



The Missing landscape of human genomic diversity in the Arabian Peninsula

Njlaa Bakhsh, Latifa Jackson, Christopher Cross, Fatimah Jackson



Background/Introduction

The Arabian Peninsula (AP) is the first site of human migration and habitation outside of Africa. As a major crossroad for human populations, the AP provides an opportunity to better understand early to modern changes in human demographic patterns through selections, admixture, gene flow, and migration. Dramatic climatic fluctuations have been recorded in the AP that contributed to contractions and expansions in water availability. These climatological perturbations are thought to have shaped genomic variations in this population. Recent reports indicate that a number of Arab nation-states have committed significant resources to genetically typing the national population, with the overall goal of determining the degree of genomic diversity in the AP. We sought to characterize currently typed genomic variation in Arabian populations to support the rationale for our proposed analyses of Saudi Arabian genomic diversity. Interestingly, in contrast to published claims, a comprehensive search of peer-reviewed reports on genomic analysis (N=20 papers) revealed no genomic data from four national genomic projects (Qatar, Saudi Arabia, Kuwait, and The United Arab Emirates). Our analysis demonstrates that while much fanfare and presumably resources have been devoted to defining the genomic landscape of the Arabian peoples, little actual data is available to either substantiate or support such an investment.

Hypothesis/Objective

The **main objective** of this study was to search available resources for data on the genomic diversity of the AP population. Our **hypothesis** is that genomic sampling of the AP population will reveal patterns of genomic diversity that may have resulted from ecological variability, and that information from this analysis can be used to trace regional origins, migration patterns, gene flow, and differential geospatial selection pressure

Methodology

We developed a search strategy and methods based on: (i) specific criteria for selection of peer-reviewed genomic reports, (ii) extraction of relevant studies, (iii) assessment of study, and (iv) statistical analysis of available data.

Data Search: We searched international electronic databases (PubMed, GeneWatch-UK, The 1000 Genome project, the International HapMap Project of the NIH National Human Genome Research Institute, and the National Center for Biotechnology Information, NCBI).

Search Criteria: We searched the databases using a combination of key words: "genome", "diversity", "DNA database", "Arabian Peninsula", "Saudi Arabia", "United Arab Emirates", "Kuwait", "Qatar", "Yemen". We limited our search to English language articles published in Journals indexed in the National Library of Medicine database. We only analyzed peer-reviewed articles with complete abstracts, background, and cross-referenced bibliography (see examples of articles in Table 1). Articles were considered eligible if they included population-level genomic data from AP countries. Duplicate publications, reviews, and abstracts from conference reports were excluded.

Results

AP countries have proposed to create a genomic DNA database, but there is no evidence of such a database that can be accessed for analysis..



Figure 1: The 1000 Genomes Project shows no data available on AP.

