Forensic Science International: Genetics Supplement Series 6 (2017) e21–e23

Contents lists available at ScienceDirect

Forensic Science International: Genetics Supplement Series

journal homepage: www.elsevier.com/locate/fsigss

Helping the identification of refugee shipwreck victims in the Straits of Sicily: An AIM-Indel reference database for the Tigray population of Ethiopia

H.R.S. Kumar[a], K. Haddish[a], N. Futwi[b], D. Lacerenza[c], G. Tewelemedhin[a], R.V. Manukonda[a], C. Di Gaetano[d], M. Fondevila[e], C. Phillips[e], C. Robino[c,*]

[a] Department of Forensic Medicine, University of Mekelle, Mekelle, Ethiopia
[b] Tigray Health Research Institute, Mekelle, Ethiopia
[c] Department of Public Health Sciences and Pediatrics, University of Turin, Turin, Italy
[d] Department of Medical Sciences, University of Turin/Italian Institute for Genomic Medicine, Turin, Italy
[e] Forensic Genetics Unit, Institute of Forensic Sciences, University of Santiago de Compostela, Santiago de Compostela, Spain

ARTICLE INFO

Keywords:
AIM
Indel
Ethiopia
DVI

ABSTRACT

One of the worst accidents in the European refugee crises occurred on 18th April 2015, when a migrant vessel sank in the Straits of Sicily, with over 800 deaths, including 350 Eritreans. Ancestry informative markers (AIMs) can contribute to the ongoing DVI process enabling the matching of human remains and reference samples from victim's relatives according to their ancestry. A reference population dataset (n = 228) of 46 AIM Insertion Deletion polymorphisms (Indels), obtained from the Tigray population (the major ethnic group in Northern Ethiopia and Eritrea), proved to be effective at discriminating Tigray and other sub-Saharan African populations. Ancestry inference was also satisfactory in comparison to Middle East, when excluding Northern Africans from this population group, and Central-Southern Asia.

1. Introduction

Tigray is the northernmost regional state of Ethiopia. The Semitic-speaking Tigray people are the fourth largest ethnic group in Ethiopia, reaching up to 4.5 millions, and the major ethnic group of neighboring Eritrea.

The availability of Tigray-specific population frequency data for forensic DNA markers is important in the context of the recent refugee crisis, since Eritrea is one of the major countries of origin of migrants crossing the Mediterranean sea to reach Europe [1].

In a drowning accidents occurred on 18th April 2015 in the Straits of Sicily, over 800 migrants died in a sunken vessel, including people from Syria, Bangladesh, West and East Africa (among them 350 Eritreans) [2]. In 2016, the vessel was raised, autopsies performed on recovered bodies, and DNA samples collected for future testing [3]. Since it is expected that ancestry informative markers (AIMs) can contribute to the identification process, enabling the matching of human remains and reference samples from victim's relatives according to their ancestry, 46 AIM Insertion Deletion polymorphisms (Indels) were investigated in the Ethiopian Tigray population.

2. Materials and methods

DNA samples from 228 volunteer Mekelle University (MU) students with four grandparents of Tigray origin were genotyped with the 46-plex AIM-Indels, as described in [4].

The Tigray sample was compared with relevant populations from a worldwide reference panel including Central South Asia (CSA), Middle East (ME), and sub-Saharan Africa (AFR) [4,5]. Algerian Mozabites included in the ME population group described in [5] were parceled out and complemented with 38 samples from Morocco (Northern Africa, NAFR). Among AFR populations described in [4], Mandinka (Senegal), Yoruba (Nigeria), and Bantu-speaking Kenyans were selected together with 48 additional samples from Western Africa (Ivory Coast). The Snipper 2.5 app suite, STRUCTURE, Structure Harvester and CLUMPP software were used for data analysis [6].

The study was authorized by MU Research Ethics Review Committee (ERC 0841/2016).

3. Results

Results of STRUCTURE analysis are shown in Fig. 1a. At the optimal

* Corresponding author.
E-mail address: carlo.robino@unito.it (C. Robino).

http://dx.doi.org/10.1016/j.fsigss.2017.09.017
Received 25 August 2017; Accepted 10 September 2017
Available online 12 September 2017
1875-1768/ © 2017 Elsevier B.V. All rights reserved.

