

FULL PAPERS /LAVORI ORIGINALI

Ruolo del sequenziamento diretto nella tipizzazione di HCV-RNA a fini clinico-epidemiologici

Silvia Gabella, Tiziano Allice, Silvia Varetto, Fabrizia Pittaluga, Valeria Ghisetti.

Laboratorio di Microbiologia, Ospedale Molinette, Torino.

Key words: HCV sequencing, 5'UTR, NS5B

HCV-RNA typing by direct sequencing

SUMMARY

Background/Aims: Genotyping and subtyping of hepatitis C virus (HCV) is epidemiologically and clinically relevant to prognosis and therapeutical management of HCV infection. Aim was to study the feasibility of an "in house" direct sequencing method for the genotyping and subtyping of HCV strains compared with a commercially available genotyping assay (Inno-LiPA, VERSANT® HCV Genotype 2.0 Assay, Bayer).

Methods: 74 clinical plasma samples cross-representing HCV genotypes 1 to 5 typed with the line probe assay Inno-LiPA (LiPA) were subjected to a laboratory-developed 5'UTR direct sequencing protocol (5'UTR Seq). A NS5B direct sequencing protocol (NS5B Seq) was apply to 23/74 samples (17 negative by the 5'UTR). Two libraries of 5'UTR and NS5B HCV prototypes were constructed; BioEdit 7.0.0 and ClustalW were used for alignments and phylogenetic analysis.

Results: 5'UTR Seq typed 47/74 LiPA positive samples (64%). Concordance between the two assays was 91% (43/47) and 72% (31/43) for HCV genotyping and subtyping, respectively. LiPA and 5'UTR Seq discordant samples were typed by NS5B Seq. NS5B Seq typed 13/17 samples negative with 5'UTR Seq. All genotypes and subtypes detected with NS5B Seq were concordant with LiPA. 5'UTR and NS5B direct sequencing pointed out a wider HCV subtype distribution than LiPA for genotype 1, 2 and 4. By the combination of 5'UTR and NS5B direct sequencing, 81% (60/74) of LiPA positive samples were typed.

Conclusion: HCV typing and subtyping with the line probe assay Inno-LiPA is highly recommended for clinical purposes. A more detailed HCV typing for epidemiologic purposes can be achieved by the direct sequencing of a region with a moderate degree of genetic variability such as NS5B. HCV NS5B analysis is more efficient in resolving viral strain genetic variability than direct sequencing of a highly conserved region such as 5'UTR.

INTRODUZIONE

Il virus dell'epatite C (HCV) è un virus ad RNA a singola elica con polarità positiva che rappresenta una delle principali cause di epatite cronica e carcinoma epatico nel mondo: circa il 3% della popolazione mondiale (circa 180 milioni) è infettata da HCV (2, 4).

HCV è presente in un singolo individuo infetto, sotto forma di popolazioni virali differenti, che costituiscono la "quasispecies" virale (3). HCV mostra una notevole variabilità genetica determinata dall'alto tasso di mutazioni (stimato $1,5-2 \times 10^{-3}$ sostituzioni per sito genomico/anno) risultante dalla combinazione di una elevata replicazione virale (circa 10^{10} e 10^{12} virioni al giorno) e dalla mancanza di attività "proof-reading" (correzione di errori di replicazione) della RNA-poli-merasi RNA-dipendente del virus. Il genoma di HCV è formato da regioni conservate che presentano una variabilità non superiore il 10% e da regioni variabili che presentano differenze anche superiori al 50% (3, 5, 9, 10). La regione genomica maggiormente conservata è la 5' non codifi-

cante (5'UTR), con una omologia di sequenza pari al 92-98% tra diversi isolati. La variabilità di HCV è maggiore nelle regioni codificanti le proteine dell'envelope, in particolare nel segmento amino-terminale di 27 aminoacidi di E2|NS1 (HVR1). In base all'analisi di tali regioni è stato possibile classificare il virus HCV in 6 genotipi e più di 50 sottotipi (9, 10).

