# Supplement 1

# Full narrative of published cases

The forensic context from Table 1 column 1 is mentioned in bold

## Nosocomial Transmission

The first documented forensic use of HIV phylogenetics was in 1990s, in which a dentist from Florida, who knew his HIV-1 positive serostatus, allegedly infected several of his patients during routine dental treatment (**First Florida Dentist case in Table 1**) [1]. After identification of an HIV positive patient who had no known risk factors for HIV infection but who had undergone an invasive procedure performed by that HIV positive dentist, six other patients of this dentist were found to be HIV-infected. This case was a high-profile investigation, inferring the phylogenetic relationships of HIV-1 in different individuals. Ou *et al*. [1] performed a phylogenetic analysis of HIV strains infecting the dentist, the patients and local HIV-positive controls, that showed a closer epidemiological link between 6 of the 7 patients and the dentist than between the 6 patients or the dentist and local HIV positive controls [1,2]. This was a civil case where the patients brought the case against the dentist’s insurer, who settled the case out of court and paid the 6 HIV-infected patients with a closer epidemiological link to the dentist. However, this case was highly controversial. DeBry *et al*. [3] criticized the original analysis for using inappropriate phylogenetic methods and inadequate sampling of local controls. According to DeBry*’s* phylogenetic analyses, there was no evidence to support the hypothesis of infection via the dentist. Both the analyses of Ou *et al*. [1] and De Bry *et al*.[3] were based on the maximum parsimony procedure for phylogeny reconstruction. Later on, Crandall [4] used an alternative method, combining cladogram estimation and a nested analysis framework, and found solid support for what was originally concluded by Ou *et al* [1]. In later years, concerns remained about the reliability of the phylogenetic evidence that was used in this case [5].

In 1991, a **second Florida dentist case** was published [6]. A dentist, upon diagnosis with AIDS, disclosed that he frequently undertook invasive procedures and did not always follow recommended infection-control procedures. An epidemiological investigation was conducted, where letters were sent to 6474 patients of the dentist. Of these, 1274 patients were tested for HIV and 24 were found to be positive (1.9%). Phylogenetic analysis of HIV sequences from this dentist and his HIV-positive clients was subsequently conducted. The analyses did not show a strong bootstrap support for monophyletic clustering between the clients and the dentist, so the phylogenetic evidence could not support the claim that the dentist had infected some of his 28 HIV-positive patients.

As in the Florida dentist case, there have also been other forensic phylogenetic investigations of HIV transmission where health care workers were involved. Goujon *et al* [7] reported a case where a HIV-1 negative patient with no risk factor, experienced HIV-1 primary infection 4 weeks after being hospitalized in the Paris area for surgery (**Paris surgery case in Table 1**). An epidemiological investigation was requested by the sanitary authorities. Within the surgical staff, only two night shift nurses (nurse 1 and nurse 2) who had been in contact with the patient during her stay in the clinic were found to be HIV-positive. The results obtained, through phylogenetic analysis conducted using *pol* reverse transcriptase and *env* C2C4 region, comparing the patients’ strains with reference strains and a large set of local controls obtained from HIV-infected individuals living in the Paris area, allowed the authors to unambiguously exclude nurse 1 as the source of infection. Nurse 2 was consistently found to cluster monophyletically with the patients, therefore was found to be a potential source of infection of the patient.

Another case, involving a **UK obstetrician/gynecologist (Table 1)**, investigated the possibility of transmission from an HIV-1 infected health care worker to a female patient, who was identified in an epidemiological investigation that was conducted for 1206 patients, upon whom the infected doctor had performed surgical procedures in the preceding 10 years. Here, the set of controls included a female HIV-1 positive who was a member of the same ‘sex circle’ as the female patient; an established male to female transmission pair and an established transmission trio (one male to three females). Arnold*et al* [8] used *env* genomic sequences to reconstruct ML and NJ trees analyzing the hypothesis of transmission. None of the results supported the hypothesis that the health care worker had infected the patient [8,9].

