded H3Africa grant (U54HG010275). M.R. holds a South African Research Chair in Genomics and Bioinformatics of African populations hosted by the University of the Witwatersrand, funded by the Department of Science and Innovation, and administered by the National Research Foundation and is principal investigator of an NIH-funded H3Africa grant (U54HG006938). The views expressed in this Review do not necessarily reflect the views of the funding institutions.

Author contributions
The authors contributed to all aspects of the article.

Competing interests
The authors declare no competing interests.

Peer review information
Nature Reviews Genetics thanks N. Hanchard, A Wonkam and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Publisher's note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

RELATED LINKS
54gene: https://www.54gene.com/
Africa CDC: http://www.africacdc.org/
African Society of Human Genetics (AfSHG): https://www.afshg.org/
H3Africa: https://h3africa.org/
H3Africa Genotyping Chip: https://www.h3abionet.org/h3africa-chip
H3Africa Guideline for the Return of Individual Genetic Research Findings: https://h3africa.org/wp-content/uploads/2018/05/H3Africa%20Feedback%20of%20Individual%20Genetic%20Results%20Policy.pdf
United Nations Sustainable Development: https://sustainabledevelopment.un.org/
United Nations World Population Prospects: https://population.un.org/wpp/
West African Genetic Medicine Centre (WAGMC): https://wagmc.org
World Economic Forum Precision Medicine Programme: https://www.weforum.org/communities/precision-medicine