HIV-1 molecular epidemiology evidence and transmission patterns in the Middle East and North Africa

Ghina Mumtaz,1 Nahla Hilmi,2 Francisca Ayodeji Akala,2 Iris Semini,3 Gabriele Riedner,4 David Wilson,5 Laith J Abu-Raddad1,6,7

ABSTRACT

The distribution of HIV-1 subtypes in a population tracks the spread and evolution of the epidemic. This study is a systematic review of all available evidence on HIV-1 molecular epidemiology and subtype distribution in the Middle East and North Africa. Sources of data included Medline and various institutional documents and databases. In several countries, a diverse distribution of HIV-1 subtypes was observed principally reflecting travel-related exogenous exposures. A trend for a dominant HIV-1 subtype was observed in a few other settings and was often linked to HIV transmission within specific high-risk groups such as subtype A and CRF35_AD among injecting drug users and subtype C among commercial sex networks. Multiple exogenous introductions of HIV-1 variants seemed common to all countries, as observed from the high diversity in subtypes, or the high genetic divergence among any specific subtype even if predominant. In several countries though, epidemic-type clustering of specific subtypes suggests established or nascent HIV epidemics among classic core risk groups for HIV infection. HIV prevention efforts in MENA must be prioritized for these high-risk groups.

HIV-1 is characterised by a high genetic variability and includes three groups, with the main group responsible for the HIV pandemic (group M) being further divided into several subtypes (A–K) and a number of mosaic strains known as circulating recombinant forms (CRF).1 With the exception of sub-Saharan Africa, where the largest diversity of subtypes and CRF has been observed, HIV-1 subtypes show a specific geographical distribution on the global map.2,3 Subtype A is the predominant variant in central and eastern Africa and in countries of the former Soviet Union. Subtype B is predominant in western and central Europe, the Americas, Australia, and is commonly found in several countries of Southeast Asia.2 Subtype C is the main genetic form in southern Africa and India, the two countries where more than 80% of the global HIV cases occur. Finally, CRF are the most predominant forms in Southeast Asia (CRF01 AE) and in west and central Africa (CRF02_AG).2–7

Molecular epidemiology investigation can be a powerful tool in developing sensitive diagnostic tools, in the management of individual infections, and most importantly in tracking transmission patterns and the spread and evolution of the epidemic.4 Molecular epidemiology research in the Middle East and North Africa (MENA) region is at an early stage of development and has yet to contribute to informed HIV prevention policy and programming. A recent review on the epidemiology of HIV in MENA indicates that in this region where HIV spread in the general population is very limited, there are emerging epidemics in high-risk populations including injecting drug users (IDU), men who have sex with men (MSM), and to a lesser extent female sex workers (FSW) in a few countries.5 Exogenous HIV exposures among nationals linked to travel abroad also appear to be the dominant observed HIV transmission pattern in a few MENA countries.5 However, there remain many gaps in the understanding of HIV transmission pathways in MENA.

The objective of this work was to review for the first time all evidence on HIV virus diversity and subtype distribution in MENA countries to complement our knowledge of HIV epidemiology and to understand transmission patterns.

METHODS

This work was part of a comprehensive systematic review of all available data about HIV in MENA, which was complemented by another detailed search focused specifically on published HIV molecular epidemiology literature in MENA. The countries included in our definition of MENA are: Afghanistan, Algeria, Bahrain, Djibouti, Egypt, Iran, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Pakistan, Palestine, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates and Yemen. Considering geographical proximity and similarity in the sociocultural context, data were occasionally included on Mauritania, Israel and Turkey.