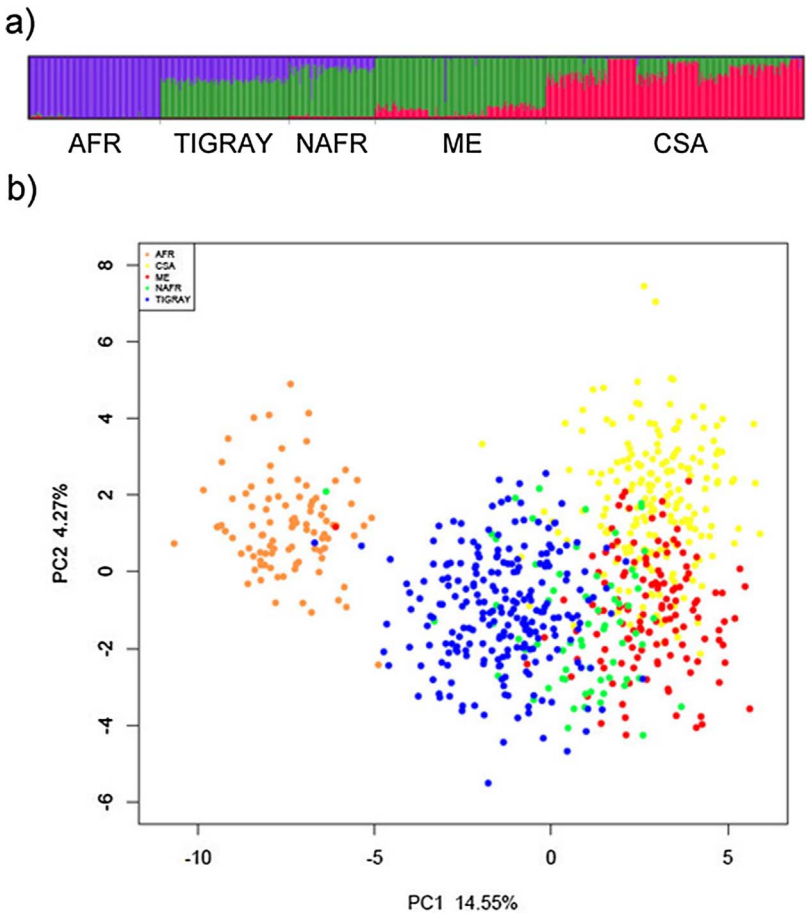

Fig. 1. a) STRUCTURE results (K = 3) considering the admixture LOCPRIOR ancestry model b) PCA results.

K value of K = 3, ancestry proportions in the Tigray sample were highly homogeneous, and well differentiated from other populations, with the exception of NAFR. These results were mirrored in principal component analysis (PCA) (Fig. 1b). The observed patterns also support the separation of NAFR from other ME populations, as previously described in [5].

Re-classification (taking each sample in turn and classifying it using the unmodified population reference panel) and cross validation (classifying each sample with a population reference panel modified to exclude only that sample) tests were performed. Classification accuracy in the Tigray sample was 85.1% (re-classification) and 47.8% (cross-validation), with most of the wrong assignations being NAFR (11.4%, and 43.0%, respectively).

The Tigray sample set was randomly split in training (n = 128) and test (n = 100) sets, the last representing hypothetical case samples. Classification probabilities of case samples were calculated, according to allele frequency distributions found in Tigray and worldwide training sets. Likelihood ratios (LRs) were derived from population pairwise comparisons. Table 1 reports classification outcomes when applying different LR threshold values. Even for small LRs (i.e. 10), the probability of misclassifying Tigray case samples was almost negligible (≤2%). However, the percentage of inconclusive tests was considerable for CSA (9%) and ME (17%), and exceedingly high for NAFR (46%).

4. Conclusions

The 46 AIM-Indels are effective at capturing the genetic differences previously observed between Semitic-speaking Ethiopians and other AFR populations [7]. Ancestry inference is also satisfactory in comparison to ME, when excluding NAFR from this population group, and CSA.

5. Discussion

Although the 46 AIM-Indels alone provide adequate information to differentiate between population groups involved in 18th April 2015 tragedy, typing of additional AIMs [8] in the Tigray population is currently under way and expected to further increase discrimination

Table 1
Classification success of Tigray case samples adopting different LR thresholds for population assignment.

LR	vs AFR			vs NAFR			vs ME			vs CSA		
	Correctly classified	Wrongly classified	Not classified	Correctly classified	Wrongly classified	Not classified	Correctly classified	Wrongly classified	Not classified	Correctly classified	Wrongly classified	Not classified
10	99%	1%	0%	52%	2%	46%	82%	1%	17%	90%	1%	9%
100	98%	1%	1%	16%	1%	83%	69%	0%	31%	79%	0%	21%
1000	94%	0%	6%	6%	0%	94%	45%	0%	55%	67%	0%	33%
10000	92%	0%	8%	1%	0%	99%	25%	0%	75%	51%	0%	49%
100000	85%	0%	15%	0%	0%	100%	11%	0%	89%	29%	0%	71%

capacity.

Role of the funding source

This work was supported by University of Turin WWS2 Project

Conflict of interest statement

None.

Acknowledgments

MU students participating in the study are gratefully acknowledged.

References

[1] T. Brian, F. Laczko, Fatal Journeys Volume 2 Identification and Tracing of Dead and Missing Migrants, International Organization for Migration, Geneva, 2016.
[2] R.M. Townsend, The European Migrant Crisis, Lulu Press Inc., Morrisville, 2015.
[3] V. Piscitelli, A. Iadicicco, D. De Angelis, et al., Italy's battle to identify dead migrants, Lancet Glob. Health 4 (2016) e512–513.
[4] R. Pereira, C. Phillips, N. Pinto, et al., Straightforward inference of ancestry and admixture proportions through ancestry-informative insertion deletion multiplexing, PLoS One 7 (2012) e29684.
[5] C. Santos, C. Phillips, F. Oldoni, et al., Completion of a worldwide reference panel of samples for an ancestry informative Indel assay, Forensic Sci. Int. Genet. 17 (2015) 75–80.
[6] C. Santos, C. Phillips, A. Gomez-Tato, et al., Inference of ancestry in forensic analysis II: analysis of genetic data, Methods Mol. Biol. 1420 (2016) 255–285.
[7] L. Pagani, T. Kivisild, A. Tarekegn, et al., Ethiopian genetic diversity reveals linguistic stratification and complex influences, Am. J. Hum. Genet. 91 (2012) 83–96.
[8] M. de la Puente, C. Santos, M. Fondevila, et al., The global AIMs nano set: a 31-plex SNaPshot assay of ancestry-informative SNPs, Forensic Sci. Int. Genet. 22 (2016) 81–88.