The interest of informative ancestry markers (AIM) and their fields of application

Lamiaa Mekhfi [1,2,*], *Bouchra* El Khalfi [1], *Rachid* Saile [3], *Hakima* Yahia [2], *and Abdelaziz* Soukri [1]

[1]Laboratory of Physiopathology, Molecular Genetic and Biotechnology, Ain Chock Faculty of Sciences, Health and Biotechnology Research Center, Hassan II University, Casablanca, Morocco.
[2]National Laboratory of Scientific and Technical Police, Institute of Forensic Sciences of National Security, Casablanca, Morocco.
[3]Laboratory of Biology and Health, Faculty of Sciences Ben M'sik, Health and Biotechnology Research Center, Hassan II University, Casablanca, Morocco,

Abstract: This review focuses on the study of biogeographic ancestry using the Accurate Ancestry Identification Panel. Autosomal markers may provide little information about the nature of an individual's admixture due to ongoing human recombination and migration. Biogeographic ancestry assessment (BGA) is a term used to describe ancestry through DNA testing. This is usually accomplished by testing specific regions of DNA called ancestry information markers (AIMs). AIMs are chosen because they expose significantly different frequencies between different populations in different parts of the world. The panels of these AIMs can be assessed using next-generation sequencing (NGS) to predict the geographical origins of a person of interest's ancestors, usually in terms of continent of origin, and sometimes by smaller geographic regions. The use of ancestry informative markers (AIM) to identify genomic ancestry can be useful for a variety of studies in evolutionary genetics, biomedical research, and forensic analyses. However, there remains a major challenge in determining AIMs for populations with complex and highly mixed ancestry.

Key words: Ancestry informative markers (AIM), Biogeographic ancestry, SNP, Forensic, Biomedical.

1. Introduction

The Maxam Gilbert and Sanger methods made DNA sequencing possible [1], [2]. However, these approaches were slow and expensive. Early sequencers allowed for low-coverage sequencing and long reads of up to 900 nucleotides. The addition of a DNA cloning step (replaced by a PCR step) resulted in more but shorter sequencing fragments. This is what gave rise to the next-generation or second-generation sequencers. NGS have significantly increased sequencing coverage since 2007 while offering shorter reads (36 to 500 nucleotides).

High-throughput sequencing quickly attracted the attention of anthropological researchers, whose primary goal was to research and determine the characterization and variability of human populations. The exploitation of human genetic polymorphism shaped by population genetics data is an objective of anthropology. As a result, gene frequencies vary between different populations as a result of this exploitation [3].

NGS can be used in a variety of fields, such as basic or transactional research [4]. The medical-scientific world is enthusiastic about these methods, as evidenced by the numerous studies that have been published in this field. These include genomic studies (massive de novo genome sequencing, genome "re-sequencing" to analyze variations such as polymorphisms or SNPs, insertions/deletions or translocations), transcriptomic studies (mRNA sequencing, miRNA sequencing) or epigenetic studies (mapping of DNA methylation sites, mapping of protein binding sites to chromatin). The major disciplines applying second-generation sequencing methods are medicine and forensics.

The development of genotyping methods and Next-Generation Sequencing has made it easier to study genetic variation and structure, which is critical to understanding our evolutionary history. Extensive genomic data are also readily available. Ancestral informative markers (AIMs) are a set of informative SNPs with significant differences in allele frequency between ancestral populations and are commonly used to determine genomic ancestry. Determining ancestry therefore examines the frequency of an allele in a population, not its complete presence or absence [5]–[10]. The determination of AIMs was conducted for the purpose of determining population stratification and biogeographic location of discrete continental or global populations [11]. Establishing

*Corresponding author: mekhfi2hiba@gmail.com

population substructure and measuring ancestry are especially important for association studies to elucidate the genetic etiology of complex problems and many factor disorders [12].

Several early investigations utilized the Human Genetic Diversity Panel (HGDP), which consists of more than 1,100 DNA samples from 51 diverse populations across the world [14]. For example, Through the examination of 987 microsatellites found in the HGDP collection, a total of six population clusters were identified, aligning with different continental regions. The analysis of approximately 642,000 autosomal SNPs within the HGDP collection not only allowed for the grouping of individuals within these broad geographic areas but also within specific populations residing within these regions [13],[15]. The forecasts available for each of the six populations are as follows [16]:

- European
- Sub-Saharan Africa
- Northeastern Africa that is, Ethiopia, Eritrea, Somalia, etc.
- Eastern Asian (China, Korea, Japan)
- South Asian (Pakistan, India, Bangladesh, etc.)
- Southeast Asia (Thailand, Malaysia, Indonesia)

In the analysis of an individual's genome, it is often represented as a mosaic of segments (Figure.1) that are believed to originate from either one ancestral population or the other, and in some cases, from both if the maternal and paternal alleles derive from distinct ancestral populations. It is now possible for mixed groups to carry out mapping using mixture linkage disequilibrium [17]. When multiple loci are inherited from both parents, autosomal markers represent a greater portion of the genome's history than mtDNA or a single locus inherited on the Y chromosome. However, genomes are limited, and any given person's genome fragment represents only a small portion of their ancestors, and not every ancestor necessarily passes on a specific genome fragment of their DNA to their descendants, and one can only obtain limited information about the origins of an ancestor individual [18].