The main source of data for this investigation was the MENA HIV/AIDS Epidemiology Synthesis Project whose mandate was to collect and synthesise all available data on HIV, sexually transmitted infections (STI) and sexual behaviour in MENA.5 Details of this project and of the search strategy used in this comprehensive systematic review were described in a previous publication.5 In brief, we undertook a literature search of Medline (PubMed) using a strategy with both free text and MeSH headings and with no time limitation. The main search criteria in relevance to the current work were (‘HIV Seropositivity’ OR ‘HIV’ OR ‘HIV Infections’) AND (‘Middle East’ OR ‘Islam’ OR ‘Arabs’ OR ‘Arab World’ OR ‘Africa, Northern’ OR ‘Mauritania’ OR ‘Sudan’ OR ‘Somalia’ OR...
Epidemiology

‘Djibouti’ OR ‘Pakistan’). We also searched for relevant data in country-level reports and databases, governmental and non-governmental organisations studies and publications, international organisations’ reports and databases, as well as other institutional reports related to HIV and STI in MENA.5

The above generic search was complemented by a specific search of the Medline database on all HIV-1 molecular data in MENA. The MeSH search criteria used were ((‘HIV-1-classification’ OR ‘HIV-1-genetics’) AND (‘Middle East’ OR ‘Islam’ OR ‘Iran’ OR ‘Arabs’ OR ‘Arab World’ OR ‘Africa, Northern’ OR ‘Mauritania’ OR ‘Sudan’ OR ‘Somalia’ OR ‘Djibouti’ OR ‘Pakistan’)). All relevant studies identified in this search were already covered by the wide-umbrella search of the Synthesis Project above. However, this search was undertaken and respective citations reviewed for further reassurance that no relevant data were missed.

Finally, we searched the Los Alamos HIV sequence database, which gathers global HIV sequence data stratified by geographical location, but without any information on data source and study methodology.7 Data from all MENA countries were extracted from this online database and compared with results of published studies.

RESULTS

Status of the evidence

As part of the comprehensive Medline search that was undertaken for the Synthesis Project, a total of 1092 citations on HIV in MENA were retrieved as of 29 May 2010, out of which 49 citations were specific to the molecular epidemiology search. After excluding non-relevant studies, there remained 24 studies that examined the nature of HIV-1 subtypes and suggested transmission pathways in the population (table 1). HIV-1 subtype data were also extracted from the Los Alamos HIV database for all MENA countries and suggested trends similar to those observed in published studies (table 2).

All the reviewed studies used standard and comparable methodologies1 31 in their molecular analyses, consisting of the amplification of specific genomic regions (mostly env, gag and/or pol) using PCR, followed by sequencing of the PCR products and phylogenetic tree analysis. It appears unlikely that there is bias in comparing findings between studies and over time due to differences in the methodology of molecular analyses.

Settings with variability of subtypes

A number of countries in MENA display a very diverse distribution of HIV-1 subtypes. This is the case of Lebanon where a complex HIV-1 subtype distribution pattern and high levels of intrasubtype diversity were found representing the travel history of its nationals, particularly to Europe and the Americas where subtype B (44%) is acquired, and to Africa and the United Arab Emirates where most of the non-B subtypes, especially subtype A (40%), are acquired.15 An analogous trend of high strain diversity including many CRF was observed in Saudi Arabia.24–26 However, the identification of a subcluster with a new CRF in this country (CRF45_02C) could indicate a local transmission network.24 HIV-1 subtypes in Yemen also displayed a complex distribution, which appeared to be a mix of