La determinazione del genotipo e del sottotipo del virus dell'epatite C ha importanti implicazioni dal punto di vista epidemiologico e clinico (1, 5, 13). La terapia con interferone-ribavirina è molto più efficace verso HCV di genotipo 2 e 3 rispetto ai virus appartenenti al genotipo 1 a cui è associata una minore risposta alla terapia (13). Dal punto di vista epidemiologico, la caratterizzazione genetica del virus è importante per identificare catene epidemiologiche di trasmissione nosocomiale e sorgenti di infezione. Infatti, il rischio di acquisizione nosocomiale di HCV è ancora presente, in particolare in unità ad alto rischio come le emodialisi, nonostante le misure di precauzione universalmente adottate in tali centri per prevenire e

ridurre la trasmissione del virus, in considerazione anche dell'elevato numero di pazienti con infezione da HCV che sono ricoverati in tali unità. Le analisi molecolari e filogenetiche sono pertanto indispensabili per individuare le vie e le fonti di trasmissione nosocomiale di HCV.

In questo lavoro, viene descritto un metodo di sequenziamento diretto sviluppato "home-made" e di analisi filogenetica del virus dell'epatite C, basato sullo studio di due regioni virali, la regione altamente conservata 5'UTR e la regione non strutturale NS5B a moderata variabilità genetica, per la definizione dei genotipi e dei sottotipi dei ceppi isolati. Il metodo è stato valutato in rapporto al sistema di genotyping commerciale Inno-LiPA (VERSANT® HCV Genotype 2.0 Assay, Bayer).

MATERIALI E METODI

Campioni

Campioni di plasma afferenti al Laboratorio di Microbiologia dell'Ospedale Molinette di Torino per la ricerca del genotipo di HCV, nel periodo giugno 2004-luglio 2006, sono stati analizzati utilizzando il saggio Inno-LiPA HCV mediante ibridazione inversa (LiPA, VERSANT HCV Genotype 2.0 Assay, Bayer). Per ogni campione è stata determinata la carica virale di HCV utilizzando la tecnica del branched DNA (bDNA, VERSANT HCV RNA 3.0 Assay, Bayer).

Sono stati selezionati 74 campioni rappresentanti i 5 genotipi principali di HCV e sono stati sottoposti a sequenziamento diretto della regione 5'UTR. Il sequenziamento diretto della regione NS5B è stato valutato in 23 campioni, comprendenti i campioni (n=4) per cui il saggio Inno-LiPA HCV e il sequenziamento 5'UTR davano risultati discordanti e i casi (n=17) in cui il sequenziamento 5'UTR è risultato negativo.

Tipizzazione mediante sequenziamento diretto della regione 5'UTR (5'UTR Seq) e della regione NS5B (NS5B Seq)

Estrazione dell'RNA e retrotrascrizione

L'RNA è stato estratto a partire da 140 µl di plasma utilizzando il kit di estrazione Viral RNA Mini Kit (Qiagen, Mi, I), eluito in 50 µl di acqua ribonucleasi-free e conservato a -80°C. Venticinque µl di RNA sono stati retrotrascritti con il Kit High Capacity (Applied Biosystem, Foster City, CA, USA) a 48°C per 60 minuti e utilizzando random primers in un volume finale di 50 µl.

Amplificazione di 5'UTR e NS5B mediante PCR

Una porzione di 259 nucleotidi della regione 5'UTR è stata amplificata utilizzando i primers P15 - senso (5'-GGGGGCGACACTCCACCA-

3') e P274-antiseno (5'-TCGCGACCCAACACTACTC-3') (7), mentre un frammento di 400 nucleotidi della regione NS5B del virus HCV, è stato amplificato con i primers NS5B-senso (5'-T G G G G A T C C C G T A T G A - TACCCGCTGCTTTGA-3') e NS5B-antiseno (5'-GGCGGAATTCCTGGTCATAGCCTCCGTGAA-3') (12).