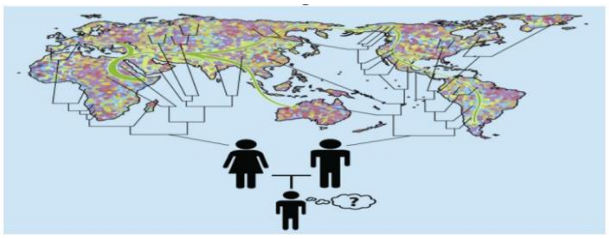

Fig.1. General ancestry. The arrows represent the move of prior human ancestors from Africa. The colorful mosaic represents global community diversity resulting from inter-continental, intra-continental and regional migration. Individual genetic genealogy testing is based on family trees, which are complex networks of mid to recent ancestors. [19]

1.1. Defining Ancestry Informative Markers

Ancestry informative markers are autosomal markers of single nucleotide polymorphism SNPs. A collection of 165 AIM-SNPs prior to selected by two laboratories [9, 10] is inserted in the Precision ID ancestry description panel, which is commercially available by Thermo Fisher Scientific (Waltham, TFS, USA,MA). The average amplicon length of these markers is 120 to 130 bp, which should allow the processing of difficult forensic specimens, such as highly degraded and low-contribution specimens [22]. AIMs, which are most fluently used to estimate the proportions of mixtures from European, Native American, Asian and African populations, provide increased power for inferring ancestry compared to a aleatory set of autosomal markers [23]. Therefore, throughput can be increased and genotyping costs can be reduced by using a smaller set of markers. [19].

2. Fields of application

2.1. Forensic context

The difficulty of identifying unknown human remains is frequently encountered by forensic science. The usual techniques by which this identification is carried out are fingerprint comparison, anthropological examination, X-rays and DNA profiling. STR autosomal is the method of body identification in the absence of specific physical characteristics. The Comparisons are made between the DNA profile and other reference profiles or profiles found in national genetic data repositories. In some cases, the body remains unidentified because the identical or related profiles are not present in the databases [24]. Ancestry informative markers analysis (AIM) can be used to infer a characteristic of genetic ancestry. Using these indicators, it is possible to define a person's genotype as being influenced in varying proportions by major continental populations [20], [21], [25]–[27].

In addition, the forensic genetics community is dedicated to the development of DNA tests that can facilitate an investigation with additional intelligence data beyond traditional identification based on STRs, it should be emphasized that STR profiling combined with thorough application of ancestry analysis has the potential to yield vital information for the advancement of the investigation and the hunt for targets [25]. Predicting an individual's geographic origin can provide useful information for conducting an investigation. Knowing a victim's ancestry could be a valuable lead, whether it's a blood stain left behind by a fleeing perpetrator or an unidentified victim's body part (Figure.2). Criminal investigations evolve from "passive comparison" to "active comparison" thanks to the characteristics inferred from DNA analysis [12]. A SNP panel has significantly different frequencies between different populations in different geographic regions of the world, especially when insufficient and/or poor-quality genomic material is obtained. AIMs can also be useful for gathering ancestry information in a variety of scenarios, including association studies, forensic and genetic assessments of populations[24] .

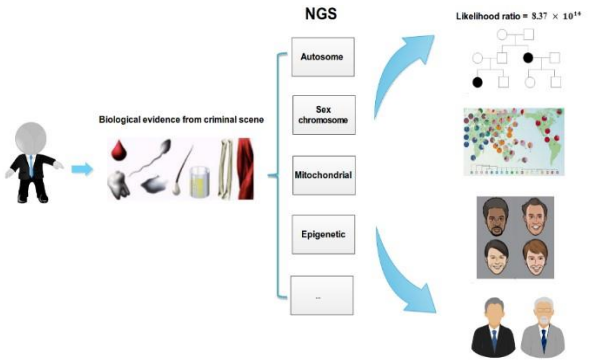

Fig.2. The NGS method can yield a wide variety of information from biological evidence samples taken from crime scenes. Biological evidence samples taken from crime scenes can yield several results at once when NGS technology is applied. These results include STRs, epigenetic information, and single nucleotide polymorphisms (SNPs) of autosomes, sex chromosomes, and mitochondrial genomes. The evidence samples can be used to deduce the physical, psychological, and geographic features of criminal suspects as well as the source population by integrating all the available data, in addition to identifying the suspects.[12]

Due to their stability, density of distribution, and full range of allele frequency patterns within populations, autosomal SNPs have become among the best markers of ancestry [28]. Therefore, it is essential to identify the small number of SNPs with the most pronounced allele frequency discontinuities between continental regions in order to create marker sets with "diagnostic" population phenotypes [29]. To assist in SNP localization, one approach is to study genetic variation that has been subject to strong positive regional selection in the recent past, resulting in localized adaptations [30]. In 2007, a 34-plex SNP assay was developed using primer extension chemistry from Applied Biosystems (AB), with the AIM of developing a sufficiently powerful, least error-prone single-tube SNP assay based on the most recent AIM, with the AIM of (1) selecting SNPs that, in the first place, provide clear differentiation between subsamples of African population groups, Europe and Asia. (2) validate allele frequencies to ensure that variation within a demographic group represents only a small proportion of the total variation; (3) Balance the final chromosomal distribution to avoid a linkage imbalance between SNP pairs (4) Establish a simple Bayesian system to predict ancestral origin and estimate error rates in classification using statistical methods, as well as test the CEPH-HGDP panel of genome diversity cell lines human (CEPH-HGDP) containing geographically confirmed samples [14], [28].

A revision of this established forensic ancestry test of 34 SNPs (table.1) was performed in 2012, replacing the underperforming SNP component rs727811 with the highly informative rs3827760, which shows an almost fixed allele specific to Southeast Asia [31]. Since all current tests for physical characteristics and legal ancestry use AB SNaPshot, the focus should be on optimizing the circumstances of the test for the seed extension in order to continue the successful use of SNP in legal studies.