<table>
<thead>
<tr>
<th>Country</th>
<th>Study</th>
<th>Sample size</th>
<th>Year*</th>
<th>Genes</th>
<th>Subtype (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afghanistan</td>
<td>Sanders-Buek8</td>
<td>4 IDU</td>
<td>2005–6</td>
<td>Pol, Env</td>
<td>CRF35 AD (100)</td>
</tr>
<tr>
<td>Algeria</td>
<td>Bouzeghoub9</td>
<td>134 Isolates</td>
<td>NA†</td>
<td>Rt, Prot, Env</td>
<td>B (56), CRF02 AG (12.7), CRF06_cpx (4), Int-CRF recombinants (9.7)</td>
</tr>
<tr>
<td>Djibouti</td>
<td>Maslin10</td>
<td>34 Isolates</td>
<td>2002</td>
<td>Pol, Env</td>
<td>C (73), CRF02 AG (18), D (6), A (3)</td>
</tr>
<tr>
<td></td>
<td>Lasky11</td>
<td>33 French military personnel in Djibouti</td>
<td>NA†</td>
<td>Env</td>
<td>C (48), B (33), A (15), E (3)</td>
</tr>
<tr>
<td>Iran</td>
<td>Sarrami-Faroo shani12</td>
<td>46 IDU and 15 haemophiliacs in Teheran</td>
<td>2003 Gag, Env</td>
<td>IDU: A (100)</td>
<td>Haemophiliacs: B (100)</td>
</tr>
<tr>
<td></td>
<td>Nader13</td>
<td>12 IDU in Mashhad</td>
<td>2005</td>
<td>Gag, Env</td>
<td>A (100)</td>
</tr>
<tr>
<td></td>
<td>Soheili14</td>
<td>13 Isolates, mostly IDU</td>
<td>NA†</td>
<td>RT, Prot, Env</td>
<td>CRF35 AD (100)</td>
</tr>
<tr>
<td></td>
<td>Mousavi15</td>
<td>39 Isolates, mostly IDU</td>
<td>2005–6</td>
<td>Pol</td>
<td>CRF35 AD (100)</td>
</tr>
<tr>
<td></td>
<td>Hamkar16</td>
<td>42 Isolates, 36% IDU</td>
<td>2009 Pol</td>
<td>A/D recombinants (48), B (43), A (5), CRF01_AE (5)</td>
<td></td>
</tr>
<tr>
<td>Israel/Palestine</td>
<td>Gehring17</td>
<td>92 Isolates, 45% MSM</td>
<td>NA†</td>
<td>Env</td>
<td>B (89), A (11), D (1)</td>
</tr>
<tr>
<td>Lebanon</td>
<td>Pieniazek18</td>
<td>25 Isolates</td>
<td>1996</td>
<td>Prot, Gag, Env</td>
<td>A (44), B (40), C (4), D (4), G (4),</td>
</tr>
<tr>
<td>Libya</td>
<td>de Oliveira19</td>
<td>44 Children</td>
<td>1998</td>
<td>Gag</td>
<td>CRF02 AG (100)</td>
</tr>
<tr>
<td>Morocco</td>
<td>Elhart20</td>
<td>200 Isolates</td>
<td>NA†</td>
<td>Gag</td>
<td>CRF35 AD (100)</td>
</tr>
<tr>
<td></td>
<td>Abid21</td>
<td>14 Isolates</td>
<td>NA†</td>
<td>Env, Tat</td>
<td>B (78.6%), A (21.4%)</td>
</tr>
<tr>
<td></td>
<td>Bakhouch22</td>
<td>98 Isolates</td>
<td>2004–7</td>
<td>Rt, Prot</td>
<td>B (74.6%), CRF02 AG (15.5), CRF01_AE (4.2), G (2.8), C (1.4), F2 (1.4%)</td>
</tr>
<tr>
<td>Pakistan</td>
<td>Khan23</td>
<td>34 IDU in Karachi</td>
<td>2006</td>
<td>Gag, Env</td>
<td>A (100)</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>Yamaguchi24</td>
<td>56 Saudi nationals in Riyadh</td>
<td>2003–4</td>
<td>Gag, Pol, Env</td>
<td>C (39.3), B (17.8), Intersubtype recombinants (16.1), CRF43_02G (10.7), CRF25_cpx (7.1), D (3.6), A (1.8), CRF02_AG (1.8)</td>
</tr>
<tr>
<td></td>
<td>Badreddine26</td>
<td>30 Blood donors</td>
<td>1998–9</td>
<td>Gag, Env</td>
<td>C (58%), B (17%), A (8%), D (8%), G (8%)</td>
</tr>
<tr>
<td>Sudan</td>
<td>Alzahrani26</td>
<td>12 Isolates in Dammam</td>
<td>2004–7</td>
<td>Gag, Env</td>
<td>D (50), C (30), recombinants (10), A (6.7), B (3.3)</td>
</tr>
<tr>
<td></td>
<td>Hierholzer27</td>
<td>30 Blood donors</td>
<td>1998–9</td>
<td>Gag, Env</td>
<td>D (50), C (30), recombinants (10), A (6.7), B (3.3)</td>
</tr>
<tr>
<td>Tunisia</td>
<td>Ben Halima28</td>
<td>25 Isolates</td>
<td>1999</td>
<td>Gag</td>
<td>B (95), CRF02 AG (5)</td>
</tr>
<tr>
<td>Turkey</td>
<td>Yilmaz29</td>
<td>27 Isolates</td>
<td>NA†</td>
<td>Gag</td>
<td>B (70.4), A (14.8), F1 (7.4), C (3.7), D (3.7)</td>
</tr>
<tr>
<td>Yemen</td>
<td>Saad20</td>
<td>19 Isolates</td>
<td>2000–2</td>
<td>Gag</td>
<td>B (47.3), C (31.6), D (10.5), A (5.3), URF (5.3)</td>
</tr>
</tbody>
</table>