In entrambi i protocolli i cicli di amplificazione erano 40 ed è stata utilizzata una temperatura di annealing di 62°C. I prodotti di PCR sono stati analizzati in un gel di agaroso al 2%.

La sensibilità della PCR per 5'UTR è stata valutata amplificando diluizioni note di un plasmide contenente il cDNA della regione 5'UTR mentre la PCR per la regione NS5B amplificando 10 diluizioni seriali di un siero contenente HCV RNA a quantità nota (1350 UI/µl, corrispondenti a 3500 copie/µl).

Sequenziamento diretto e analisi filogenetica

I frammenti della regione 5'UTR e NS5B ottenuti mediante le rispettive reazioni di PCR, sono stati purificati utilizzando il sistema QIAquick PCR purification kit (Qiagen, Mi, I) e sequenziati direttamente con il sequenziatore automatico ABI Prism 310 Genetic Analyzer instrument (Applied Biosystem, Foster City, Ca, USA).

Le sequenze sono state visualizzate utilizzando il software DNA Sequencing Analysis™ Version 3.7 ed allineate con ClustalW (bootstrap method: 1000 replicates).

Le analisi filogenetiche sono state valutate con il software BioEdit Versione 7.0.0 e sono state costruite due librerie di riferimento costituite da sequenze comprendenti i genotipi ed i sottotipi del virus dell'epatite C selezionate dal database GeneBank (<http://www.ncbi.nlm.nih.gov/Genbank/index.html>) delle regioni 5'UTR (n=71) e NS5B (n=30).

Il genotipo ed il sottotipo di ciascuna sequenza ottenuta è stato attribuito in base all'analisi filogenetica e visualizzato con il software TreeView 1.6.6. Tutte le sequenze ed in particolare quelle il cui sottotipo è risultato dubbio mediante una prima analisi filogenetica, sono state analizzate con il programma di allineamento Genotyping (<http://www.ncbi.nlm.nih.gov/projects/genotyping/formpage.cgi>), utilizzando le stesse sequenze di riferimento HCV (n=71 e n=30) selezionate per l'analisi filogenetica delle regioni 5'UTR ed NS5B.

RISULTATI

Tipizzazione di HCV a livello di genotipo e sottotipo mediante Inno-LiPA (LiPA) e sequenziamento della regione 5'UTR: confronto tra le due metodiche

La distribuzione dei genotipi e sottotipi dei 74 campioni HCV analizzati con il test LiPA era la seguente: genotipo 1, n=2 (3% di tutti i campioni tipizzati), 1a, n=7 (9.5%), 1b, n=17 (23%), 1a/1b, n=3 (4%), 2, n=1 (1.4%), 2a/2c, n=16 (22%), 2b, n=1 (1.4%), 3a, n=15 (20%), 4, n=10 (13.5%), 4e, n=1 (1.4%) e 5a, n=1 (1.4%). Tredici campioni (18%) sono stati genotipizzati ma non sottotipizzati, mentre per 16 campioni non è stato possibile discriminare tra il sottotipo 2a e 2c e per 3 campioni tra il sottotipo 1a e 1b.

Il protocollo di sequenziamento diretto per l'analisi delle regioni 5'UTR e NS5B prevede due amplificazioni single step specifiche e indipendenti di HCV con una sensibilità di 50 copie per la regione 5'UTR e di 75 copie per la regione NS5B.

Il sequenziamento è stato possibile solamente quando il prodotto di amplificazione era superiore a 2500 copie di genoma virale.

Utilizzando il sequenziamento 5'UTR sono stati tipizzati 47/74 (64%) campioni identificati con LiPA (Tabella 1). Ventisette campioni positivi con il test LiPA non sono risultati sequenziabili poiché il prodotto di amplificazione era in quantità insufficiente. I 47 campioni tipizzati con 5'UTR Seq sono stati classificati in 5 genotipi e 9 sottotipi: 1a, n=7 (15% dei campioni tipizzati); 1b, n=13 (28%); 1c, n=3 (6.4%); 2a, n=3 (6.4%); 3a, n=13 (28%); 4a, n=4 (8.5%); 4d, n=1 (2.1%); 4m, n=1 (2.1%) e 5a, n=1 (2.1%). Di un solo campione appartenente al genotipo 4 non è stato possibile determinare il sottotipo.