Therefore, it is possible to obtain the greatest amount of information from any biological remains, which will facilitate the interpretation of the mixture and increase the statistical weight of evidence. Thus, NGS technology would increase the efficiency and cost-effectiveness of legal cases by greatly facilitating the identification of mixed DNA samples and the analysis of complex paternity cases. Well-known examples:

- The Madrid bombings in 2004. The number of individuals was determined with the Bayesian classifier [32]. The main SNP analysis, with 34plex was validated to attribute the descent of DNA from seven unexplained case samples, a sample taken of a bag containing undetonated explosives, and personal items recovered from various locations connected to the suspects. In four cases, these profiles yielded probabilities providing a clear attribution of ancestry. One of the suspects, presumed to be North African according to AIM-SNP analysis of toothbrush DNA, was identified late in the investigation as being of Algerian origin [25].

- In 2014, a fully charred body was discovered in a municipal landfill near Paris. Although the individual X_ind was identified as an adult male, no direct physical description was possible. It was not possible to physically identify his remains due to his condition. A judicial investigation has been launched. Autopsy, dental examination and genetic analysis did not reveal any discriminating features. The biological profile was also supplemented by anthropological and radiological analyses, but no specific information was found. The color of X_ind brown eyes and its likely origin in the Mediterranean basin or Southwest Asia were revealed by further genetic analyses involving markers used for biogeographic studies and the determination of phenotypic traits [24].

Table 1: SNP data in 34-plex composite [31]

Markers		Primers		
Internal code	dbSNPrs-number	PCR forward primer	PCR reverse primer	SNaPshot single base extension (SBE) primer
P02	rs5997008	GTCAACACTAGAGTATTTGCCCATC	ACAAACCCAAAGACTGTTCTGC	gac[aactaggtgccacgtcgtgaaagtctgac]2 aactctcaCAGGATCGATTGGTTCC
P01	rs2304925	CCCATTAACTCATCAAAGTGGTGAT	CCCCACTCCACCGCTAAT	[aaagtctgacaactaggtgccacgtcgtg]2 aaagtctgacaaCCACTCCACCGCTAAT
A07	rs917118	GCCCTTTAGGGTCGGTTC	GTAAGAGATGACTGAGGTCAACGAG	t[ct]2TGACTGAGGTCAACGAGC
P03	rs1321333	GTCAGTAAGACGGTAACTCC	CTAACACAAGCCTAAATCCAG	AAGACGGTAACTCCATGGCTG
P04	rs2814778	AACCTGATGGCCCTCATTAGT	ATGGCACCGTTTGGTTCAG	agtctgacaactaggtgccacgtcgtgaaagtc tgacaactaggtgccacgtcgtgaaagtctgacat CTCATTAGTCCTTGGCTCTTA
A29	rs1024116	CCATGTGTTCTAATAAAAAGGATTGC	TGGGAAGTGAGCAAAAGTAAATACA	cttGTTCTAATAAAAAGGATTGCTCAT
P05	rs7897550	CGATGTGTCTTACGGAATACTAGGT	AGAGCTGACAGGCAAAAATGCTAT	t[ct]2TGTGCAGGATTGAAATATAATT
A21	rs722098	GGAAGTACACATCTGTTGACAGTAA TGA	GGGTAAAGAAATATTCAGCACATCC	agtctgacaaTGACAGTAATGAAATATCCTT G
P06a	rs10843344	TGTACAATGGTAGATGTGTGCTCAG	GATAGCTCTGGTGTTGCATTATTGT	t[ct]5AGTACTTTGCCAAAGAAACTAAA
P08	rs12913832	ACGTTGGATGCGAGGCCAGTTTCATT TGAG	ACGTTGGATGAAAACAAAGAGA AGCCTCGG	[ct]8CCAGTTTCATTTGAGCATTAA
P07	rs239031	TAGCTGTGAGATAGAAATCCTGGAC	ACTACCCTAATCTCAGCTTCCACTC	t[ct]8cAATCTCAGCTTCCACTC
P09a	rs1978806	AGAGTTTGACATGATGGTGCTCTA	TCTTGTTTCTAAGCAGGAAAGTTG	t[ct]8cGCAGGAAAGTTGTATTCTGATA
A40	rs2040411	TCTGGAATGCCAGTTCTTTTGT	CAGAACGCCTATGAAAACCAGT	[ct]7cCTCTGTATTTTCTTACTCTAAGTG C
P10	rs773658	ACAAACGGAAAGTAGTATTGGACTG	AGAAGGGGCACAGCAATTTAGTA	[ct]11cGGGAAGAATAGAGTCAATCAA
P11	rs10141763	ACAGACTTGGTTCCCTGAAGTCTA	GTAGATTGTAGGCAAGTCGTAAAGG	t[ct]10cGTGTGAGTTGTGTGATAATCTA
P12	rs182549	AAGTACTGGGACAAAGGTGTGAG	AGAAGTCAGAATACCCCTACCCTAT	[ct]16cAGGTGTGAGCCACCG
P13	rs1573020	CTATCTGCCACCTGAGAGAGTATTG	AGGTGTCAGCTTCTTCTGACCAT	[ct]14GAGTATTGCCAGCCTGATTC
P14	rs896788	GTAATGCCTCTGTGGCCCTAT	ATTCCGTCCACATCTTCACTG	[ct]17tctACAGTCACCAGCCAC
P15	rs2065160	AAGAATGGCCTCTCGATGAGTA	GATGATACCTACGCATAGTCTGTTTA CTTC	t[ct]14GCATAGTCTGTTTACTTCATTTG
P16a	rs2572307	GTGTAGCTATGCCATCATTCAATC	ATCCTTAGAAGGGTGCTAAACTGAG	t[ct]15catcATTCAATCAATAGTCATAAA C
P17	rs2303798	CCAGCCTGCACCACTGTC	AGAGATGTGTTCAGGAAGAGGCTA	t[ct]19cAGAAGAGGGACGTGGG
P18	rs2065982	CTTGGGGCAGTCTTTAAGTCCT	AGGAAGTGGTCAGTGCCAGTAG	[ct]18GGAAAAAAAAGTCCTCTTTGGTA T
P19	rs3785181	CTCTGTTCAGTTTCAAAGTTCTGG	TTGTGTTCAAAAATTTCAATTAGGTT	t[ct]20AGGGCATCTTATCTTGAGC
P20	rs881929	AGCTACCTGGTGTCTAACTC	TTGACCCAGTGGTTCTGAGC	[ct]7aaactaggtgccacgtcgtgaaagtctgacaa GCCTCTTGCCAGCTCTG
P21	rs1498444	GGCTATTACCACATTAAGAGAAACT GC	CAGCCTCTCAATGCAAATGAT	t[ct]20cGGTTGTCAGAATATTTGCTACA
P22a	rs1426654	AATTCAGGAGCTGAACTGCC	TGTTCAGCCCTTGGATTGTC	[ct]24ttCGCTGCCATGAAAGTTG
P23	rs2026721	GAAGACTTTTTGCAAGCACGAG	GGCAAATGCTGTAAGAATCCAT	t[ct]21AAATGATTGACATAGTAGGCTAT TG
P24	rs4540055	TGTGCCTCTGATCACTTTTGAATAC	CCTAGCCAACTCCAGAGTTCAT	[ct]27cGAAGCAGTGATCAGCAC
A52	rs1335873	GTGGATGATATGGTTTCTCAAGG	TTCAACAAACGTGTGATGCTCT	t[ct]25AGGTACCTAGCTATGTACTCAGT AT
P25a	rs16891982	GAATAAAGTGAGGAAAACACGGAGT	GTTTCTCATCTACGAAAGAGGAGTC	[ct]28GGTTGGATGTTGGGGCTT
P26	rs730570	CAGCACCCTGTAAAGTCCAG	CAGCACTCACCTGCATCTCA	t[ct]27tCATTAATCACACAAATTTTGCAT
A13	rs1886510	GTCCTTGTCAATCTTTCTACCAGAG	GGATTTTCACAACAACACTTGC	[ct]27ACAAGATTTTCACAACAACACTT GC
P27	rs5030240	CCAAAGTGCCAGGATCACAG	TCCCTAGAAATCCTTCAGCC	[ct]32cCACAGGAGTGAGCCACTGC
P28	rs3827760	GCTCAGCTCCACGTACAAC	CTGTCATGCCCCCAATCTC	tctgacaactaggtgccacgtggtgaaagtctgacaa ctaggtgccacgtcgtgaaagtctgacaactctca GGYGCCAYGTTTTCACAb