*Year of data collection.
†Not available in original document.
CRF, circulating recombinant form; IDU, injecting drug user; MENA, Middle East and North Africa; MSM, men who have sex with men.
The high diversity in subtypes and/or the high genetic divergence among any specific subtype suggest that HIV infections in these countries are driven by travel-related exogenous exposures among nationals or HIV acquisition by their sexual partners upon their return. Data from case notification reports confirm this postulated trend. For example in Lebanon, 45.4% of notified cases up to 2004 were linked to travel abroad, as were half of reported AIDS cases in Yemen. In these settings, local spread is usually limited to immediate sexual partners and does not lead to an epidemic.
The high population mobility in MENA including migration, tourism, studying and working abroad, and military deployment among others explain in large part the observed pattern of exogenous exposures and subsequent subtype diversity, as well as explaining the observed Africanisation of subtypes in some settings. Southern Algeria and Morocco are both important migration hubs for migrants from sub-Saharan African countries towards Europe and this is reflected in the high diversity of non-B subtypes and CRF in these regions. In Turkey, non-B subtype infections are also thought to be mainly caused by immigrants or residents from Africa, the Balkans and the Middle East. The high prevalence of subtype C in Saudi Arabia is expected in view of the country’s ties with neighbouring Horn of Africa countries. The very high diversity of subtypes observed in Lebanon reflects the frequent population movements and return of expatriates such as from West Africa. The high level of emigration to western Europe explains the predominance of subtype B in Tunisia, which is thought to have been initially introduced mainly by young male IDU or individuals with haemophilia infected in Europe.

Patterns of local epidemic-type clustering

On the other hand, the dominance of specific subtypes with limited sequence variability in Afghanistan, Djibouti, Iran, Libya, Pakistan and Somalia in addition possibly to one subcluster in Saudi Arabia, suggests that HIV acquisitions in some countries are occurring in transmission chains that have been propagating locally for at least a number of years. This suggests active high-risk networks where HIV is being transmitted via typical founder effect, which translates into local epidemic-type clustering in one or multiple risk groups (figure 1B). This trend was documented in Russia, for example, where injecting drug use represented the main mode of HIV transmission for the longest time and where subtype A was the predominant clade driving the epidemic among IDU.

This conjecture is supported in Iran and Pakistan as examples, by robust evidence for concentrated epidemics among IDU. By the end of 2007, IDU accounted for at least 67% of HIV notified cases in Iran. In Pakistan, the contribution of IDU to the total notified cases has been rising, and IDU accounted for the largest proportion of HIV infections in 2008. The similarities between HIV-1 subtypes in Afghanistan, Iran and apparently Pakistan suggest that strains have been circulating between these countries, probably through the return of Afghani refugees, and may indicate that social networks of drug producers, refiners, distributors and IDU transcend national borders.

The local epidemic-type clustering is also supported by epidemiological evidence from Djibouti and from some parts of Somalia, where there is evidence of concentrated HIV epidemics in commercial heterosexual sex networks, with rates of HIV up to 70% being reported among street-based FSW in Djibouti. Clustering within subtype C in HIV-1 infections in Djibouti and Somalia is in accordance with previous studies in the Horn of Africa, which have identified subtype C to be the dominant clade, accounting for up to 98.5% of infections among FSW in Ethiopia. Similarly, in Sudan the epidemic appears to be dominated by East African viruses also circulating in neighbouring countries. Libya is yet another example where clustering of subtypes was observed, but in the context of a long-standing nosocomial transmission scenario.