Il confronto tra le due tecniche di tipizzazione LiPA e 5'UTR Seq è stato quindi effettuato considerando i campioni risultati positivi in entrambi i test (n=47). La concordanza a livello di genotipo virale tra le due metodiche era del 91% (43/47). I campioni discordanti erano 2 classificati come genotipo 2a/2c con LiPA e risultati come genotipo 1, sottotipo a e 3, sottotipo a, nell'analisi della sequenza 5'UTR, e 2 di genotipo 4 attribuiti al genotipo 1, sottotipo a, mediante sequenziamento. Con il saggio LiPA è stato possibile definire il sottotipo di 31 su 47 (66%) campioni mentre utilizzando il sequenziamento della regione 5'UTR sono stati sottotipizzati 46 su 47 (98%) campioni. La concordanza, a livello di sottotipo virale, ha evidenziato che un sottotipo 1a determinato con il test LiPA è stato classificato come 1c mediante 5'UTR Seq, un genotipo 1a/1b come 1c, un genotipo 1 come sottotipo 1b e quattro genotipi 4 come 4a, ed infine un sottotipo 4e come 4m (Tabella 2).

Analisi dei campioni risultati discordanti con i test LiPA e 5'UTR Seq e determinazione del genotipo con il sequenziamento diretto della

regione NS5B dei campioni negativi con 5'UTR Seq

Il secondo protocollo di sequenziamento diretto, sviluppato nel nostro laboratorio per l'analisi della regione NS5B (NS5B Seq) di HCV, è stato utilizzato per valutare 4 campioni discordanti con il test LiPA e 5'UTR Seq, e 17 campioni LiPA positivi non sequenziabili con il test 5'UTR. I genotipi ottenuti con il sequenziamento della regione NS5B sono risultati concordanti con quelli determinati mediante il test LiPA.

In 13 su 17 campioni negativi mediante 5'UTR Seq è stato possibile definire il genotipo che è risultato sempre concordante con quello ottenuto con LiPA (Tabella 3).

Confrontando i sottotipi, un solo isolato classificato come 1a con il test LiPA è risultato 1b con NS5B Seq. Inoltre l'analisi della regione NS5B ha permesso di determinare il sottotipo di 9 isolati non tipizzabili con la tecnica LiPA, in particolare per i sottotipi 1a (n=2 campioni), 2c (n=6) e 4a (n=1) (Tabella 3).

Le Figure 1 (5'UTR) e 2 (NS5B) riportano gli alberi filogenetici ottenuti confrontando i campioni sequenziati (indicati con S) con le sequenze di isolati virali per le regioni 5'UTR (n=71) e NS5B (n=30) utilizzati come riferimento. I due alberi filogenetici permettono una buona discriminazione dei vari genotipi e dei sottotipi. Per poter definire in modo più preciso le analisi filogenetiche effettuate sui vari isolati, ciascuna sequenza (5'UTR e NS5B, comprese quelle dei ceppi di riferimento) è stata analizzata con il programma Genotyping.

I genotipi e sottotipi ottenuti sono risultati concordanti con quelli ottenuti mediante analisi filogenetiche delle sequenze 5'UTR e NS5B fatta eccezione per un isolato (S28) indicato come genotipo 4 dalle analisi con i test 5'UTR Seq e LiPA e come genotipo 1 con il programma Genotyping. Per 9 isolati per i quali una prima analisi di sequenza non ha reso possibile stabilire il sottotipo in base alla topologia dell'albero filogenetico creato sulle sequenze 5'UTR, in particolare (Figura 1) per due genotipi 1 (S41 e S74), tre genotipi 2 (S19, S21 e S45) e quattro genotipi 4 (S16, S17, S24 e S48), in base all'analisi con il software Genotyping, l'attribuzione è stata rispettivamente al sottotipo 1c (n=2), 2a (n=3) e 4a (n=4).

