Hepatitis C virus (HCV) infection is a major health problem in Egypt, where the sero-prevalence is 10–20-fold higher than that in the United States. To characterize the HCV genotype distribution and concordance of genotype assessments on the basis of multiple genomic regions, specimens were obtained from blood donors in 15 geographically diverse governorates throughout Egypt. The 5′ noncoding, core/E1, and NS5B regions were amplified by reverse transcription–polymerase chain reaction and analyzed by both restriction fragment length polymorphism (RFLP) and phylogenetic tree construction. For the 5′ noncoding region, 122 (64%) of 190 specimens were amplified and analyzed by RFLP: 111 (91%) were genotype 4, 1 (1%) was genotype 1a, 1 (1%) was genotype 1b, and 9 (7%) could not be typed. Phylogenetic analyses of the core/E1 and NS5B regions confirmed the genotype 4 preponderance and revealed evidence of 3 new subtypes. Analysis of genetic distance between isolates was consistent with the introduction of multiple virus strains 75–140 years ago, and no clustering was detected within geographic regions, suggesting widespread dispersion at some time since then.

An estimated 170 million people worldwide have hepatitis C virus (HCV) infection [1]. HCV infection frequently persists and may cause cirrhosis and hepatocellular carcinoma [2–4]. In the United States and most developed nations, where the prevalence of infection is 1%–2% [5, 6], HCV is the leading cause of chronic liver disease. HCV infection is expected to be even a greater public health problem in Egypt, where 10%–20% of the general population is infected [7–10].

As in other regions of the world, in Egypt some HCV infections are caused by transfusions [8, 11], injection drug use [12], and perinatal exposure [13]. However, most appear to be caused by percutaneous medical procedures (both traditional and nontraditional) [14–16]. For example, in one Egyptian village where the prevalence was 24%, 95% of antibody-positive residents had a percutaneous exposure—such as injection treatment of schistosomiasis, venous catheterization, dental treatments, and circumcision—in settings without sterilization capability [16].

Individual cases of HCV infection in Egypt may be accounted for by these percutaneous exposures. However, these practices are common to many developing countries, leading the uniquely high population prevalence unexplained. One hypothesis is that extensive nationwide transmission occurred as a result of a countrywide public health campaign to treat Schistosoma infections by inoculation [17, 18]. There is evidence of such an effect in the sharp decrease in HCV prevalence among persons born after the campaign was discontinued in the 1970s and in some (but not all) retrospective risk factor assessments [15, 16, 19–22]. It is possible to evaluate this hypothesis indirectly by examination of the distribution of HCV sequence in isolates collected around the country.

The HCV genome consists of >9500 bp [23]. The 5′ noncoding region (5′NCR) is highly conserved and is used for diagnostic testing, whereas sequences encoding the viral envelope may have <60% identity between isolates collected worldwide. On the basis of the phylogenetic relatedness of sequences from the core/E1 and NS5B genomic regions, ≥6 distinct HCV genotypes (HCV types 1–6) have been described, and each contains multiple subtypes (e.g., HCV 1a, 1b) [24–27]. An indirect method to determine the HCV genotype involving analysis of the gel migration pattern of polymerase chain reaction (PCR)–amplified 5′NCR sequence after restriction-enzyme digestion has been developed [28].

Interferon responsiveness, the pathogenesis of infection, and the performance of diagnostic assays all may vary according to genotype [29–33]. However, the distribution of genotypes...
has not been systematically studied in Egypt, the country with the highest HCV prevalence. The presence of novel subtypes in Egypt has been suggested [34], but small PCR amplicon size (152 nucleotides [nt] from NS5B) prevented reliable resolution of subtypes [35], and current recommendations for subtype assignment require data from more than 1 genomic region [36].

HCV transmission between individuals has been assessed by comparison of viral sequences [37-40]. Likewise, the hypothesis that HCV was extensively transmitted in Egypt during a 1970s countrywide campaign to treat schistosomiasis would be supported by a higher-than-expected degree of viral homology throughout the country. Previously, serum specimens from blood donors throughout Egypt were collected, carefully processed, and stored [41]. To characterize the genetic distribution of HCV throughout Egypt and examine the hypothesis of recent nationwide transmission, we ascertained the HCV genotype of specimens from the national blood donor study and analyzed the genetic relatedness of core/E1 and NS5B sequences.

