Quaternary International xxx (2015) 1–6

Contents lists available at ScienceDirect

Quaternary International

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

Western Eurasian genetic influences in the Indonesian archipelago

Pradiptajati Kusuma [a, b], Murray P. Cox [c], Nicolas Brucato [a, d], Herawati Sudoyo [b], Thierry Letellier [a], François-Xavier Ricaut [a, *]

[a] *Laboratoire d'Anthropologie Moléculaire et Imagérie de Synthèse UMR-5288, Université de Toulouse, Toulouse, France*
[b] *Genome Diversity and Diseases Laboratory, Eijkman Institute for Molecular Biology, Jakarta, Indonesia*
[c] *Statistics and Bioinformatics Group, Institute of Fundamental Sciences, Massey University, Palmerston North, New Zealand*
[d] *Center for Linguistics, University of Leiden, Leiden, The Netherlands*

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Western Eurasia
India
Arabia
Indonesia
Mitochondrial DNA
Y chromosome

ABSTRACT

Western Eurasia, notably the Near East and South Asia (Indian sub-continent), has interacted with Indonesia through Indian Ocean trade (the Maritime Silk Route) for more than 2000 years. The Indianization, and later Islamization, of Indonesia was enacted largely through trading activities, but also spread with help from the many Indianized and Islamic kingdoms that reigned over parts of the Indonesian archipelago during this time. Western Eurasian interaction left behind not only imported trade goods and cultural features, but also genetic traces. To locate the primary areas of Western Eurasian genetic influence in Indonesia, we have assembled published uniparental genetic data from ~2900 Indonesian individuals. Frequency distributions show that Western Eurasian paternal lineages are found more commonly than Western Eurasian maternal lineages. Furthermore, the origins of these paternal lineages are more diverse than the corresponding maternal lineages, predominantly tracing back to South West and South Asia, and the Indian sub-continent, respectively. Indianized kingdoms in the Indonesian archipelago likely played a major role in dispersing Western Eurasian lineages, as these kingdoms overlap geographically with the current distribution of individuals carrying Western Eurasian genetic markers. Our data highlight the important role of these Western Eurasian migrants in contributing to the complexity of genetic diversity across the Indonesian archipelago today.

1. Introduction

Western Eurasia (WE), notably the Near East and South Asia, had established very close contact with Island South East Asia, particularly Indonesia, through cultural and trading networks in the Maritime Silk Route as early as the 1st century B.C.E. (Ardika and Bellwood, 1991; Ardika et al., 1997; Calo, 2014; Lawler, 2014). Contacts between the Indian sub-continent and Southeast Asia occurred even earlier (from the late 2nd millennium B.C.E.) if we consider the evidence from transfers of plants (Fuller, 2006; Asouti and Fuller, 2008), and trade networks were well established by the 1st millennium B.C.E. (Bellina and Glover, 2004). Indianization and Islamization in Indonesia led to the development of Hindu, and later Islamic, kingdoms between the 5th and 15th centuries, which reinforced interactions and influence from WE to all regions of western and central Indonesia (Gonda, 1975; Kanchan, 1990; Beaujard, 2012). These interactions brought not only new ideas and trade goods into the Indonesian archipelago, but were also accompanied to some degree by gene flow, which contributed to a number of modern Indonesian populations. Considering the historical events of the last two millennia, it has been variously postulated that WE genetic influx into Indonesia could have originated from populations in India (via the main Indian kingdoms – Pallava and latter Chola in Southeast India; Gurjara and Rashtrakuta in north and east India; and later Indian sultanates) (Mabbett, 1977; Lansing, 1983; Lukas, 2003; Beaujard, 2012), the Near East (via Persian influence, the Sassanid empire and later the Arab Muslim empire) (Jacobsen, 2009; Beaujard, 2012) or Europe following Portuguese and Dutch colonialism (Taylor, 2009). The genetic evidence so far suggests that the primary contribution came from the Indian subcontinent (Karafet et al., 2005, 2010; The HUGO Pan-Asian SNP Consortium, 2009), but there is still debate regarding

---

* Corresponding author.
E-mail addresses: pradiptajati.kusuma@univ-tlse3.fr (P. Kusuma), m.p.cox@massey.ac.nz (M.P. Cox), n.g.a.brucato@hum.leidenuniv.nl (N. Brucato), herawati@eijkman.go.id (H. Sudoyo), thierry.letellier@u-bordeaux2.fr (T. Letellier), francois-xavier.ricaut@univ-tlse3.fr (F.-X. Ricaut).

http://dx.doi.org/10.1016/j.quaint.2015.06.048
1040-6182/

Please cite this article in press as: Kusuma, P., et al., Western Eurasian genetic influences in the Indonesian archipelago, Quaternary International (2015), http://dx.doi.org/10.1016/j.quaint.2015.06.048

the importance of Indian gene flow, and its geographic diffusion into and impact on the Indonesian archipelago. Some scholars propose either colonization by Indian exiles (Majumdar, 1963), large-scale migration establishing Indian colonies in South East Asia (Mabbett, 1977), or more punctuated contact by Indian traders (Lansing, 1983; Sandhu and Mani, 1993), but the available data are insufficient to suggest a reliable chronological framework for these gene flow scenarios at the scale of the whole Indonesian archipelago. Historical evidence mentions the "… constant flux of people from Gujarat and neighboring regions [in the Indian sub-continent] …" to Java early in the 9th century, related to the construction of the Plaosan temple (Beaujard, 2012, pp. 94). Indeed, genetic traces of paternal Indian ancestry have been detected at low frequency in Indonesia (<10%) (Karafet et al., 2005, 2010) across several islands (Sumatra, Java, Borneo and Bali), but in comparison, no Indian maternal lineages have been detected (Tumonggor et al., 2013). The influx of Indian ancestry has also been observed from genome-wide analysis using a limited number of nuclear SNPs (The HUGO Pan-Asian SNP Consortium, 2009) in just a few populations from western Indonesia (Sumatra). However, Indian markers in autosomal data were not detected in large-scale genotyping of eleven populations across Island Southeast Asia (Pugach et al., 2013).