2.2. Biomedically

Studies of genetic ascendancy in medical research have added to our understanding of genetics by identifying new loci associated with traits and diseases that traditional association studies would not have been able to identify. [33]. Studies of associations will enable more effective personalized medicine development, prevent results that could lead to incorrect interpretations, and enable risk prediction based on medication responses [34]. The NGS approach could provide a more rapid tool for determining the clinical course of patients, providing indications of the likelihood of survival, recurrence or favorable response to treatment. Beyond experimental evidence of feasibility, larger, independent, multicenter validation studies are and

will be needed to determine whether in-depth genomic characterization and dynamic monitoring of tumor clones can lead to better management of cancer patients [35]. NGS holds exceptional potential to advance personalized medicine.

The relationship between genetic ancestry and individual and population health is a potentially important area of investigation because it could have social and political consequences [36], [37]. AIM panels may be useful for controlling for population substructure, in the context of association studies on multifactorial problems [39].

Previous research has demonstrated how AIM panels intended for mixed populations can provide information on the connections between individual descent and specific characteristics, including breast cancer [40], hypertension [24, 25], skin pigmentation [41] and type II diabetes [42]. It is noticeable that among the top twenty most "informative" SNPs (InfocalcInscores) found in current analyses are heart rate (rs11931264)[43], Blood Glucose (rs1516510) [44], High Blood Pressure (rs225555) [45], Triglyceride levels in type II diabetes (rs2240466) [44], abdominal fat in women (rs7927727) [45], and cognitive processing (rs2839627) [46]. Four SNPs contained in our AIM panel (rs4858613, rs4377353, rs931885, and rs242105) had previously been shown to be ancestry-informative for South Central Asian groups in a study conducted by Dr. Petros Drineas' group at Purdue University without surveying any Indian tribes. [46].

The COVID-19 pandemic has raised ethical questions in medical, political and social decision-making. Although scientific studies are being conducted on a global scale, medical professionals are not yet aware of its specific nature. When infected with the coronavirus, some people only have mild or hard-to-detect symptoms. Others have severe symptoms, struggle for weeks to breathe, and some die [47]. However, there is a large difference in the degree of infection of the coronavirus pandemic between different ethical groups (Figure.3). The Covid-19 pandemic has revealed the importance, diversity and relevance of bioethics. Bioethicists can help to see things more clearly, they come from various fields (philosophy, social sciences, law, medicine, nursing). However, they have often played a crucial role in the current crisis [48].