In 1993 an HIV-positive patient with no additional risk factors was identified during an epidemiological investigation that was conducted on 413 patients that had undergone surgical procedures by an identified HIV-positive surgeon [10]. However, the patient had also been transfused with 2 blood units with possibly infected blood products (the donor of 1 of the units was later found to be HIV-positive). This study investigated the possible pathways of transmission of HIV-1 to the patient on the basis of nucleotide sequence data of virus derived from 3 infected persons: the patient, the surgeon and the blood donor. As in previously described investigations, three different phylogenetic inference methods (ML, NJ and MP) were used. 10 HIV-1 strains, mainly isolated in the US, were used as controls. All methods consistently grouped the patient’s sequences with those from the blood donor (**blood transfusion case**, Table 1), while the surgeon’s sequences were more closely related to other published isolates. This analysis exonerated the surgeon, suggesting that it was more likely that the blood donor was the source of the HIV-1 infection in the patient [10].

In a potential **patient to surgeon infection**,Blanchard [11] investigated the hypothesis of HIV transmission from a patient to an orthopedic surgeon. The surgeon was found to be HIV positive in 1994 and reported percutaneous injury during surgery on a multi-transfused patient back in 1983. After the diagnosis, an epidemiological investigation was performed on the 3004 persons that he had operated in the 10-year interval. One of the patients was found to be HIV positive. The results obtained in the phylogenetic analysis indicated that the virus sequence from the surgeon and this patient formed a monophyletic cluster. The authors concluded that there was a high likelihood that nosocomial transmission had occurred. However, no local control sequences were added to the analyses: only reference sequences representative of the different HIV-1 subtypes. Therefore, the investigation reported does not allow to make statements about the likelihood of nosocomial transmission.

## HIV-1 Outbreaks in Prisons

Yirrel *et al.* [12] used molecular techniques to identify possible sources of an HIV outbreak resulting from needle sharing in the **Glenochil prison** (Central Scotland) . This epidemiological investigation was precipitated by the reporting of eight symptomatic cases of acute hepatitis B and 2 diagnosed cases of primary infection with HIV in the prison between April and June 1993. The prison doctor also reported widespread drug injecting and needle sharing, and the possibility of an outbreak was recognized. 227 (60%) of the inmates agreed to counseling and subsequently to an HIV test, in which 12 other cases of HIV infection were found. Phylogenetic analysis showed that 13 of the 14 HIV-positive prisoners had been infected from a common source, since they were infected with an almost identical virus. These results strongly suggested, not only that these men were infected while in Glenochil prison, but also that the infection came from a single source. Infection was probably transmitted by a single source directly or indirectly along a chain to either 12 (one of the 13 being the source) or 13 (original source not sampled) inmates over a short period of time. The 14th person was clearly infected with a different viral strain.

To the best of our knowledge, this was the only phylogenetic court investigation in an HIV-1 outbreak in prison published so far.

## HIV-1 Transmission in the Context of Sexual Abuse

Several court cases were conducted related to infection with HIV after rape, the first one was a **Swedish Rape case.** In Stockholm, Sweden, a male intravenous drug user (IDU) was convicted of rape and deliberately transmitting HIV-1 to a female victim. Albert *et al* [13] became involved during the appeal process. HIV-1 *gag* and *pol* genes from the suspect and the victim were sequenced and compared with sequences from 21 local controls. These controls were from the same geographic area and included both individuals from the same risk group as the suspect (IDUs), as well as from homosexual men. The analysis of the HIV-1 *gag* and *pol* gene sequences revealed that the case sequences were significantly more closely related to each other than to controls, and that the patients shared two distinct genetic variants. As such, it was concluded that it was very likely that HIV strains carried by the male and the female were epidemiologically closely linked. Herein, Albert *et al* used three distinct phylogenetic approaches for tree reconstruction: maximum-likelihood (ML), maximum parsimony (MP) and neighbor-joining (NJ). The results of this study together with other facts contributed to the guilt verdict being upheld during the appeal. This case set the stage for the use of phylogenetic analysis of HIV as forensic evidence [14].