Our study is based on the largest dataset of maternal (mtDNA) and paternal (Y chromosome) lineages assembled for the Indonesian archipelago to date. However we need to bear in mind some limitations of this dataset; namely (1) the resolution of the data (at the haplogroup level), which only allows the tracing of haplogroup origins at a regional scale (the Indian sub-continent, Middle East and Europe), and (2) the source of gene flow, which is inferred from the hypothesized origin of WE haplogroups (which, in turn, is inferred from where these are found at high frequencies in populations today). In addition, there are chronological limitations caused by (3) the inability of dating the time of WE gene flow to Indonesia from the data used in this study, that lead us (4) to rely on the archaeological and historical chronology (e.g., gene flow starting at least from the late 1st millennium B.C.E. with a likely intense phase during the period of the Hindu kingdoms in Indonesia (7th–16th century AD) (Beaujard, 2012; Calo, 2014). These assumptions and approaches are reasonable and standard in the field. Our primary aim is to use a phylogeographic approach (assessing the frequency of WE haplogroups in Indonesia) to (1) map Western Eurasian genetic influences in Indonesia, and to a lesser extent to (2) identify the main regional sources of these influences (the Indian sub-continent, Middle East and Europe). Considering early cultural and trade networks with the Indian sub-continent, we will test the hypothesis that India is the main source of WE influence in Indonesia, and that this WE gene flow is probably sex biased. We propose that genetic contributions from Western Eurasia and the Indian sub-continent are mostly restricted to the areas where Indianized kingdoms were most powerful. We also show that genetic contact with both Indian men and women was involved.

2. Materials and methods

2.1. DNA samples and data sources

Indonesian samples used in this study have been described previously (Karafet et al., 2010; Gunnarsdóttir et al., 2011; Tumonggor et al., 2013; Kusuma et al., 2015). In brief, the mtDNA dataset includes 2841 individuals, and the Y chromosome dataset includes 2095 individuals, from 7 major Indonesian islands, representing 22 and 17 different ethnic groups, respectively (Table 1). These two datasets represent the largest Indonesian genetic assemblage studied to date.

Table 1
List of populations used in this study.

Group	Islands	Population code	Population	N mtDNA	N Y chromosome	Ref.
Western Indonesia	Sumatra	1	Gayo	62	*NA*[a]	Eijkman Institute archived samples
		2	Batak Toba	42	37	Karafet et al., 2010; Tumonggor et al., 2013
		3	Besemah	36	38	Gunnarsdóttir et al., 2011
		4	Semende	36	37	Gunnarsdóttir et al., 2011
		5	Nias	59	60	Karafet et al., 2010; Tumonggor et al., 2013
		6	Mentawai	128	74	Karafet et al., 2010; Tumonggor et al., 2013
	Borneo	7	Lebbo'	19	15	Kusuma et al., 2015
		8	Ma'anyan	159	90	Kusuma et al., 2015
		9	SK Dayak[b]	64	*NA*	Eijkman Institute archived samples
		10	EK Dayak[b]	*NA*	85	Karafet et al., 2010
	Java	11	Java	51	61	Karafet et al., 2010; Tumonggor et al., 2013
	Bali	12	Bali	487	634	Karafet et al., 2010; Tumonggor et al., 2013
Eastern Indonesia	Sulawesi	13	Mandar	54	54	Karafet et al., 2010; Tumonggor et al., 2013
		14	Kajang	46	*NA*	Tumonggor et al., 2013
		15	Toraja	50	*NA*	Tumonggor et al., 2013
		16	Bugis	50	*NA*	Tumonggor et al., 2013
		17	Bajo	27	27	Kusuma et al., 2015
	Lesser Sunda	18	Sumba	634	349	Karafet et al., 2010; Tumonggor et al., 2013
		19	Flores	469	388	Karafet et al., 2010; Tumonggor et al., 2013
		20	Lembata	92	89	Karafet et al., 2010; Tumonggor et al., 2013
		21	Pantar	29	*NA*	Tumonggor et al., 2013
		22	Alor	23	27	Karafet et al., 2010; Tumonggor et al., 2013
	Maluku	23	North Maluku	224	*NA*	Eijkman Institute archived samples
		24	Maluku (Hiri and Ternate)	*NA*	30	Karafet et al., 2010

[a] NA = Not Available for analysis.
[b] Samples of SK Dayak were collected from various Dayak ethnics in South Kalimantan, while the EK Dayak were collected from various Dayak ethnics in East Kalimantan.

2.2. Haplogroup assignment

Mitochondrial DNA hypervariable region I sequences were compiled from published sources (Table 1) and aligned against the revised Cambridge Reference Sequence (rCRS) (Andrews et al.,

Please cite this article in press as: Kusuma, P., et al., Western Eurasian genetic influences in the Indonesian archipelago, Quaternary International (2015), http://dx.doi.org/10.1016/j.quaint.2015.06.048