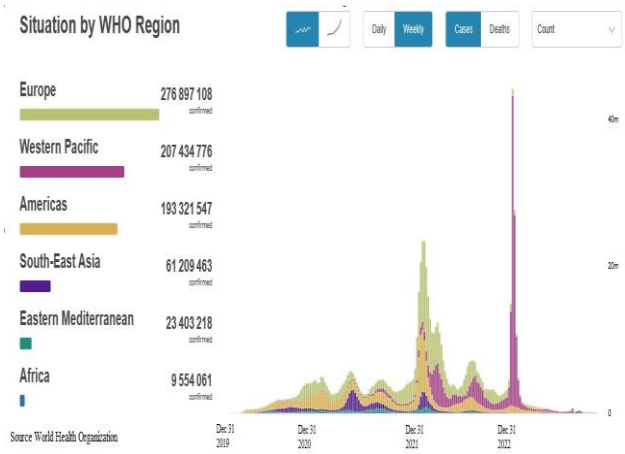

Fig. 3. Who graph shows data to track the geographic location of the pandemic between different ethical groups. [49]

The current genetic resources need to be improved. especially for the populations in Africa and the Middle East or where there aren't enough data. Additionally, keep in mind that given the high rate of consanguinity that may be overlooked in other demographic groups, the genetic characteristics of Arab populations may present a priceless opportunity [50]. The National Arab Genome Project at the Arab United Emirates (EAU) is one of the efforts that aims to increase Arab geographical representation through the use of NGS technologies [51].

2.3. Anthropologically

It is essential to comprehend the genetic roots of ethnic groups in order to address the social and cultural differences that give rise to socioeconomic disparities. These approaches enable us to comprehend the ways in which marriage models interact with historical and cultural factors like epidemics and colonization to affect human variety. During the comparison, one might choose to examine various character types, such as physiology, morphology, behavior, and/or ecology [52]. An example of the first case is the use of post-bregmatic depression variant of the skull trait [37, 38].

In anthropology, descent is often used as a marker of social race. We examined this idea and its implications through a content analysis of Journal of Forensic Sciences publications from 2009 to 2019 [54].

3. Conclusion

Estimates of genetic ancestry are based on haploid markers (Y chromosome haplotypes or mitochondrial DNA [mtDNA]) or diploid autosomal markers, which are sometimes chosen because they "provide information about ancestry." In epidemiology, ancestry estimates are often based on allele frequencies of autosomal SNPs. Both types of markers are used by population geneticists and anthropologists. Sequences from the mtDNA D-loop region or SNP haplotypes on the Y chromosome allow direct comparison of lineages. Unlike probabilistic estimates, exact matches between haploid genetic markers

are easy to understand: if one person's Y chromosome haplotype matches another person's Y chromosome haplotype, this indicates a common paternal ancestor [19].

In addition, Although mtDNA and Y chromosome markers often show informative geographic differences, special attention is needed if comparisons of uniparental marker panels are chosen because both regions share a common mtDNA and Y chromosome lineage arising from a previous gene [56]. As a result, it must be distinguished that estimates of the addition of residents movements to local change patterns vary across markers and studies [57].

Our worldview needs to change thoughtfully, and we need to take a fundamental approach. To do this, we must carefully select informative characters to study and formulate hypotheses based on meaningful questions. It's time to pick up where we left off and look at actual biological differences in humans [53]. There is a growing concern in the ancestral histories of other geographical regions, as well as in the deeper time scales of populations in Africa and India, which have identified deep divisions between geographical regions exhibiting complex patterns of past migration and mixing. Clearly, more research on the genetic variety of humans is required; These research will identify potential global distributions of these scaled geographic patterns and provide these patterns with interpretive values [58].

Limiting the Use of AIM Ancestry Informative Markers

- It's crucial to remember that the diversity of human social structures, mixed-race marriage models, and demographic histories make it difficult to resolve the population's genetic makeup. The problem of disentangling recently mixed ancestry becomes more complex when people who have recently migrated among their ancestors are included.
- AIM evaluation methods and génotypage panels at the génome scale can be used to quantify autosomal variation. Notwithstanding the possibility that the idea could herald the acceptance of genetic anthropology, whole genome sequencing remains prohibitively expensive and is therefore out of reach for most academic researchers or commercial tests.
- While the HGDP Human Genetic Diversity Panel collection is useful for bringing together widely dispersed human populations, it does not allow for dense sampling in a particular geographic region. For this reason, the accuracy of inference about ancestry within and across regions is limited [19].

References

1 W. Gilbert, A. Maxam, *The Nucleotide Sequence of the lac Operator*, PNAS, **70**, 12, 3581–3584, (Dec. 1973), doi: 10.1073/pnas.70.12.3581.

2 F. Sanger, S. Nicklen, A. R. Coulson, *DNA sequencing with chain-terminating inhibitors*, PNAS, **74**, 12, 5463–5467, (Dec. 1977).

3 C. Terzian, C. Biémont, *Les théories de l'effet fondateur, Mayr, et la génétique des populations*, GSE, **20**, 1, 111, (1988).

4 S. Pääbo, *The mosaic that is our genome*, N, **421**, 6921, (Jan. 2003).

5 M. Daya, L. van der Merwe, U. Galal, M. Möller, M. Salie, E.R. Chimusa, J.M. Galanter, P.D. van Helden, B.M. Henn, C.R. Gignoux, E. Hoal, *A Panel of Ancestry Informative Markers for the Complex Five-Way Admixed South African Coloured Population*, PLSONE, **8**, 12, e82224, (Dec. 2013).