In **Germany**, another **sexual abuse** casereported by Banaschack *et al* [15] investigated potential sexual abuse, of a 10-year old girl by her stepfather and consequent infection with HIV-1. Phylogenetic analysis of the *gag* and *env* genes of the suspect and the victim, together with 5 control sequences, indicated that their strains were very similar and clustered together in the phylogenetic tree. However, no information was provided on how control sequences were selected.

In **Denmark**, a court case was conducted after 10 boys revealed that their recreation center **coach** (“the man”), who trained them in wrestling, had abused them sexually between 1994 and 1998 [16]. After examination of all boys, one 12-year-old boy was found to be HIV-positive. The man was therefore charged for having performed intercourse with a child under the age of 15 and, since it was assumed that the man was aware of his own HIV-positive status, was also charged for recklessly having caused danger to another person’s life or health. During the court case, genetic analysis was performed of the HIV-1 strains found in the man and the child to determine whether viral sequences from the child were significantly more closely related to sequences from the man than to sequences from relevant controls. Local controls were collected from 14 unrelated HIV-1 infected individuals from the same geographic area and representing patients from different transmission risk groups. Three regions of the genome were analysed (*gag*, *pol* and *env*) and the phylogenetic analyses was performed as in Albert *et al* [16]. Here, the fact that 4 years had passed since the estimated time of transmission complicated the analyses. Analysis of the viral divergence between the man and the child was not comparable to what was found in previous transmission investigations. However, other investigations on related transmission cases have shown that the divergence differs from case to case and from gene to gene. Yet, when using phylogenetic analysis, it was found that, in all examined genomic regions, the viral sequences from the man and the child formed a separate cluster when compared to the local controls; and a unique amino acid deletion was found in the *env* gene, indicating a close epidemiological link between the virus from the man and the child. The man was sentenced to 6 years of imprisonment in March 2000 at the Copenhagen High Court.

In 2007, 14 MSM complainants from **The Netherlands** filed a criminal case against three other MSM defendants, accusing them of administering sedative drugs, sexual abuse, and deliberate subcutaneous injections with HIV-1-infected blood without consent, between the end of 2005 and the beginning of 2007 [17]. The complainants knew the three defendants, who organized both private meetings and **sex parties** at their home. The allegations were that at these parties, HIV-1 was transmitted from defendants to complainants, through the act of the defendants that injected their own blood subcutaneously into the complainants on several occasions. However, transmission could also have happened sexually. During the investigation, twelve of the 14 complainants and all defendants were found to be HIV-positive. Furthermore, phylogenetic analyses of *pol* and *env* genes was performed using NJ and Bayesian methods, together with reference sequences representative of the Dutch HIV epidemic, generated from blood plasma of subtype B-infected outpatients visiting the Academic Medical Center in Amsterdam, The Netherlands, between 2005 and 2009. The analysis identified the same dual infection in seven individuals (six complainants and one defendant), with two distinct HIV-1 subtype B strains, whereas three other complainants were infected with strain 1 only and two others with strain 2 only. Another complainant harbored a triple HIV-1 infection, two strains similar as in the other dually infected, and an unrelated subtype B strain. After trial and retrial, one was sentenced to 8 years in prison, another to 5 years and the third to 18 months.

## Sexual Transmission in the Context of Alleged HIV Non-Disclosure

Birch*et al* [18] reported a case of potential transmission of HIV in **Australia** from a HIV-positive man to several male and female contacts. A criminal case was conducted against this man for knowingly and **recklessly transmitting HIV**. Phylogenetic analysis was performed on genomic sequences from the p17 region of *gag* and a region bracketing the V3 region of *env*. The blood samples were obtained from the defendant, 3 complainants, the sex partner of 1 of these complainants, and a group of local HIV-positive controls, presumably epidemiologically unrelated to the suspect. The results indicated relatedness between HIV strains of the defendant and 2 of his male contacts, and the sex partner of 1 of these male contacts. However, it did not demonstrate closer relatedness to the viral strain from the third female contact, as compared to the used controls. Information obtained by contact tracing suggested that the defendant and the third female contact had been sex partners for many years and that, if transmission of HIV had taken place between them, it could have happened up to 9 years prior to the investigation. In contrast, the other two contacts had been infected for <2 years. Also it showed that one of the controls separating the defendant and the third female contact was in fact a contact of the defendant. It is therefore possible that the third female contact was indeed infected by the defendant, but that their strains had separated since the event due to further onward transmission. Time since infection was therefore identified as an important variable in such forensic analyses [18].