Materials and Methods

Study subjects. One hundred ninety serum specimens were obtained from subjects in 15 geographically diverse governorates (table 1, figure 1). These specimens were selected on the basis of geographic diversity, serum availability, and high titers of anti-HCV antibodies. They repeatedly tested positive for antibodies to HCV, as determined by a second-generation immunoassay (HCV EIA 2.0 enzyme; Abbott Diagnostika, Wiesbaden, Germany), with signal-to-cutoff ratio of >3. A recombinant immunoblot assay (RIBA HCV 2.0 strip immunoblot assay; Chiron, Emeryville, CA) was used to confirm the specificity of highly positive HCV. In a separate subset of 75 specimens with signal-to-cutoff ratios of >3, 70 (90%) were positive by recombinant immunoblot assay (authors’ unpublished data), which is consistent with the findings of others [42, 43].

Reverse-transcription (RT) and nested PCR amplification. Total RNA was extracted from 100 μL of plasma or serum with reagent (Trizol LS; Life Technologies, Gaithersburg, MD), followed by chloroform extraction and isopropanol precipitation in the presence of 20 μg of glycogen (Boehringer Mannheim, Indianapolis). The RNA pellet was washed with 75% (vol/vol) ethanol and then air dried briefly and redissolved in 100 μL of diethyl pyrocarbonate–treated water with 6 mM dithiothreitol (Promega, Madison, WI) and 20 U of RNasin ribonuclease inhibitor (Promega). After incubation at 55°C for 10 min, 10 μL of purified RNA was used for RT and PCR in a 100-μL combined reaction mixture. This reaction mixture included 1.5 mM MgCl2, 10 mM dNTPs, 20 U of RNasin, 20 μM both first-round primers, 10 U of Moloney murine leukemia virus reverse transcriptase (Perkin-Elmer, Foster City, CA), and 2.5 U of Taq DNA polymerase (Life Technologies). After incubation at 42°C for 40 min, the reverse transcriptase was inactivated by incubation at 95°C for 4 min. DNA amplification was then done by thermal cycling for 35 cycles of 95°C for 1 s, 55°C for 15 s, and 72°C for 60 s. From this first-round reaction mixture, 10 μL was used in a second reaction with inner nested primers, with the same components and conditions as the first, omitting reverse transcriptase and RNAsis. The 5NCR primers [44] (position of 5′ base relative to the HCV-1 genome [45]) were as follows: outer: NF5 (-295), 5′-GTGGAGGAACTAATCTTTCGACCA-3′, and NR5 (7), 5′-TGCTCATGGTGACGCTTCACGAA-3′; inner: KF2 (-279), 5′-TTACGGCAAGAAGCGTGCTAG-3′, and NR4 (-8), 5′-ATACAGGCAAGCACAAGG-3′. Core/E1 primers were as follows: outer: 493S_H77 (493), 5′-GCAAACGGAAACCTCTCTGTTGCTC-3′, and 987R_H77 (987), 5′-CTTGGGACAGCCATTTCATCATCAT-3′; inner: 502S_H77 (502), 5′-AACCTTCTTCCTGGGTCCTCTCTAT-3′, and 975R_H77 (975), 5′-GTTGTCATCATATCATCCCCGATCAT-3′. For some specimens, a second set of forward primers was used: outer: CE1_F1 (346), 5′-TGGCAGAATTTGGTAAAGGTTCTGC-3′; inner: CE1_F2 (479), 5′-ACGGCGTGAATCTGAAACGGG-3′; this resulted in a product 14 nt larger than the first set. For NS5B, a single round of PCR was done with primers 242 (7904), 5′-TGGGGATCCCTGGTCATAGCCTCCGT-3′, and 243 (8304), 5′-GGCGGAATCTCGTCATAGCCTCCG-3′ [46].

Subtype determination by means of restriction fragment length polymorphism (RFLP) analysis. Subtype determination was done by use of HaeIII-RsaI and MvaI-HinDII digestions of the 5′NCR PCR product [47]. For specimens designated genotype 1 by the method above, a second digestion with BsrUI was done to distinguish subtypes 1a and 1b [48].

