
Shoshanna Maya\textsuperscript{a,1}, Siew Lin Ngui\textsuperscript{a,1}, Sarah Collins\textsuperscript{b,1}, Sam Lattimore\textsuperscript{b,1}, Mary Ramsay\textsuperscript{b,1}, Richard S. Tedder\textsuperscript{a,c,d,1}, Samreen Ijaz\textsuperscript{a,*,1}

\textsuperscript{a} Blood Borne Virus Unit, Microbiology Service – Colindale, Public Health England, 61 Colindale Avenue, London NW9 5EQ, UK
\textsuperscript{b} Immunisation, Hepatitis and Blood Safety Department, Centre for Infectious Disease Surveillance and Control, Public Health England, 61 Colindale Avenue, London NW9 5EQ, UK
\textsuperscript{c} Division of Infection and Immunity, University College London, Gower Street, London WC1E 6BT, UK
\textsuperscript{d} Transfusion Microbiology, NHS Blood and Transplant, Colindale Avenue, London, NW9 5BG, UK

\textbf{Article info}\\
\textbf{Article history:} \\
Received 17 October 2014 \\
Received in revised form 19 December 2014 \\
Accepted 23 December 2014

\textbf{Keywords:} \\
Hepatitis C virus \\
Molecular epidemiology \\
Genotype \\
Risk/behavioural factors

\textbf{ABSTRACT}\\
\textbf{Background:} Analysis of laboratory testing data collected through the Sentinel Surveillance programme has provided a method for identifying individuals who have recently acquired their hepatitis C virus (HCV) infection. Access to samples from these individuals provided a rare opportunity to undertake molecular characterization studies.

\textbf{Objectives:} To describe the epidemiology and genetic diversity of hepatitis C in recent seroconverter infections and to predict how this will impact on HCV treatment and control.

\textbf{Study design:} One hundred and forty seven samples were available from individuals, identified to have recently acquired their HCV infection. Genotype determination with additional phylogenetic analysis was carried out on NS5B sequences. Analysis across the NS3 region investigated the presence of antiviral resistance mutations. Where possible, molecular data was linked to demographic and risk/behavioural factor information.

\textbf{Results:} The majority of new infections occurred in males with a mean age of 37 years. The most commonly observed genotypes were 1a (49%) and 3a (42%) and injecting drug use (58%) was the most common risk factor. Genotype distribution differed between persons who inject drugs and those with other risk factors suggesting two possible epidemics. Phylogenetic analysis indicated possible transmission networks within specific risk groups. Amino acid changes associated with antiviral resistance were noted in the NS3 region in some samples.

\textbf{Conclusions:} Continued surveillance of linked molecular, virological, demographic and epidemiological information on recently acquired infections will contribute to understanding the on-going HCV epidemic in England.

\textcopyright 2015 Elsevier B.V. All rights reserved.

1. Background

Preventing the transmission of hepatitis C virus (HCV), remains a key priority for reducing prevalent infections and subsequent sequelae of HCV-related disease in England. Primary interventions aimed at reducing incident infections by the provision of sterile injecting equipment and reducing injection through opiate substitution therapy have had an impact on the prevalence of HCV in persons who inject drugs (PWIDs) [2,3,17]. In addition, treatment, which leads to viral clearance will clearly impact on both the burden of disease and onward transmission [11].

As acute HCV infections are frequently asymptomatic, a number of methods have been used to estimate HCV incidence [2]. The sentinel surveillance of blood borne virus testing has collected information on testing for viral hepatitis, HIV and HTLV-1, from 24 sentinel laboratories across England. This has formed a valuable platform enhancing routine national surveillance by monitoring both confirmed infections and testing [1]. More recent methodological developments in data retrieval have allowed the identification of recent hepatitis C infections through the demonstration of seroconversion among individuals undergoing repeat testing [9].
The Sentinel Surveillance provides a novel way of identifying HCV incident infections in persons undergoing regular testing but importantly gives a rare opportunity to undertake molecular characterization of viruses linked to incident infections. These data are currently lacking but can provide a unique insight into recently acquired HCV infections and their contribution to the on-going HCV epidemic in England. Linked epidemiological and molecular investigations in this population can inform on possible transmission networks through risk and phylogenetic analysis and therefore instruct on possible targets for intervention. Understanding the genotype distribution amongst recent infections is also an essential factor in informing prevention measures and determining their effectiveness. These data also have the potential to inform on estimates for the future burden of chronic disease in this country.

2. Objectives

We describe here the epidemiology and genetic diversity of hepatitis C among individuals with evidence of a recently acquired infection (seroconverters) and attempt to predict how these data will impact both on the treatment and the control of HCV infections nationally.

