Forensic Science International: Genetics 20 (2016) 45–52

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Allele frequency data for 15 autosomal STR loci in eight Indonesian subpopulations

Samantha J. Venables[a,b,*], Runa Daniel[c], Stephen D. Sarre[d], Nurtami Soedarsono[e,f], Herawati Sudoyo[g], Helena Suryadi[g], Roland A.H. van Oorschot[c], Simon J. Walsh[h], Putut T. Widodo[e], Dennis McNevin[a]

[a] National Centre for Forensic Studies, Faculty of Education, Science, Technology and Mathematics, University of Canberra, Canberra, Australia
[b] School of Science and Health, Western Sydney University, Sydney, Australia
[c] Office of the Chief Forensic Scientist, Victoria Police Forensic Services Department, Melbourne, Australia
[d] Institute for Applied Ecology, University of Canberra, Canberra, Australia
[e] DNA Laboratory, Centre of Medical Health and Services, Indonesian National Police, Jakarta, Indonesia
[f] Department of Oral Biology, Faculty of Dentistry, Universitas Indonesia, Jakarta, Indonesia
[g] Eijkman Institute for Molecular Biology, Jakarta, Indonesia
[h] Forensics, Australian Federal Police, Canberra, Australia

ARTICLE INFO

Article history:
Received 28 April 2015
Received in revised form 25 August 2015
Accepted 27 September 2015
Available online 3 October 2015

Keywords:
Indonesia
Identifiler
Short tandem repeat (STR)
Allele frequency
Population sub-structure
Database

ABSTRACT

Evolutionary and cultural history can affect the genetic characteristics of a population and influences the frequency of different variants at a particular genetic marker (allele frequency). These characteristics directly influence the strength of forensic DNA evidence and make the availability of suitable allele frequency information for every discrete country or jurisdiction highly relevant. Population sub-structure within Indonesia has not been well characterised but should be expected given the complex geographical, linguistic and cultural architecture of the Indonesian population. Here we use forensic short tandem repeat (STR) markers to identify a number of distinct genetic subpopulations within Indonesia and calculate appropriate population sub-structure correction factors. This data represents the most comprehensive investigation of population sub-structure within Indonesia to date using these markers. The results demonstrate that significant sub-structure is present within the Indonesian population and must be accounted for using island specific allele frequencies and corresponding sub-structure correction factors in the calculation of forensic DNA match statistics.

1. Introduction

Understanding genetic population sub-structure within populations and having suitable allele frequency information available is critical to the assessment of the strength of forensic DNA evidence. The presence of genetic sub-structure affects the interpretation of forensic DNA evidence, where the probability of observing a second copy of a particular genotype in a certain population is used to assign weight to the evidence [1]. Forensic practitioners use a relevant population database with allele frequency information to calculate this genotype probability. Since population sub-structure causes individuals to become more genetically related through remote inbreeding, and related individuals are more likely to share a DNA profile than unrelated individuals [2], it causes inaccuracies in the match probability calculations required for likelihood ratios [3]. Frequently, a correction is applied to match probability calculations to accommodate the variations that exist between subpopulations [4].

Population sub-structure within Indonesia has not been well characterised for forensic short tandem repeat (STR) markers and little allele frequency information exists. However, it is conceivable that Indonesia would exhibit some level of genetic sub-structure owing to the distribution of the population across the archipelago and the complex migration and settlement history of the area.

The Indonesian archipelago consists of over 17,000 islands of various sizes. Although approximately 6000 of the islands in the Indonesian archipelago are inhabited, the majority of the

* Corresponding author at: School of Science and Health, Western Sydney University, Sydney, Australia.
E-mail addresses: s.venables@westernsydney.edu.au, sam.venables@hotmail.com (S.J. Venables).

http://dx.doi.org/10.1016/j.fsigen.2015.09.014
1872-4973/

240 million inhabitants are dispersed across a small number of larger islands (i.e. Sumatra, Java, Kalimantan, Sulawesi, Nusa Tenggara and Papua) [5]. Indonesia is home to over 300 traditional groups and 726 languages, of which 719 are living [6,7].

This population diversity is a result of the migration and settlement history of the region. The first major migration into the area occurred approximately 50,000 years ago when the ancestors of modern day Aboriginal Australians and Papua New Guinea highlanders moved into Sahul [8–12]. A second wave of migration began approximately 4000 years ago in which Austronesian language speakers from Taiwan moved into the Pacific and this also had a significant impact on Indonesia [10,12–15]. More recent settlements based on trade relationships (e.g. Chinese, Dutch, Spanish) have also left their mark on Indonesian culture [7,16].

Here we show that genetic population sub-structure exists in Indonesia for forensic short tandem repeat (STR) loci. We use our data to suggest discrete sub-populations, based on island groupings, which minimise deviations from Hardy Weinberg and linkage equilibrium (HWE and LE). This information will advance knowledge of genetic relationships between populations within the Indonesian archipelago and allow for the formation of defensible population databases for estimating random match probabilities. Our data represent the most comprehensive investigation of population sub-structure within Indonesia to date using forensic STR markers.