DNA purification and nt sequencing. Specimens for sequencing were purified from low-melting-temperature agarose by use of a gel extraction kit (Qiagen; Qiagen, Chatsworth, CA) according to the manufacturer’s protocol. Sequencing was done with an automated sequencer (PRISM, version 2.1.1; ABI, Foster City, CA). The inner forward PCR primer was used to determine all sequences, and 46 of 68 core/E1 sequences were confirmed by use of the reverse primer. Sequences were assembled with the BioEdit program, version 7.1 (T. Hall, North Carolina State University, Raleigh; available at http://www.mbio.ncsu.edu/RNaseP/info/programs/BIOEDIT/bioedit.html), and primer sequences were removed. The

<p>| Table 1. Results of restriction fragment length polymorphism (RFLP) genotyping of hepatitis C virus isolates. |</p>
<table>
<thead>
<tr>
<th>Governorate</th>
<th>Region</th>
<th>n</th>
<th>NA</th>
<th>1a</th>
<th>1b</th>
<th>4</th>
<th>NT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alexandria</td>
<td>North Central</td>
<td>6</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Assiut</td>
<td>South</td>
<td>21</td>
<td>15</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Beheira</td>
<td>North Central</td>
<td>21</td>
<td>12</td>
<td>1</td>
<td>0</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>Beni Suef</td>
<td>South Central</td>
<td>14</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Cairo</td>
<td>Central</td>
<td>13</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Daqahliya</td>
<td>North Central</td>
<td>15</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Damietta</td>
<td>North Central</td>
<td>10</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Qalyubiya</td>
<td>North Central</td>
<td>16</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Qena</td>
<td>South</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>9</td>
<td>1</td>
</tr>
<tr>
<td>Matrouh</td>
<td>West</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Menoufiya</td>
<td>North Central</td>
<td>16</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>11</td>
<td>0</td>
</tr>
<tr>
<td>Minya</td>
<td>South Central</td>
<td>21</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>0</td>
</tr>
<tr>
<td>Red Sea</td>
<td>South East</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>South Sinai</td>
<td>North East</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Suez</td>
<td>South Central</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>190</td>
<td>68</td>
<td>1</td>
<td>1</td>
<td>111</td>
<td>9</td>
</tr>
</tbody>
</table>

NOTE. NA, not amplified; NT, not typed by this method.
* As depicted in figure 1.
Figure 1. Map of Egypt, indicating regions from which specimens were obtained. Governorates composing each region are listed in table 1. Below each region label, each symbol represents a serum specimen from that region and hepatitis C virus subtype assigned by both restriction fragment length polymorphism (RFLP) and phylogenetic analysis (squares) or RFLP only (circles). The asterisk (*) overlying square in North Central region indicates only the specimen for which RFLP and phylogenetic analysis were discordant (subtype 1a by RFLP, subtype 4a by phylogenetic analysis).

GenBank accession numbers for the sequences presented herein are AF271797–AF271891.

Phylogenetic analysis. Phylogenetic trees were generated by use of the PHYLIP suite of programs, version 3.572c [49]. Genetic distance matrices were calculated with the DNADIST program, maximum likelihood model, with a transition-to-transversion ratio of 4.25 [50]. The distance matrix was then used to generate a tree by use of the neighbor-joining algorithm as implemented in NEIGHBOR. In a similar manner, bootstrap confidence values were calculated after randomly permuting the sequence alignment 100 times with the SEQBOOT program [51]. Permuted trees were generated by use of the NEIGHBOR program with random addition, and a consensus topology was derived by use of CONSENSE. Nonsynonymous and synonymous genetic distances were calculated with the method of Nei and Gojobori [52].

Reference sequences for the different genotypes of HCV were obtained from GenBank. The designations of subtypes 1a, 1b, 1c, 2a, 2b, 3a, 3b, 4a, 5a, and 6a are generally accepted. Other designations are more controversial and are enclosed in quotation marks herein. GenBank accession numbers (strain designations) for the reference sequences used are as follows: 1a, M62321 (HCV-1); 1b, D90208 (HCV-J); 1c, D14853 (G9); “1d,” L38378 (N16); “1e,” L38361 (CAM1078); “1f,” L38350 (FR2); 2a, D00944 (J6); 2b, D10988 (J8); 2c, D50409 (BEBE1); 3a, D17763 (NZL1); 3b, D49374 (Tr Kj); 4a, Y11604 (ED43); “4b,” U10235, L16677 (Z1); “4c,” L38338, L29607 (GB358); “4d,” U10192, L16656 (DK13), D86537 (SD006) D86538 (SD008); “4e,” L29587, L29590 (CAM600); “4f,” L29597 (G27) L29592 (CAMG22); “4g,” L29617, L29618 (GB549); “4h,” L29610, L29611 (GB438); “4i,” L36439, L36437 (CAR4/1205); “4j,” L36438 (CAR1/501); “4l,” D86534 (SD002); “4l,” D86539 (SD015); “4m,” D86543 (SD035); 5a, Y13184 (H1480); 6a, Y12083 (HK2); “6b,” D84262 (Th580); “7a,” D84263 (VN235); “7b,” D84264 (VN405); “9a,” D84265 (VN004); “10a,” D63821 (JK049); “11a,” D63822 (JK046) [24, 25, 35, 51–64].