Transmission links between epidemics

One of the insightful applications of HIV molecular research is that it can indicate whether there is a link between two epidemics occurring simultaneously in two different risk groups in one country. This was reported in many regions of the world such as in Thailand, where in the first decade of the epidemic, HIV-1 subtypes were segregated by risk group, with IDU being infected with subtype B and patients infected through the heterosexual route being infected with CRF01_AE.

Similarly in Iran, molecular epidemiology investigations indicated that the HIV epidemic among IDU is independent of the epidemic among individuals with haemophilia as seen by the different subtype clustering in each transmission network: IDU were found to be exclusively infected with subtype A, which has been possibly introduced to Iranian IDU by a pilgrim visiting Mashhad, Iran’s holiest city, whereas individuals with haemophilia were exclusively infected by subtype B. In Israel/Palestine, whereas subtype B was the predominant clade, subtype C was exclusively observed among the Ethiopian Israeli population, indicating two distinct HIV transmission pathways in the population that possibly reflect the geographical site of virus acquisition.

Recent phylogenetic analyses have also linked epidemics in MENA such as the emerging epidemic among MSM in Pakistan.
and that which emerged among IDU earlier in this decade in this country (S. Ali, personal communication, 2010).

**Recommendations for future research and interventions**

With the weak HIV surveillance systems in this region, it is useful to expand HIV molecular research, especially as some of the available studies have been conducted a decade ago. Overall in MENA, only a very small number of HIV cases have ever been subtyped. Despite the high cost of molecular analyses, identifying the subtypes and studying the phylogenetic clustering of at least small subsamples of notified HIV cases and HIV-seropositive cases in surveyed populations is pertinent. It can pinpoint the most affected risk groups and the prevalent modes of transmission in the population, and help establish some reliable information on the subtype distributions. However, the priority in MENA countries that are still developing their HIV surveillance infrastructure would be to invest in integrated biobehavioural surveillance using state of the art sampling methodologies for hard-to-reach populations. This kind of surveillance is essential in order to facilitate the prioritisation of HIV response to where needs are and to avert the possibility of HIV efforts being dissipated on non-consequential programmes.

The molecular evidence gathered in this review, supported by other epidemiological evidence (G Muntaz, 2010, unpublished data), indicates the absence of a sustainable general population HIV epidemic in MENA, with the exception possibly of southern Sudan. As such, any current HIV prevention efforts focused on the general population are misplaced. Whereas it is challenging to identify, prevent or even control travel-related exogenous exposures, the number of infections arising from these random introductions remains relatively small. Instead, surveillance and prevention efforts should be prioritised for high-risk groups including IDU, MSM, and FSW, in whom low intensity and concentrated epidemics are emerging or are already established. This would prevent further HIV transmission or the emergence of localised and hidden HIV epidemics among these population groups.

**Funding**

The MENA HIV/AIDS Epidemiology Synthesis Project was funded through a joint partnership of the World Bank, the MENA Regional Support Team of the Joint United Nations Programme on HIV/AIDS (UNAIDS) and the Eastern Mediterranean Regional Office (EMRO) of the WHO. LJA and GM are also grateful for the Qatar National Research Fund for supporting this work (NPRP 08-068-3-024).

**Competing interests**

None.

**Contributors**

LJA conceived and led the design and conduct of this study; co-led the literature reviews and collection, analysis and interpretation of the data, and contributed to the drafting of the manuscript. GM led the drafting of the manuscript and co-led the literature reviews and collection, analysis and interpretation of the data. NH contributed to the literature reviews and assisted in the general conduct of the research. FAA, IS, GR and DW contributed to the interpretation of the results, contributed to the drafting of the manuscript, and provided data at the regional and country level. All authors were involved in the finalisation of the manuscript and approved the final version.

**Provenance and peer review**

Not commissioned; externally peer reviewed.

**REFERENCES**