6 K. Vongpaisarnsin, J. B. Listman, R. T. Malison, J. Gelernter, *Ancestry informative markers for distinguishing between Thai populations based on genome-wide association datasets*, LM., **17**, 4, 245–250, (Jul. 2015).

7 H.C. Santos, A.V.R. Horimoto, E. Tarazona-Santos, F. Rodrigues-Soares, M.L. Barreto, B.L. Horta, M.F. Lima-Costa, M.H. Gouveia, M. Machado, T.M. Silva, J.M. Sanches, N. Esteban, W.CS. Magalhaes, M.R. Rodrigues, F.S.G. Kehdy, A.C. Pereira, The Brazilian EPIGEN Project Consortium, *A minimum set of ancestry informative markers for determining admixture proportions in a mixed American population: the Brazilian set*, EJHG, **24**, 5, 725–731, (May 2016).

8 L. Mendoza, D. P. Aguirre, J. J. Builes, *Ancestry evaluation of an Afro-descendant population sample of the department of Chocó-Colombia*, FSIGSS, **6**, e292–e293, (Dec. 2017).

9 R. Santangelo, F. González-Andrade, C. Børsting, A. Torroni, V. Pereira, N. Morling, *Analysis of ancestry informative markers in three main ethnic groups from Ecuador supports a trihybrid origin of Ecuadorians*, FSIG **31**, 29–33, (Nov. 2017).

10 Y. Wang, D. Lu, Y.-J. Chung, S. Xu, *Genetic structure, divergence and admixture of Han Chinese, Japanese and Korean populations*, H, **155**, 1, 19, (Dec. 2018).

11 R. Das, P. Upadhyai, *An Ancestry Informative Marker Set Which Recapitulates the Known Fine Structure of Populations in South Asia,* GBE, **10**, 9, 2408–2416, (Sep. 2018).

12 Y. Yang, B. Xie, J. Yan, *Application of Next-generation Sequencing Technology in Forensic Science*, GPB, **12**, 5, 190–197, (Oct. 2014).

13 J.Z. Li, D.M. Absher, H. Tang, A.M. Southwick, A.M. Casto, S. Ramachandran, H.M. Cann, G.S. Barsh, M. Feldman, L.L. Cavalli-Sforza, R.M. Myers, *Worldwide Human Relationships Inferred from Genome-Wide Patterns of Variation,* S, **319**, 5866, 1100–1104, (Feb. 2008).

14 H.M. Cann, C. de Toma, L. Cazes, M.F. Legrand, V. Morel, L. Piouffre, J. Bodmer, W. F. Bodmer, B. Bonne-Tamir, A. Cambon-Thomsen, Z. Chen, J. Chu, C. Carcassi, L. Contu, R. Du, L. Excoffier, G.B. Ferrara, J.S. Friedlaender, H. Groot, D. Gurwitz, T. Jenkins, R.J. Herrera, X. Huang, J.

Kidd, K.K. Kidd, A. Langaney, A.A. Lin, S.Q. Mehdi, P. Parham, A. Piazza, M.P. Pistillo, Y. Qian, Q. Shu, J. Xu, S. Zhu, J.L. Weber, H.T. Greely, M.W. Feldman, G. Thomas, J. Dausset, L. L. Cavalli-Sforza, *A Human Genome Diversity Cell Line Panel*, S, **296**, 5566, 261–262, (Apr. 2002).

15 M. Jakobsson, S.W. Scholz, P. Scheet, J.R. Gibbs, J.M. VanLiere, H-C. Fung, Z.A. Szpiech, J.H. Degnan, K. Wang, R. Guerreiro, J.M. Bras, J.C. Schymick, D.G. Hernandez, B.J. Traynor, J. Simon-Sanchez, M. Matarin, A. Britton, J. van de Leemput, I. Rafferty, M. Bucan, H.M. Cann, J.A. Hardy, N.A. Rosenberg, A.B. Singleton, *Genotype, haplotype and copy-number variation in worldwide human populations*, N, **451**, 7181, (Feb. 2008).

16 *Specialist forensic DNA testing using Next Generation*. Accessed: (Nov. 14, 2023). [Online]. Available: https://cdnmedia.eurofins.com/european-west/media/1929693/17904_ngs_short_guide_-web.pdf

17 M.W. Smith, J.A. Lautenberger, H. Doo Shin, J.P. Chretien, S. Shrestha, D.A. Gilbert, S.J. O'Brien - *Markers for Mapping by Admixture Linkage Disequilibrium in African American and Hispanic Populations* (Nov. 14, 2023). [Online]. Available: https://www.cell.com/action/showPdf?pii=S0002-9297%2807%2961324-X

18 K. M. Weiss, J. C. Long, *Non-Darwinian estimation: My ancestors, my genes' ancestors*, GR., **19**, 5, 703–710, (Jan. 2009).

19 C.D. Royal, J. Novembre, S.M. Fullerton, D.B. Goldstein, J.C. Long, M.J. Bamshad, A.G. Clark, *Inferring Genetic Ancestry: Opportunities, Challenges, and Implications*, AJHG, **86**, 5, 661–673, (May 2010).

20 K.K. Kidd, W.C. Speed, A.J. Pakstis, M.R. Furtado, R. Fang, A. Madbouly, M. Maiers, M. Middha, F. ọise R. Friedlaender, J.R. Kidd, *Progress toward an efficient panel of SNPs for ancestry inference*, FSIG, **10**, 23–32, (May 2014).