Results

In another study, HCV antibodies were detected in 656 (24.8%) of 2644 specimens from blood donors in 24 of 26 Egyptian governorates [41]. From these, 190 geographically representative specimens from 15 governorates were selected for genetic testing.

Amplification of the 5’NCR. The 5’NCR was amplified in 122 (64.2%) of 190 specimens. Table 1 shows the number of specimens amplified from each governorate. After correction
for multiple comparisons, no statistically significant difference was detected in the proportion of specimens in which HCV could be amplified in each governorate. The RFLP patterns of the 122 PCR products were analyzed, and genotypes were assigned: 111 (91%) were consistent with genotype 4, 1 (1%) was subtype 1a, 1 (1%) was subtype 1b, and 9 (7%) could not be typed by this method.

Amplification of core/E1 and NS5B regions. A subset of specimens amplified with 5’NCR primers was selected for amplification of the core/E1 region to permit further analysis. This subset consisted of the 9 specimens for which the restriction fragment pattern did not fit a reference standard (nontypeable), 2 that were classified as genotype 1, and an additional 61 with an RFLP pattern consistent with genotype 4. The latter specimens were chosen with a goal of obtaining at least 5 specimens per governorate. By use of the core/E1 primers, 68 (94%) of 72 specimens were amplified. Analysis of the NS5B region also was done for 43 of these specimens, including the 9 nontypeable and 2 genotype 1 specimens assessed by RFLP; 20 (47%) of 43 specimens were amplified with the NS5B primers. All NS5B sequences lacked nonsense substitutions, whereas 2 of 68 core/E1 sequences had single nonsense substitutions. No length polymorphism was observed.

Phylogenetic analysis of the core/E1 and NS5B regions. To confirm and complete the genotype assignments, these 68 amplified specimens were studied by use of phylogenetic analysis. Phylogenetic trees constructed from these sequences demonstrated that all of them clustered with genotypes 1 and 4, consistent with the RFLP results, and revealed 3 new HCV subtypes (figure 2).

Most of the Egyptian sequences clustered with genotype 4, forming 5 distinct groups. The largest group, comprising 48 sequences on the core/E1 tree and 13 sequences on the NS5B tree, clustered with the subtype 1a reference sequence ED43. The mean (± SD) genetic distance (in substitutions per position) within this cluster was 0.11 ± 0.02 for core/E1 and 0.09 ± 0.02 for NS5B, consistent with isolates within a subtype [35]. This close relationship was not surprising, because ED43 was obtained from an Egyptian who received multiple blood transfusions in Cairo in 1982 [60]. The remaining type 4 sequences (14 in core/E1 and 4 in NS5B) formed 4 distinct groups with strong (100%) bootstrap support. Two of these groups clustered with Middle Eastern NS5B reference sequences provisionally identified as subtypes “4f” and “4m” [62]. Because these sequences were included in the core/E1 tree, their clades on that tree were also designated “4f” and “4m” (figure 2A). The other 2 groups were given arbitrary identifiers “4e” and “4g,” because they did not cluster with any reference sequence, suggesting that they may belong to previously unrecognized subtypes. These groups’ genetic distances from reference sequences for genotype 4 were comparable to the distances among the reference sequences (table 2), demonstrating a star phylogeny (branches of equal length surrounding a central node).

When smaller NS5B amplicons from previous studies of Egyptian sequences [27, 34] were included in separate phylogenetic analyses, sequences from 2 subjects with liver disease (Segi44 and Segi193) were strongly associated with the “4β” sequences (bootstrap value 98% [data not shown]). The small size of the amplicons and lack of sequence from other genomic regions prevented further correlation. BLAST utility search of the GenBank database by use of each “4α” and “4β” sequence did not reveal any sequence with homology >81% in core/E1 or >84% in NS5B.