In 2005, Lemey *et al* [19] reported the analysis of a possible HIV-1 transmission case, in which six females presumably became HIV infected subsequent to a **sexual assault**. In connection with **professional activities** in Belgium, the defendant, who emigrated from Rwanda in 1992, was in contact with the complainants for several periods between the beginning of 1993 and the end of 1995. For five of the complainants, seroconversion could be dated between 1993 and 1995. At the time of the first sampling for the court investigation, the defendant was found to be HIV seropositive. An epidemiological investigation was conducted. Fragments from *pol* and *env* genes were sequenced from all patients involved in the case. Local controls were collected from two local hospitals. Local controls were selected based on the available epidemiological information from medical files of the defendant and the complainants. The primary epidemiological criteria included the same geographical area (preferentially from the same hospitals as attended by victims), diagnosed as HIV infected between 1992 and 1996, age- and risk group matching (heterosexual). Database controls were collected using a BLAST algorithm to retrieve the 10 most similar publicly available sequences to each defendant and complainants (excluding doubles). Phylogenetic analyses were conducted using different ML algorithms, Bayesian and Maximum Parsimony approaches. Phylogenetic analyses consistently inferred a monophyletic cluster for the defendant and the complainant sequences in both genome regions, distinct from local controls and database sequences. This led to the conclusion that the hypothesis that the defendant had transmitted the virus to the complainants could not be discarded. Together with other compelling testimonies, this forensic evidence led to the conviction of rape.

The **Glenochil** prison outbreak discussed above led to the first prosecution of HIV transmission in Scotland, data and samples from the prison outbreak were seized to build a case against a former inmate, that was found guilty of **recklessly transmitting HIV** to a female sexual partner [20], illustrating how public health data can be (mis)used in court.

## Breastfeeding

Another interesting case was reported in 2012 [21]. A 5-month old baby born from an HIV negative mother was found to be HIV positive after being admitted to the hospital for bronchopneumonia related respiratory distress. At first, nosocomial infection was suspected, but upon investigation by the hospital, the mother reported that her HIV-positive sister (aunt of the baby) had intermittently breast fed her baby since she was 6 weeks old. This led to an investigation of the aunt and of her own baby, who was also found to be HIV positive. Phylogenetic analysis was performed using the HIV-1 genomic sequences from the two babies and the aunt, together with 100 HIV-1 reference strains from the Free State Province, South Africa, closely related to the three case sequences. Phylogenies were subsequently estimated using Bayesian and ML methods. This, together with the results of the contact tracing, supported the hypothesis of transmission of HIV-1 between the aunt and the baby, in the face of a scenario of **shared breastfeeding** between the mother and the aunt of the baby. This investigation was presented as an emergency call to the common practice of breastfeeding by non-biological caregivers as the most important factor associated with HIV infection in discordant mother–child pairs in South Africa.