21 R. Kosoy, R. Nassir, C. Tian, P.A. White, L.M. Butler, G. Silva, R. Kittles, M.E. Alarcon-Riquelme, P.K. Gregersen, J.W. Belmont, F.M. De La Vega, M.F. Seldin, *Ancestry informative marker sets for determining continental origin and admixture proportions in common populations in America*, HM, **30**, 1, 69–78, (Jan. 2009).

22 A. B. Felkl, E. Avila, A. Z. Gastaldo, C. G. Lindholz, M. Dorn, C. S. Alho, *Ancestry resolution of South Brazilians by forensic 165 ancestry-informative SNPs panel*, FSIG, **64**, (May 2023).

23 I. Halder, M. Shriver, M. Thomas, J. R. Fernandez, T. Frudakis, *A panel of ancestry informative markers for estimating individual biogeographical ancestry and admixture from four continents: utility and applications*, HM, **29**, 5, 648–658, (May 2008).

24 C. Hollard, C. Keyser, T. Delabarde, A. Gonzalez, C. Vilela Lamego, V. Zvénigorosky, B. Ludes, *Case report:on the use of the HID-Ion AmpliSeq™ Ancestry Panel in a real forensic case*, IJLM, **131**, 2, 351–358, (Mar. 2017).

25 C. Phillips, L. Prieto, M. Fondevila, A. Salas, A. Gómez-Tato, J. Álvarez-Dios, A. Alonso, A. Blanco-Verea, M. Brión, M. Montesino, Á. Carracedo, M.V. Lareu, *Ancestry Analysis in the 11-M Madrid Bomb Attack Investigation*, PLSONE, **4**, 8, e6583, (Aug. 2009).

26 C. Phillips, W. Parson, W. Parson, B. Lundsberg, Carla Santos, A. Freire-Aradas, M. Torres, M. Eduardoff, C. Børsting, P. Johansen, M. Fondevila, N. Morling, P. M. Schneider, A. Carracedo, M. Lareu, *Building a forensic ancestry panel from the ground up: The EUROFORGEN Global AIM-SNP set*, FSIG, **11**, 13–25, (Jul. 2014).

27 C. Phillips, A. F. Aradas, A. K. Kriegel, M. Fondevila, O. Bulbul, C. Santos, F.S. Rech, M.D. P. Cárceles, A. Carracedo, P. Schneider, M. Lareu, *Eurasiaplex: A forensic SNP assay for differentiating European and South Asian ancestries*, FSIG, **7**, 3, 359–366, (May 2013).

28 C. Phillips, A. Salas, J.J. Sanchez, M. Fondevila, A. Gómez-Tato, J. Álvarez-Diós, M. Calaza, M. C. Cal, D. Ballard, M. Lareu, A. Carracedo, *Inferring ancestral origin using a single multiplex assay of ancestry-informative marker SNPs*, Forensic Sci. Int. Genet., **1**, 3–4, 273–280, (Dec. 2007), doi: 10.1016/j.fsigen.2007.06.008.

29 N. Rosenberg, J. Pritchard, J. Weber, H. Cann, K. Kidd, L. Zhivotovsky, M. Feldman, *Genetic Structure of Human Populations*, S, **298**, 5602, 2381–2385, (Dec. 2002).

30 J. Akey, M. Eberle, M. Rieder, C. Carlson, M. Shriver, D. Nickerson, L. Kruglyak, *Population History and Natural Selection Shape Patterns of Genetic Variation in 132 Genes*, PLSB, **2**, 10, e286, (Sep. 2004).

31 M. Fondevila, C. Phillips, C. Santos, A. F. Aradas, P. Vallone, John M. Butler, M. Lareu, A. Carracedo, *Revision of the SNPforID 34-plex forensic ancestry test: Assay enhancements, standard reference sample genotypes and extended population studies*, FSIG, **7**, 1, 63–74, (Jan. 2013).

32 R. Van De Schoot, D. Kaplan, J. Denissen, J. B. Asendorpf, F. J. Neyer, M. A. G. Van Aken, *A Gentle Introduction to Bayesian Analysis: Applications to Developmental Research*, CD, **85**, 3, 842–860, (May 2014).

33 E. Suarez-Pajes, A. Díaz-de Usera, I. Marcelino-Rodríguez, B. Guillen-Guio, C. Flores, *Genetic Ancestry Inference and Its Application for the Genetic Mapping of Human Diseases*, IJMS, **22**, 13, 6962, (Jun. 2021).

34 G. Sirugo, S. M. Williams, S. A. Tishkoff, *The Missing Diversity in Human Genetic Studies*, C, **177**, 1, 26–31, (Mar. 2019).

35 P. L. Bedard, A. R. Hansen, M. J. Ratain, L. L. Siu, *Tumour heterogeneity in the clinic*, N, **501**, 7467, 355–364, (Sep. 2013).