A minority of Egyptian sequences clustered with genotype 1 reference sequences. One of these clustered with the subtype 1b reference sequence HCV-J but was not represented on the core/E1 tree. The other genotype 1 sequences clustered on both trees as a monophyletic group (bootstrap value 99%–100%), which was designated “1g” because it was distinct from previously characterized subtypes 1a–“1f.” The mean (± SD) genetic distance within the “1g” cluster was 0.11 ± 0.03 for core/E1 and 0.06 ± 0.01 for NS5B, consistent with isolates within a subtype [35]. The genetic distance between the “1g” group and reference sequences was 3–4 times larger (table 3), consistent with distances among subtypes. No core/E1 reference sequence for subtype “1e” is available, and a BLAST utility search of the GenBank database by use of each “1g” sequence did not reveal any core/E1 sequence with homology >73%.

Comparison of genotyping results. Of 61 specimens for which a genotype assignment was made on the basis of >1 region, 60 were concordant. A specimen from Beheira was assigned subtype 1a by RFLP but was assigned genotype 4a by phylogenetic analysis of the core/E1 sequence. Therefore, there were no subtype 1a specimens detected among the 122 specimens assessed. Of the 9 specimens nontypeable by RFLP, 3 were subtype 4a by phylogenetic analysis, 1 was not amplified in either the core/E1 or NS5B PCR, and the other 5 were subtype “1g.” Sequencing of the 5’NCR from the 5 subtype “1g” specimens revealed the presence of a C at position −167 (relative to the HCV-1 genome [45]) rather than the T present in prototypic genotype 1 sequences, resulting in an additional HinII site with cleavage between bases −171 and −170. This HinII site was present in the sequences from subtypes 3a, 3b, 4a, 5a, and 10a; however, subtype “1g” sequences had an Msel site with cleavage between positions −164 and −163 (like prototypical genotype 1 sequences). The effect on Msel-HinII RFLP analysis was reduction in size of the 63-nt band by 7 bases, yielding a new pattern that was uniquely shared by the subtype “1g” specimens. Both C and U at position −167 can pair with the complementary G residue at −118 in the stem of loop IIIb in the HCV internal ribosome entry site [65, 66].

Lack of regional clustering. As illustrated in figure 1, subtype 4a was the majority subtype in each governorate examined. In addition, inspection of the core/E1 and NS5B trees revealed no clustering by geographic origin. For instance, the Egyptian subtype “1g” clade was composed of sequences from 5 different
Figure 2. Phylogenetic trees based on hepatitis C virus (HCV) sequences obtained via reverse transcription–polymerase chain reaction of serum specimens from Egypt. Trees were generated by Fitch-Margoliash method after phylogenetic bootstrap analysis with neighbor-joining method, based on 100 permutations of sequence alignment. A, Tree for core/E1 region; B, tree for NS5B region. Reference sequence identifiers are followed by their subtype designations, enclosed in quotation marks when designation is provisional. Core/E1 tree includes reference sequences from genotypes 1 and 4, with subtype 5a as outgroup. Because inclusion of other reference sequences had no effect on clustering of experimental sequences (data not shown), they were omitted from figure. Numbers adjacent to bifurcations indicate number (of 100 trees) in which cluster was observed (values <60 omitted).

governorates, and, within the Egyptian subtype 4a clade, there was no clustering of sequences from the same region or governorate. Because of the limited number of non–subtype 4a sequences, we also performed an analysis to detect correlation between geographic and genetic distance within subtypes. Figure 3 shows the genetic distance for each pair of sequences plotted versus the geographic distance between the capitals of their respective governorates by use of only within-subtype comparisons. This plot reflects the findings from the phylogenetic trees, in that the genetic distances varied over the same range in core/E1 and NS5B, for sequences from both genotypes 1 and 4. Despite a large number of comparisons over a broad range of geographic distances (0–1000 km), there was no correlation between genetic and geographic distance.

Discussion

The chief objective of this investigation was to characterize the genotype distribution of HCV in Egypt, the country with the highest worldwide prevalence. The importance of this aim relates to differences in interferon responsiveness, natural history of infection, and the performance of diagnostic tests attributed to HCV genotypes [29, 31–33]. There is also evidence that the different HCV subtypes differentially inhibit the interferon-induc-
1. On the NS5B tree, subtype "1g" was most closely related to and all other type 4 reference sequences.

Table 2. Genetic distances between Egyptian subtypes of hepatitis C virus genotype 4 and reference sequences.

<table>
<thead>
<tr>
<th>Region of genome</th>
<th>Mean (SD) genetic distance from sequences of other subtypes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>&quot;4a&quot;</td>
</tr>
<tr>
<td>Core/E1</td>
<td>0.31 (0.04)</td>
</tr>
<tr>
<td>NS5B</td>
<td>NA</td>
</tr>
</tbody>
</table>

NOTE. NA, not applicable (no "4a" NS5B sequence was obtained). Genetic distance is defined as no. of substitutions per position.