3. Study design

3.1. Patients and samples

Patients and samples included in this investigation were from 24 testing laboratories participating in the English Sentinel Surveillance Programme and were identified between 2008 and 2011. Individuals were considered to have recently acquired their HCV infection on the basis of demonstrating seroconversion for antibody to HCV (anti-HCV) in serial samples within three years of a previous negative test result. Over this period 610 recent HCV infections were identified, 375 of which were confirmed as true seroconverters. The identification and confirmation of these seroconverters are to be discussed elsewhere. Briefly, HCV seroconversion could not be reliably confirmed in the remaining 235. Individuals with a previous negative and first antibody positive result from different sentinel laboratories were excluded. Additional information including RNA testing was requested and cases were excluded where this was not available. The time elapsed between last negative and first positive across the seroconverters was a median of 365 days (mean—416 days; IQR—175–616 days). Those diagnosed within three years of a negative would therefore be a mixture of acute and chronic infections, but all would be recently acquired.

One hundred and forty seven samples from these individuals were available for molecular characterization studies. Demographic and risk factor information associated with these patients were also collected where available. As self-reported ethnicity was not generally available, ethnicity was determined using name analysis software.

3.2. Molecular characterisation

Genotyping was performed by restriction fragment length polymorphism [16] and NS5B sequencing [10]. Total nucleic acid was extracted from 200 μl of plasma on the Qiagen Universal BioRobot automated extraction platform using the QIAamp one-for-all nucleic acid protocol (Qiagen, UK). Additional PCR amplification across the NS5 region was undertaken on all genotype 1a samples. A one-step protocol was used for first round amplification of NS3 sequences using the following primers: NS3p1aOS (ATATAGGACCGGTGAGTTTG) and 1abgIIAs (ATATAGTCTCATGGTGCYCTTAGG). Sequencing primers for the NS3 region were: 3df (GACAAAAACAGTGAGGAGG) and 3Cr (AATGACA TCGGCCGTCCCTCG) 1aRIN1 (TATATGCAAGGGGATG) and 1aN2 (TGATCCTCAAGGGGTG) [19]. The NS3 and NS5B PCR amplicons were purified and sequenced on the ABI 3730 automated capillary sequencer (ABI, Paisley, UK).

3.3. Analyses

Sequences were aligned in Bioedit version 7.0.9 [15], using the ClustalW alignment tool. Phylogenetic analyses of NS5B and NS3 region sequences were conducted using Mega version 4 [32] and the HCV genotypes defined by NS5B phylogenetic analysis. NS3 sequence alignments were scanned for previously published polymorphisms considered to confer phenotypic resistance to directly acting protease inhibitors [6,12,14,21,22,28]. Statistical analysis was performed in Stata version 13.1.

4. Results

4.1. Demography of patients

Gender information was available for 116 of the 147 individuals and over two thirds were male (84/116, 72%). The mean age of patients was 37 years ranging from 18 to 84 years, 84% of infections occurred in people aged 20–49 years. Ethnicity was assigned for 111 individuals, the majority of whom were white-British (71%), followed by South Asian (12%), white-Irish (8%) and white-other (9%).

4.2. Risk exposures

Of the 85 cases with known risk and behavioural factor information associated with exposure, slightly over half were PWIDs (56/85, 66%) of whom two were also known to be HIV positive. Recent HCV infections were also observed in men who have sex with men (MSM; 13/85, 15%), ten of whom were known to be HIV-infected; the HIV status was unknown for the remaining three. Two individuals (2/85, 2%) had potential exposure through both injecting drug use and sex between men (one of whom was also HIV co-infected). Other potential risk factors associated with infection included dialysis abroad (7/85, 8%), dialysis (3/85, 4%), sexual exposure (3/85, 4%) and needlestick injury (1/85, 1%). HIV co-infection was noted in a further eight cases in whom there was no declared risk.

Association analysis between risk/behavioural factors and basic demography was undertaken where the linked information was known (fisher’s exact test p = 0.005). The use of injecting drugs was the predominant risk in both males (36/57, 53%) and females (7/14, 50%). Dialysis/dialysis abroad was shown to be the main risk factor associated with recent HCV infection in the majority of the South Asian patients within this study (8/10, 80%).

4.3. HCV genotype distribution

The HCV genotype could be determined in samples from 105 of 147 patients. Viral load was either not detected or too low for HCV genotype analysis in the remaining 42 samples. The majority of the viruses belonged to either genotype 1a (49/105, 47%) or 3a (47/105, 45%) with the remainder being genotypes 1b (3/105, 3%), 2a (1/105, 1%), 2b (3/105, 3%), 3i (1/105, 1%) and 4d (1/105, 1%).