36 M. Bamshad, *Genetic Influences on Health: Does Race Matter?* JAMA, **294**, 8, p. 937, (Aug. 2005).

37 M. Via, E. Ziv, E. Burchard, *Recent advances of genetic ancestry testing in biomedical research and*

direct to consumer testing, CG, **76**, 3, 225–235, (Sep. 2009).
38 R. S. Cooper, B. Tayo, X. Zhu, *Genome-wide association studies: implications for multiethnic samples*, HMG, **17**, R2, R151–R155, (Oct. 2008).
39 J. Pardo-Seco, F. Martinón-Torres, A. Salas, *Evaluating the accuracy of AIM panels at quantifying genome ancestry*, BMCG, **15**, 1, p. 543, (Jun. 2014).
40 S.J. Serrano-Gómez, M. C. Sanabria-Salas, J. Garay, M. Baddoo, G. Hernandez-Suarez, J. Mejía, Ó. García, L. Miele, L. Fejerman, J. Zabaleta, *Ancestry as a potential modifier of gene expression in breast tumors from Colombian women*, PLSONE, **12**, 8, e0183179, (Aug. 2017).
41 M. Shriver, E. Parra, S. Dios, C. Bonilla, H. Norton, C. Jovel, C. Pfaff, C. Jones, Aisha Massac, N. Cameron, A. Baron, T. Jackson, G. Argyropoulos, L. Jin, C. Hoggart, P. McKeigue, R. Kittles, *Skin pigmentation, biogeographical ancestry and admixture mapping*, HG, **112**, 4, 387–399, (Apr. 2003).
42 E. J. Parra, *Relation of type 2 diabetes to individual admixture and candidate gene polymorphisms in the Hispanic American population of San Luis Valley, Colorado*, JMG, **41**, 11, e116–e116, (Nov. 2004).
43 M.H. Shahin, D. Conrado, D. Gonzalez, Y. Gong, M. Lobmeyer, A. Beitelshees, E. Boerwinkle, J. Gums, A. Chapman, S. Turner, R. Cooper-DeHoff, J. A. Johnson, *Genome-Wide Association Approach Identified Novel Genetic Predictors of Heart Rate Response to β-Blockers*, JAHA, **7**, 5, e006463, (Mar. 2018).
44 C. Liu, Q. Yang, L. Adrienne Cupples, J. B. Meigs, J. Dupuis, *Selection of the most informative individuals from families with multiple siblings for association studies*, GE, **33**, 4, 299–307, (May 2009).
45 A. C. Morrison, E. Boerwinkle, S. T. Turner, R. E. Ferrell, *Regional Association-based Fine-mapping for Sodium-Lithium Countertransport on Chromosome 10*, Am. J. Hypertens., **21**, 1, 117–121, (Jan. 2008).
46 P. Paschou, J. Lewis, A. Javed, P. Drineas, *Ancestry informative markers for fine-scale individual assignment to worldwide populations*, JMG, **47**, 12, 835–847, (Dec. 2010).
47 A. Nellore, A. Nguyen, R. Thompson, *Covid-19 : la réaction de notre corps dépendrait de nos gènes – et on commence à comprendre lesquels sont impliqués*, The Conversation. Accessed: (Nov. 15, 2023). [Online]. Available: http://theconversation.com/covid-19-la-reaction-de-notre-corps-dependrait-de-nos-genes-et-on-commence-a-comprendre-lesquels-sont-impliques-137958
48 B. Williams-Jones, S. N. Abtroun, *Covid-19 : comment la bioéthique peut aider à faire face à des choix déchirants*, The Conversation. Accessed: (Nov. 15, 2023). [Online]. Available: http://theconversation.com/covid-19-comment-la-bioethique-peut-aider-a-faire-face-a-des-choix-dechirants-136460
49 *WHO Coronavirus (COVID-19) Dashboard.* Accessed: (Nov. 16, 2023). [Online]. Available: https://covid19.who.int
50 A. N. Abou Tayoun, H. L. Rehm, *Genetic variation in the Middle East—an opportunity to advance the human genetics field*, GM, **12**, 1, 116, (Dec. 2020).
51 M. Al-Ali, W. Osman, G. K. Tay, H. S. AlSafar, *A 1000 Arab genome project to study the Emirati population*, JHG, **63**, 4, 533–536, (Apr. 2018).
52 E. A. DiGangi, J. D. Bethard, *Uncloaking a Lost Cause: Decolonizing ancestry estimation in the United States*, AJPA, **175**, 2, 422–436, (Jun. 2021).
53 A. H. Ross, S. E. Williams, *Ancestry Studies in Forensic Anthropology: Back on the Frontier of Racism*, B, **10**, 7, 602, (Jun. 2021).
54 K. Stull, E. Bartelink, A. Klales, G.E. Berg, M.W. Kenyhercz, E. L'Abbé, M.C. Go, Kyle A. McCormick, C. Ariscal, *Commentary on: Bethard JD, DiGangi EA. Letter to the Editor—Moving beyond a lost cause: Forensic anthropology and ancestry estimates in the United States.* JFS, **66**, 1, 417–420, (2021).
55 A. H. Ross, *Regional isolation in the Balkan region: An analysis of craniofacial variation*, AJPA, **124**, 1, 73–80, (May 2004).
56 S. Plaza, F. Calafell, A. Helal, N. Bouzerna, G. Lefranc, J. Bertranpetit, D. Comas, *Joining the Pillars of Hercules: mtDNA Sequences Show Multidirectional Gene Flow in the Western Mediterranean*, AHG, **67**, 4, 312–328, (Jul. 2003).
57 C. Flores, N. Maca-Meyer, J. A. Perez, M. Hernandez, V. M. Cabrera, *Y-Chromosome Differentiation in Northwest Africa*, HB, **73**, 4, 513–524, (2001).
58 C. Royal, J. Novembre, S. Fullerton, D. Goldstein, J. Long, M. Bamshad, A. Clark, *Inferring Genetic Ancestry: Opportunities, Challenges, and Implications*, AJHG, **86**, 5, 661–673, (May 2010).