## Cases where Timing of Transmission was Investigated

A court case in **Libya** regarding the involvement of **foreign medical staff** in infecting hundreds of children received a lot of media attention [22]. In May 1998, the Al-Fateh Children’s Hospital (AFH) in Benghazi noted the first case of HIV-1 infection in a child with no known risk factor. In September, 111 other children were identified as HIV-positive. In total, 418 children were infected in the outbreak and at least 75 were co-infected with HCV. In March 1998, six foreign health workers (five Bulgarian nurses and a Palestinian doctor) had joined the medical staff. One year later, they were accused of purposefully infecting more than 400 children with HIV-1 and detained in prison. Although there was a European scientific report from 2003 that concluded that nosocomial infection was more likely, caused by pre-existing poor hygiene practices, and that suggested that the outbreak was already present in the hospital prior to the arrival of the foreign health workers, the Libyan court considered these findings imprecise and decided not to consider the report. Afterwards, a second scientific report produced by Libyan researchers concluded that the outbreak was not caused by nosocomial infection and that the data available did not contradict the hypothesis of a deliberate transmission of HIV to the infected children. As a consequence, the foreign medical staff was condemned to death. An international appeal was launched, which led to a retrial that was started in 2006. *Lawyer*s *without borders*, who were representing the defendants, appealed to international AIDS experts for a new inquiry, which was answered by Oliveira *et al* in a paper published in Nature just one week before the verdict was due [22]. Phylogenetic analyses were performed for 44 and 61 children for which HIV-1 and HCV sequences, respectively, were available. Using Bayesian and ML approaches, the authors found that the HIV-1 patients’ sequences formed one monophyletic cluster, while HCV sequences formed three separate monophyletic clusters. Furthermore, dating of the origin of the Most Recent Common Ancestor (MRCA) of these HIV an HCV clusters using Markov chain Monte Carlo (MCMC) analysis, indicated that the probability that this MRCA was originated at or after 1 March 1998 (the date of arrival of the foreign medical staff) was almost zero, the clusters originated from a few year to a decade earlier [22], indicating that the HIV and HCV strains were already circulating and prevalent in the hospital years before the arrival of the foreign medical staff that were accused of the transmissions. Thus, the results are consistent with the hypothesis of nosocomial transmission, and indicated a longstanding infection control problem in that hospital. These results were used to defend the accused, however, the death sentence was upheld. The paper by de Oliveira *et al* [22], together with many please was able to convince the Libyan government to allow the prisoners to serve their sentence in Europe, where they were promptly pardoned and released.

Matsuda[23] reported a case of a recently diagnosed 22-year-old female with no risk factor for HIV-1 acquisition. The patient had evidence of advanced disease at diagnosis and laboratory data was also compatible with long-standing infection. Contact tracing was performed to try to identify a source for the patient’s infection. After a thorough questionnaire, the only potential risk identified was **sharing of manicure instruments** with an HIV-1 infected female cousin, who was known to be infected for 17 years and had been lost to follow-up. Phylogenetic analysis of partial HIV-1 *pol* and *env* sequences from the infected patient and from the older cousin, revealed highly related sequences, with a putative common ancestry date coincident with the purported period of time when the manicure instruments had been shared.

Although not an HIV case, another interesting criminal case in which timing of transmission was investigated involved an outbreak of HCV, detected among patients having minor surgery in two **Valencian** hospitals [24–26]. The epidemiologists examined the conditions under which each of 66,000 surgeries had happened. The only significant common factor among those infected with HCV was an **anesthetist** who worked in these two hospitals, who was found to be a morphine addict. No additional links were found between these patients, and the anesthetist was charged with having infected the patients through needle sharing when administering part of the morphine dose to himself before giving the remainder of the dose to the patients. Public health authorities and the judge in charge of the corresponding epidemiological and judicial investigations requested expertise in evolutionary biology to discriminate who out of the potential victims had actually been infected by a common source, to provide an individual assessment of the likelihood that the anesthetist was the source, and to obtain an estimate of the date of infection for each patient and how this related to date of surgery. As the defendant (the anesthetist) was infected with HCV genotype 1a, only patients infected with this genotype were further included in the analysis. Samples from 322 HCV-1a infected patients who had been in contact with the defendant were sequenced, together with 44 samples from HCV-1a infected patients, from the city of Valencia, unrelated to the defendant (local controls). For this purpose, multiple HCV clones from each individual were sequenced and NJ and ML phylogenetic trees for thousands of these sequences were obtained. All trees shared a highly supported clade, which included the sequences from the defendant. All patients with sequences in this clade were considered as belonging to the outbreak: these included sequences from 274 patients and sequences from the defendant. All the sequences derived from the local controls and sequences from 47 patients, initially considered to belong to the outbreak, clustered outside this “outbreak clade”. Then, the Bayesian method with constant population size and relaxed molecular clock implemented in BEAST were used to estimate the time of the MRCA (tMRCA) that joined each of the 274 patients to the outbreak clade. The infection dates of 24 patients in the outbreak, known from laboratory seroconversion information, were used to calibrate the clock and the surgery dates were compared with the Bayesian Highest Posterior Density (HPD). In 176 cases (65%) the HPD interval for the Bayesian estimate of infection, although very wide, included the most likely estimated date given by the prosecution to the court, based on documents and testimonies from several sources. The results helped the court to convict the anesthetist of professional malpractice leading to the infection of 274 of his patients [24–26].