2. Materials and methods

2.1. Study population

DNA samples from 1500 unrelated Indonesian individuals were obtained by the Eijkman Institute for Medical Research (Jakarta, Indonesia). These samples represent 31 subpopulations within Indonesia and span 22 distinct geographical regions (Fig. 1 and SI Table 1). Blood samples were collected from volunteers according to three criteria; (a) the participants provided informed consent, (b) the participant's family declared the same ancestry for the past three generations and (c) the participants were not related for the past three generations.

To strengthen the data set, we obtained AmpFℓSTR® Identifiler® genotype data for 441 additional individuals from 5 subpopulations (Fig. 1 and SI Table 1) from the Indonesian National Police (INP) for inclusion in the data analysis phase of this study.

Ethics approval for this project was obtained from the University of Canberra Committee for Ethics in Human Research (Project 09-127).

2.1.1. Reference populations

De-identified STR genotype data for three Australian populations were obtained for use as a quality control measure [17,18] for comparison to the Indonesian populations analysed in the present study. These Australian populations are; Australian Caucasians from the Australian Capital Territory where individuals self-declared their ethnicity as Caucasian (Australian Caucasian; $N = 223$); indigenous (Aboriginal) Australians from the Northern Territory where individuals were known to be indigenous and residing in remote tribal communities (Australian Aboriginal; $N = 560$); and indigenous (Aboriginal) Australians from the Northern Territory where individuals self-declared their ethnicity as indigenous (Aboriginal) Australian (Declared Australian Aboriginal; $N = 850$). The three Australian populations have been used in the genetic distance analyses as a quality control measure.

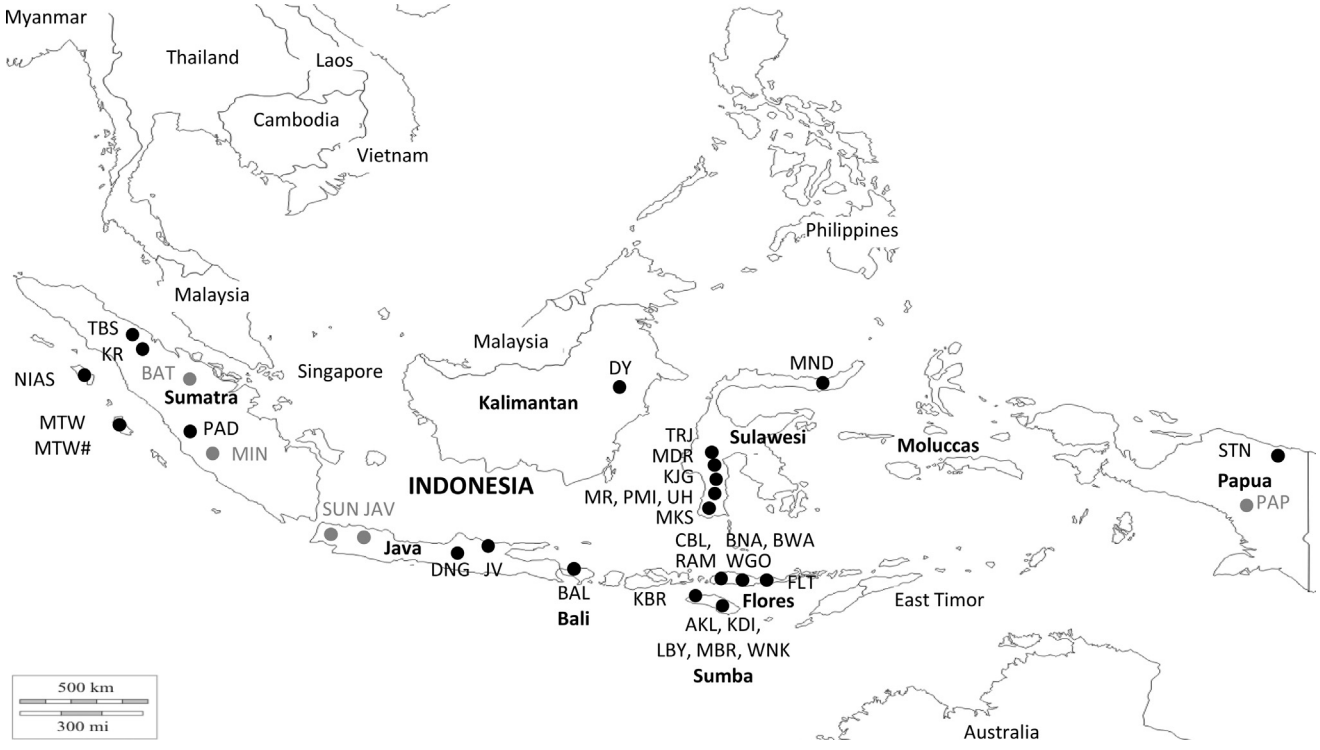

Fig. 1. Collection sites in the Indonesian archipelago for samples used in this study. The Eijkman Institute for Molecular Biology collection sites are indicated in black and the Indonesian National Police sites are indicated in grey.