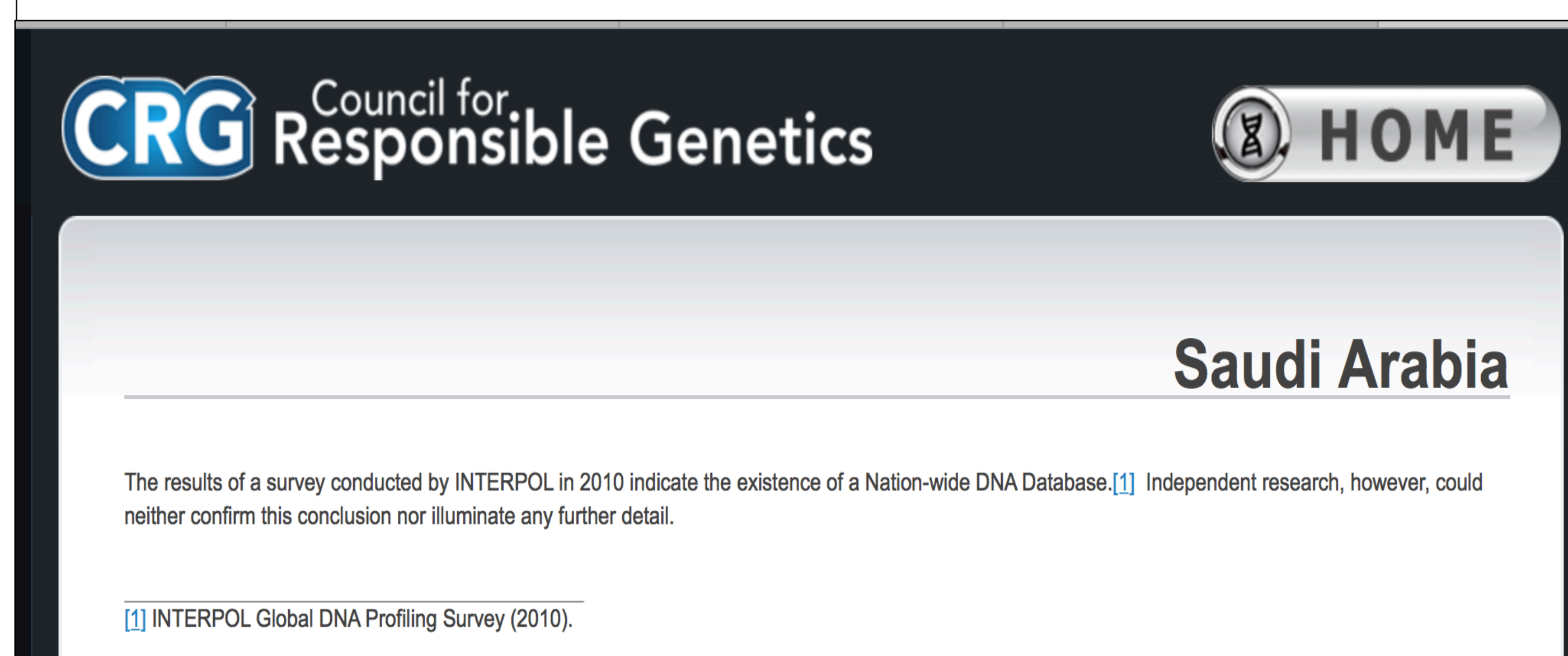


Figure 2: GeneWatch revealed that Saudi Arabia is planning to generate a DNA database program.

Table 1: Examples of peer-reviewed reports analyzed in this study

First Author	Year	Title/Journal
Alsmadi	2012	Genetic Substructure of Kuwaiti Population Reveals Migration History. <i>BioMed Central</i> .
Thareja	2015	Sequence and Analysis of a Whole Genome from Kuwaiti Population Subgroup of Persian Ancestry. <i>BMC Genomics</i> . BioMed Central. 2015.
Abu-Amro	2007	Eurasian and African Mitochondrial DNA Influences in the Saudi Arabian Population. <i>BMC Evolutionary Biology</i> 2007, 7:32.
Abu-Amro	2011	Mitochondrial DNA Lineages of African Origin Confer Susceptibility to Primary Open-angle Glaucoma in Saudi Patients. <i>BMC Evolutionary Biology</i>
Petraglia	2015	The Middle Paleolithic of Arabia: Implications for Modern Human Origins, Behavior and Dispersals, <i>Antiquity</i>

Conclusions/Discussion/Future Directions

The AP links Europe, Asia, and Africa, and has served as a hub for the early dispersal of people between these regions. Climatic fluctuations across the AP have contributed to significant migrations of people across the AP. However, the effect of these changes on genomic diversity in the region have not been comprehensively examined. In order to understand the origin and evolution of human populations in the AP, future studies should aim to:

1. Comprehensively determine the genomic landscape of the AP in order to understand how ecological variability, and migration patterns may have influenced the the genomic diversity of this important region.
2. Develop an interdisciplinary strategy for collecting and analyzing genomic data generated in accordance with the demographic structure of the AP population groups.
3. Compare and contrast the genomic landscape of the AP relative to the known genomic diversity of continental Africa and Eurasia; this information will then be used to trace regional origins, migration patterns, and gene flow, and to predict the genomic makeup of the foundational AP subsets with links to Africa and Eurasia, which will further guide both applied and theoretical studies of the genomic complexity of the AP.
4. Utilize recent advances in techniques for sequencing of nuclear, mitochondrial, Y- chromosome, and autosomal DNA that have made it possible to study human genetic variation and the genetics of human populations. These approaches have helped to determine founder populations of modern peoples as well as the human migrations that shaped these variations.

References

1. Abu-Amro, Khaled K., Larruga, JM., Cabrera, VM., Gonzales, AM. Mitochondrial DNA structure in the Arabian Peninsula. *BMC Evolutionary Biology* 2008. 8:45.
2. Abu-Amro, Khaled K, Ana M González, Jose M Larruga, Thomas M Bosley, and Vicente M Cabrera. Eurasian and African mitochondrial DNA influences in the Saudi Arabian population. *BMC Evolutionary Biology* ,2007
3. Thareja, Gaurav, Sumi Elsa John, Prashantha Hebbar, Kazem Behbehani, Thangavel Alphonse Thanaraj, and Osama Alsmadi. "Sequence and Analysis of a Whole Genome from Kuwaiti Population Subgroup of Persian Ancestry." *BMC Genomics*. BioMed Central, 18 Feb. 2015.