Genotype 1a was the dominant infection in males (36/63, 57%), whilst 3a was dominant in females (15/23, 65%) (fisher’s exact test p = 0.027).
Table 1
Risk/behavioural factor data linked to HCV genotypes 1a and 3a. Data shown in square brackets \([n]\) indicates the number of cases known to also be HIV-infected.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>PWID</th>
<th>MSM</th>
<th>Dialysis</th>
<th>Dialysis abroad</th>
<th>Other</th>
<th>Unknown</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a</td>
<td>12 (30%)</td>
<td>8 [7] (80%)</td>
<td>3 (100%)</td>
<td>4 (67%)</td>
<td>5 (71%)</td>
<td>17 [4] (44%)</td>
</tr>
<tr>
<td>3a</td>
<td>24 [1] (60%)</td>
<td>0</td>
<td>0</td>
<td>2 (33%)</td>
<td>2 (29%)</td>
<td>19 [1] (48%)</td>
</tr>
<tr>
<td>Other</td>
<td>4 (10%)</td>
<td>2 [1] (20%)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3 (8%)</td>
</tr>
<tr>
<td>TOTAL</td>
<td>40</td>
<td>10</td>
<td>3</td>
<td>6</td>
<td>7</td>
<td>39</td>
</tr>
</tbody>
</table>

Fig. 1. Phylogenetic tree of 328 bp region of NS5B from genotype 1 (a) and genotype 3 (b) sequences. Those sequences linked to this study are prefixed by ‘HCV’; GenBank Accession numbers and the associated genotype are indicated. Possible linked clusters or pairs based on a nucleotide similarity of between 97.8–100% are indicated.
There were striking differences in the genotype distribution between risk/behavioural groups (Table 1) (Fisher’s exact test \( p = 0.002 \)). PWIDs were more likely to harbour genotype 3a infections whilst individuals within the non-PWID related risk groups were more commonly found to harbour genotype 1a viruses.

Additional genotype distribution analysis indicated genotype 1a to be found in 11 of the 13 HIV-infected individuals in whom genotype information was available, seven of whom were known to be MSMs.

### 4.4. HCV phylogeny

NS5B sequence information was obtained for 94 of the 105 viruses. Complete concordance was seen between RFLP and NS5B genotyping. Analysis identified one possible transmission cluster involving eight sequences and six pairs as being potentially linked on the basis of the viruses within each pair or cluster being 97.8–100% homologous at the nucleotide level in the NS5B region (Fig. 1a and b; Table 2). The cluster and five of the pairs comprised genotype 1a viruses. Where data were known, it showed the cluster to consist of seven males, four of the seven were known to be HIV co-infected and one was an MSM (Table 2). Four of the five genotype 1a pairs were male and identified as either HIV co-infected or PWIDs; the fifth pair was made up of two females and both had undergone dialysis. The final 3a pair was comprised of males who were known to be PWIDs.

### 4.5. Analysis of the NS3 region — identification of protease inhibitor resistant variants

Sequencing of the NS3 region in 34 genotype 1a samples showed nine (28%) samples to have the reported antiviral protease resistance mutations present as the dominant population. One sample had the polymorphisms: V36M and Q80K; and a second sample had two polymorphisms: V36M and V55A. Six samples had the Q80K polymorphism, and one sample contained the Q80R/Q polymorphism.

### 5. Discussion

Analysis of laboratory testing data collected through the Sentinel Surveillance Programme has provided a method for identifying individuals who have recently acquired their HCV infection. This procedure depends on accurate identification of results from the same individual; and therefore, may miss seroconversions in individuals who are tested in different settings when patient identifiers are not routinely used. Identification of recent infections using this method may have some bias towards individuals within certain risk groups as they are more likely to have repeat HCV tests; and therefore, more likely to be identified when they are infected. However, it is the access to samples from these individuals that has provided a rare opportunity to undertake molecular characterisation studies.

The demographic profile of individuals identified to have a recent seroconversion infection was similar to previous reports of laboratory confirmed HCV cases in England with the majority of infections to be in males aged between 25 and 39 [2]. The most commonly observed genotypes associated with the recent infections were 1a (49%) and 3a (42%) again reflecting the prevalent genotype estimates reported nationally [2,20]. Our data confirm injecting drug use continues to be the most common risk factor contributing to new infections in England in both males and females and for genotype 3a viruses to be commonly found in PWIDs.

MSM was reported for only 10 of the 21 (48%) new HCV infections in individuals who were HIV-infected. Infections in HIV-infected individuals were predominately due to genotype 1a viruses and found only in males possibly reflecting un-reported sex between men. In addition to drug use, the observation of transmission in known HIV positives is not surprising with recent studies from high income countries have demonstrated a high incidence of recent HCV infections in HIV-infected MSMs [8,18,23,25–27,29–31,33–36]. Recent PHE data demonstrated a 21% increase in new HCV infections amongst MSMs in London since 2011 [20]. Whilst some of this rise was attributed to an increase in testing for HCV infection the report also suggested that an increase in high-risk drug and sexual practices leading to a rise in HIV transmission could also be leading to a rise in HCV infections.