Table 3. Genetic distances between Egyptian hepatitis C virus subtype "1g" and reference sequences.

<table>
<thead>
<tr>
<th>Region of genome</th>
<th>Mean (SD) genetic distance between &quot;1g&quot; sequences and reference sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1a</td>
</tr>
<tr>
<td>Core/E1</td>
<td>0.46 (0.03)</td>
</tr>
<tr>
<td>NS5B</td>
<td>0.19 (0.02)</td>
</tr>
</tbody>
</table>

NOTE. Genetic distance is defined as no. of substitutions per position. NA, not applicable; no subtype reference sequence available for this region.
dependent on the substitution rates that are used. Higher rates had previously been estimated from HCV variation in a single person over a 13-year period [72] and a chimpanzee over an 8-year period [73]; however, the rates based on a human cohort, especially those based on synonymous substitutions, are likely to be superior [50].

Estimates of divergence times must be interpreted with extreme caution [74]. The apparently long time of divergence for Egyptian HCV sequences may be due to the original introduction of multiple HCV lineages with preexisting differences, because molecular clock calculations assume that differences among the sequences have occurred only since the introduction of a single ancestor. In addition, for sequences more distantly related than sequences of the same type, the effects of saturation begin to result in falsely reduced genetic distance and may operate in some regions within a subtype [75].

In addition to the broad distribution of genetic distances that was found, regional clustering of HCV sequences was not detected. In fact, there was no relationship between genetic and geographic distances. Such a finding might occur if there had been substantial transmission among regions. It is not known whether injection equipment and medication was transported throughout the country or used just within a single district. However, such practices would explain these data. Significant population movement could also explain these findings, and segments of the Egyptian population are highly mobile for employment reasons. Alternatively, the Egyptian HCV epidemic may be much older than is currently presumed. A larger nationwide study of persons of various ages is needed to evaluate these hypotheses further.

A limitation of all molecular virologic studies is that only the isolates that amplify can be examined. In this study, the proportion of specimens with positive results by use of PCR of the 5’NCR varied widely among the governorates. The data are consistent with stochastic (random) variation but do not exclude the possibility that variants present in some regions are less amenable to amplification with these PCR primers. However, other investigators have found that these primers amplify a broad range of specimens [22, 44]. We did not find a correlation between the proportion of specimens positive by PCR of the 5’NCR in a

### Table 4. Mean pairwise genetic distances between hepatitis C virus sequences of same subtype.

<table>
<thead>
<tr>
<th>Region, subtype</th>
<th>All sites</th>
<th>Synonymous sites</th>
<th>Nonsynonymous sites</th>
</tr>
</thead>
<tbody>
<tr>
<td>Core/E1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;1g&quot;</td>
<td>0.114</td>
<td>0.339</td>
<td>0.054</td>
</tr>
<tr>
<td>4a</td>
<td>0.102</td>
<td>0.330</td>
<td>0.041</td>
</tr>
<tr>
<td>NS5B</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;1e&quot;</td>
<td>0.064</td>
<td>0.173</td>
<td>0.031</td>
</tr>
<tr>
<td>4a</td>
<td>0.091</td>
<td>0.308</td>
<td>0.034</td>
</tr>
</tbody>
</table>
governorate and the subtype that predominated there. Such differences might be explained by regional differences in specimen handling, differences in levels of viremia (due to host or virologic factors), or both. The former is supported by the observation of visible hemolysis in specimens from Assiut, where the proportion of PCR-positive specimens among anti-HCV-positive subjects was lowest. Amplification of the core/E1 region was very sensitive in the specimens tested, but these were restricted to specimens that had already tested highly positive by EIA and were also positive by PCR amplification of the 5’NCR.

We conclude from this study that the Egyptian HCV epidemic is composed of multiple lineages of genotypes 1 and 4. Despite a large, geographically diverse sample and molecular analysis of 3 genomic regions, we were unable to discover any regional associations that would suggest a pattern of spread. One explanation for the lack of such a pattern might be simultaneous dissemination of multiple HCV strains to the entire population, as might occur in a vaccination or nationwide treatment program that involves injection.

Acknowledgments

We thank Qing Mao and Oliver Laeyendecker for technical assistance and Tom Hall (Department of Microbiology, North Carolina State University, Raleigh) for making the BioEdit program freely available.

References

34. Rapicetta M, Argentini C, Dettori S, Spada E, Pellizzer G, Gandin C. Mor-