## Cases Where Transmission Direction was Investigated

In 2002, Metzker *et al* [27] claimed for the first time that the finding of paraphyly for the defendant’s sequences was consistent with the hypothesis of transmission from the defendant to the complainant. This was a criminal case that resulted in the conviction of a **Louisiana gastroenterologist** on the charge of attempted second-degree murder. The prosecution’s case was based on evidence indicating that the gastroenterologist made a mixture of blood-products from two patients under the doctor’s care, one infected with HIV-1 and the other with hepatitis C, and intentionally infected his former girlfriend, a nurse, during an argument by intramuscular injection. The complainant was tested after the incident and was found to be negative for HIV-1, but in January 1995, she tested positive and subsequently accused the defendant of infecting her with HIV-1. In the epidemiological investigation, the complainant reported having sexual contacts with seven men, including the defendant, who all tested negative for HIV. This was the first case, which made use of phylogenetic analysis in a criminal court case in the United States. Bayesian, MP and minimum evolution-based phylogenetic analyses of HIV-1 sequences isolated from the victim, the patient, and a local population sample of HIV-1 positive individuals showed that sequences from the complainant were more closely related to the gastroenterologist patient’s HIV-1 sequences.

In 2010, Scaduto et al [28] reported two HIV-1 transmission court cases, so called the Washington case and the Texas case. In both cases, the authors showed that direction of transmission can be established based on paraphyly. During transmission of HIV, a genetic bottleneck occurs, resulting in the paraphyly of source viruses with respect to those of the recipient. This paraphyly, when detected, establishes the direction of transmission, from which the source can then be inferred in the context of contact tracing. In the **Washington case**, the defendant was charged with intending to inflict “great bodily harm” by administering, exposing or transmitting HIV to 17 female partners through unprotected sexual relations. In the **Texas case**, the defendant was charged with intentionally, knowingly and recklessly causing “serious bodily injury” by exposing six female partners to HIV through unprotected sexual contact. Based on phylogenetic analysis and blinded sampling, an index case was inferred for each study based on the observed paraphyly, and after unblinding, this index case was the defendant. In both cases, contact tracing excluded other intermediate contacts, indicating that this paraphyly could correspond to direct transmission.

In a more recent case, in 2013 in China, a woman accused the man with whom she lived with, that he had not disclosed his HIV-positive status to her, and had deliberately infecting her with HIV [29]. In August 2013, samples were collected from both individuals, as well as from 18 local controls, which were selected from individuals with similar epidemiological characteristics, including the same geographical area and risk group (heterosexual contact), and diagnosis around the time of the alleged transmission event (between January and March 2013), while not necessarily of the same subtype as the cases. Within all three analysed gene regions, the mean genetic distances of the woman’s virus to the man’s were significantly smaller than those of the controls’ compared to the man’s. Neighbor-joining phylogenetic analyses of all three genes showed that the women’s sequences clustered together with the man’s sequence with high bootstrap support. Furthermore, bootscanning analyses showed that *gag* sequences of the woman and the man had the same recombination pattern between subtype B and subtype C, which was different from those of CRF07\_BC and CRF08\_BC. Finally, the phylogenetic analyses of the quasispecies showed that the man’s sequences were paraphyletic to the woman’s sequences and the woman’s *env* sequences showed genetic distances lower than the man’s, while they were significantly closer to the man’s sequences than to the group of control sequences. The authors then concluded that, in the context of epidemiological, serological and phylogenetic evidence, there was support for the transmission from the man to the woman.