The finding that infections in non-PWIDs were more likely to be genotype 1a is useful in supporting a lack of overlap between infections in the population of HIV infected MSMs and PWIDs, and suggests the concept of two possible epidemics. This is further supported by the identification of the 1a cluster of closely related viruses possibly linked to the same transmission network. HCV sequences with a divergence of between zero and 2.2% have
previously been considered to indicate possible linkages between infections [24]. Based on this parameter, our sequence data identified a further six pairs of sequences that were very closely related (Fig. 1) which appeared to be linked to specific risk groups. However, cluster analysis based on short HCV sequences may not be totally reliable and further investigations with additional epidemiological data would need to be undertaken in order to confirm any links. Nevertheless, these data show the value of molecular phylogenetic studies in recent infections which can guide public health interventions.

No risk factor information was available for 39 individuals where genotype information was available. It is tempting to extrapolate from the genotype possible routes of transmission. However, what this highlights is the importance of ascertaining risk data in order to understand any emerging or changing risk/behavioural practices and inform on possible interventions.

Viruses bearing single nucleotide polymorphisms (snps) associated with NS3 protease resistance have been reported in persistently infected treatment-naïve patients [4–7,12,13,22]. The snps identified in this study were those associated with low-level resistance to Telaprevir and Boceprevir. In the context of this study, it is not possible to confirm that these snp-bearing viruses were transmitted from a treated individual. The snps could have been independently arisen during HCV replication in the infected individuals. Nevertheless, the detection of snps that are phenotypically resistant to directly acting drugs indicates that they could impact on the virological response to treatment in the naïve patient. The Q80K polymorphism was also noted in 18% of the genotype 1a samples sequenced across the NS3. The pre-existence of the Q80K polymorphism at baseline has been shown to have a substantial impact on the efficacy of Simeprevir, one of the ‘second wave’ protease inhibitors.

The findings from this study may impact on future planning for HCV interventions in terms of prevention and management. Available data have shown intervention measures in PWIDs can improve outcomes in HCV interventions in terms of prevention and management.

Competing interests
None.

Ethical approval
Not required.

Acknowledgements
Laboratories Contributing Data to the Sentinel Surveillance Programme: Justin Dawkins, Hamid Jalal, Kate Rolfe, Addenbrookes Hospital, Cambridge; Chas Ashley, Peter Muir, David Wright, Bristol Regional PHE Laboratory; Rolf Meigh, Castle Hill Hospital, Cottingham, Hull; Mark Atkins, Jeremy Merritt, Chelsea and Westminster Hospital, London; John Croll, Chester Microbiology Laboratory, Countess of Chester Hospital, Chester; Tony Vicca, Diana Princess of Wales Hospital, Grimsby; Ferial Ahmad, Imad Ibrahim, Eating Hospital, London; Julia Taylor, Sheila Waugh, Freeman General Hospital, Newcastle; Tracey Leech, Celia Pennman, CIDS, Public Health England, London; Manoj Vallapil, Newcastle PHE laboratory, Newcastle General Hospital, Newcastle; Elizabeth Boxall, Janet Mowbray, Erasmus Smit, West Midlands PHE laboratory, Heart of England Foundation Trust, Birmingham; Antony Hale, Mike Wallis, Leeds Teaching Hospitals NHS Trust, Leeds; David Johnson, Mark Zuckerman, Kings College Hospital, London; Aide Finch, Paul Klappler, Ken Mutton, Manchester Medical Microbiology Partnership, Manchester Royal Infirmary, Manchester; Mohammed Osman Hasan Ibrahim, Royal Sussex County Hospital, Brighton; Will Irving, Lisa Prichett, Yursi Taha, Queens Medical Centre, Nottingham; Josephine Silles, PHE Collaborating centre, North Middlesex University Hospital, London; Louise Hesketh, Microbiology Laboratory, Royal Preston Hospital, Preston; Lynne Ashton, Ian Hart, Royal Liverpool Hospital, Liverpool; Tony Oliver, Derren Ready, Barts and the London NHS Trust, London; Hasan Al-Ghusein, Phil Rice, St George’s Hospital, London; Alan Reid, Gillian Underhill, St Mary’s Hospital, Portsmouth; Mike Kidd, Ian Riddoch, Univeristy College London; Mark Baker, James Nash, William Harvey Hospital, Ashford, Kent.

Conflict of interest
No conflict of interest.

Funding

References