# Experimental techniques, untested in court

In a recent paper, Romero-Severson *et al* [30] used modeling to study the phylogenetic relationships between HIV populations and associated phylogenetic tree topologies. Using both simulations and real known transmission cases, they argued that the direction of transmission can often be established when paraphyly exists, although this paraphyletic relationship is often lost with time passed after the transmission event. Furthermore, they argue that intermediary links can be excluded when multiple lineages were transmitted given that a certain type of phylogenetic relationship (combination of paraphyletic and polyphyletic) occurs in cases of direct transmission. Finally, they find that, when the sampled individuals’ HIV populations both are monophyletic, a common source was likely the origin.

Next generation sequencing, both genome wide and deep sequencing, is also starting to be used to improve the estimates of direction, and the estimate of a time window of infection. According to recently unpublished results by Puller *et al*, (<http://biorxiv.org/content/early/2017/04/21/129387>), virus diversity increases linearly with time since infection. The precision of this approach to estimate time of infection depends on the genomic region and codon position, with highest precision when the 3rd codon positions and the entire pol gene are used and a precision that is better than most alternative biomarkers. Using this approach, time since infection can be estimated many years after infection, since viral diversity increases approximately linearly during at least 8 years after infection, which allows estimation of time since infection during this time period

# References

1 Ou CY, Ciesielski CA, Myers G, Bandea CI, Luo CC, Korber BT, *et al.* Molecular epidemiology of HIV transmission in a dental practice. *Science* 1992; **256**:1165–1171.

2 Smith TF, Waterman MS. The continuing case of the Florida dentist. *Science* 1992; **256**:1155–1156.

3 DeBry RW, Abele LG, Weiss SH, Hill MD, Bouzas M, Lorenzo E, *et al.* Dental HIV transmission? *Nature* 1993; **361**:691.

4 Crandall KA. Intraspecific phylogenetics: support for dental transmission of human immunodeficiency virus. *J Virol* 1995; **69**:2351–2356.

5 Leitner T, Fitch W. The Phylogenetics of known transmission histories. In: *The Evolution of HIV.* The Johns Hopkins University Press; 1999.

6 Jaffe HW, McCurdy JM, Kalish ML, Liberti T, Metellus G, Bowman BH, *et al.* Lack of HIV transmission in the practice of a dentist with AIDS. *Ann Intern Med* 1994; **121**:855–859.

7 Goujon CP, Schneider VM, Grofti J, Montigny J, Jeantils V, Astagneau P, *et al.* Phylogenetic analyses indicate an atypical nurse-to-patient transmission of human immunodeficiency virus type 1. *J Virol* 2000; **74**:2525–2532.

8 Arnold C, Balfe P, Clewley JP. Sequence distances between env genes of HIV-1 from individuals infected from the same source: implications for the investigation of possible transmission events. *Virology* 1995; **211**:198–203.

9 Hochuli V, Hyndman O, Porter K. Response to news that an obstetrician/gynaecologist has AIDS. *Commun Dis Rep CDR Rev* 1995; **5**:R7-11.

10 Holmes EC, Zhang LQ, Simmonds P, Rogers AS, Brown AJ. Molecular investigation of human immunodeficiency virus (HIV) infection in a patient of an HIV-infected surgeon. *J Infect Dis* 1993; **167**:1411–1414.

11 Blanchard A, Ferris S, Chamaret S, Guétard D, Montagnier L. Molecular evidence for nosocomial transmission of human immunodeficiency virus from a surgeon to one of his patients. *J Virol* 1998; **72**:4537–4540.

12 Yirrell DL, Robertson P, Goldberg DJ, McMenamin J, Cameron S, Leigh Brown AJ. Molecular investigation into outbreak of HIV in a Scottish prison. *BMJ* 1997; **314**:1446–1450.

13 Albert J, Wahlberg J, Leitner T, Escanilla D, Uhlén M. Analysis of a rape case by direct sequencing of the human immunodeficiency virus type 1 pol and gag genes. *J Virol* 1994; **68**:5918–5924.

14 Albert J, Wahlberg J, Uhlén M. Forensic evidence by DNA sequencing. *Nature* 1993; **361**:595–596.

15 Banaschak S, Werwein M, Brinkmann B, Hauber I. Human immunodeficiency virus type 1 infection after sexual abuse: value of nucleic acid sequence analysis in identifying the offender. *Clin Infect Dis Off Publ Infect Dis Soc Am* 2000; **31**:1098–1100.

16 Machuca R, Jørgensen LB, Theilade P, Nielsen C. Molecular investigation of transmission of human immunodeficiency virus type 1 in a criminal case. *Clin Diagn Lab Immunol* 2001; **8**:884–890.

17 van der Kuyl AC, Jurriaans S, Back NKT, Sprenger HG, van der Werf TS, Zorgdrager F, *et al.* Unusual cluster of HIV type 1 dual infections in Groningen, The Netherlands. *AIDS Res Hum Retroviruses* 2011; **27**:429–433.

18 Birch CJ, McCaw RF, Bulach DM, Revill PA, Carter JT, Tomnay J, *et al.* Molecular analysis of human immunodeficiency virus strains associated with a case of criminal transmission of the virus. *J Infect Dis* 2000; **182**:941–944.

19 Lemey P, Van Dooren S, Van Laethem K, Schrooten Y, Derdelinckx I, Goubau P, *et al.* Molecular testing of multiple HIV-1 transmissions in a criminal case. *AIDS Lond Engl* 2005; **19**:1649–1658.

20 Bird SM, Brown AJ. Criminalisation of HIV transmission: implications for public health in Scotland. *BMJ* 2001; **323**:1174–1177.

21 Goedhals D, Rossouw I, Hallbauer U, Mamabolo M, de Oliveira T. The tainted milk of human kindness. *Lancet Lond Engl* 2012; **380**:702.

22 de Oliveira T, Pybus OG, Rambaut A, Salemi M, Cassol S, Ciccozzi M, *et al.* Molecular epidemiology: HIV-1 and HCV sequences from Libyan outbreak. *Nature* 2006; **444**:836–837.

23 Matsuda EM, Coelho LPO, Pimentel VF, Onias HB, Brigido LF de M. An HIV-1 transmission case possibly associated with manicure care. *AIDS Res Hum Retroviruses* 2014; **30**:1150–1153.

24 González-Candelas F, Bracho MA, Wróbel B, Moya A. Molecular evolution in court: analysis of a large hepatitis C virus outbreak from an evolving source. *BMC Biol* 2013; **11**:76.

25 Echevarría JM, León P, Domingo CJ, López JA, Elola C, Madurga M, *et al.* Laboratory diagnosis and molecular epidemiology of an outbreak of hepatitis C virus infection among recipients of human intravenous immunoglobulin in Spain. *Transfusion (Paris)* 1996; **36**:725–730.

26 Vandamme A-M, Pybus OG. Viral phylogeny in court: the unusual case of the Valencian anesthetist. *BMC Biol* 2013; **11**:83.

27 Metzker ML, Mindell DP, Liu X-M, Ptak RG, Gibbs RA, Hillis DM. Molecular evidence of HIV-1 transmission in a criminal case. *Proc Natl Acad Sci U S A* 2002; **99**:14292–14297.

28 Scaduto DI, Brown JM, Haaland WC, Zwickl DJ, Hillis DM, Metzker ML. Source identification in two criminal cases using phylogenetic analysis of HIV-1 DNA sequences. *Proc Natl Acad Sci U S A* 2010; **107**:21242–21247.

29 Chen M, Ma Y, Yang C, Yang L, Chen H, Dong L, *et al.* The combination of phylogenetic analysis with epidemiological and serological data to track HIV-1 transmission in a sexual transmission case. *PloS One* 2015; **10**:e0119989.

30 Romero-Severson EO, Bulla I, Leitner T. Phylogenetically resolving epidemiologic linkage. *Proc Natl Acad Sci U S A* 2016; **113**:2690–2695.