Tabella 1. Campioni di HCV sottoposti al test Inno-LiPA e sequenziati sulla regione 5'UTR.

Inno-LiPA HCV	5'UTR Seq	
Campioni analizzati	Campioni positivi	Campioni negativi
74	47 (64%)	27 (36%)

Tabella 2. Genotipi e sottotipi ottenuti mediante il test commerciale Inno-LiPA HCV ed il sequenziamento della regione 5'UTR.

Inno-LiPA HCV		5'UTR Seq	
Genotipo	N° isolati		Genotipo
1a	4	→	1a
1a	1	→	1c
1b	12	→	1b
1	1	→	1b
1	1	→	1c
1a/1b	1	→	1c
2a/2c	3	→	2a
2a/2c	1	→	1a
2a/2c	1	→	3a
3a	12	→	3a
4	1	→	4
4	4	→	4a
4	1	→	4d
4	2	→	1a
4e	1	→	4m
5a	1	→	5a
Totale = 47		Totale = 47	

Le righe in grigio scuro indicano i genotipi discordanti
 Le righe in grigio chiaro indicano i sottotipi discordanti

Tabella 3. Confronto dei genotipi e sottotipi ottenuti mediante il test commerciale Inno-LiPA e il sequenziamento della regione NS5B di campioni risultati non sequenziabili sulla regione 5'UTR.

Inno-LiPA		NS5B Seq	
Genotipo	N° isolati		Genotipo
1a	1	→	1b
1°, 1a/1b	3	→	1a
2	1	→	2c
2b	1	→	2b
2a/2c	5	→	2c
3a	1	→	3a
4	1	→	4a
Totale = 13		Totale = 13	

La riga in grigio indica i sottotipi discordanti

CONCLUSIONI

Il sistema di tipizzazione di HCV basato sul sequenziamento diretto della regione 5'UTR sviluppato nel nostro laboratorio ha permesso di giungere ad una attribuzione di genotipo nel 64% dei campioni considerati. L'analisi della sequenza di 5'UTR di HCV non è risultata sempre efficiente probabilmente perché basata su una reazione di PCR con un solo step di amplificazione che genera una quantità di prodotto non sempre sufficiente come input per la successiva reazione di sequenza. Infatti, il sequenziamento è stato possibile solamente quando il prodotto di amplificazione superava le 2500 copie di genoma virale. Per rendere più efficiente il protocollo potrebbe essere studiata e utilizzata una seconda amplificazione interna alla regione 5'UTR, trasformando la reazione di amplificazione in una nested-PCR. Inoltre, la capacità di discriminare a livello di sottotipo in questa regione (soprattutto nel genotipo

1 e 2) è nota per essere meno efficiente che per altre regioni genomiche di HCV (es. NS5B) e questo comporta la necessità di uno studio particolare del disegno dei primers per migliorare l'efficienza di discriminazione a livello di genotipo e sottotipo virale.

I campioni che non sono stati sequenziabili, ad un successivo esame mediante sequenziamento della regione NS5B, sono stati risolti a livello di genotipo (ed sottotipo) nella maggior parte dei casi (60 su 74 isolati, 81%). Il sequenziamento sulla regione NS5B ha permesso di caratterizzare il 76% dei campioni risultati non sequenziabili sulla regione 5'UTR fornendo risultati completamente concordanti con quelli ottenuti con il test LiPA. L'analisi delle sequenze NS5B ha permesso di determinare più correttamente i sottotipi di HCV, in particolare all'interno del genotipo 2 (tutti gli isolati 2a/2c con il test LiPA, sono stati identificati come 2c con il test NS5B Seq, mentre con il test 5'UTR Seq 3 su 5 isolati sono risultati sottotipo 2a).

In conclusione, il metodo LiPA per l'identificazione dei genotipi di HCV da 1 a 6 e relativi sottotipi virali, è un ottimo test da utilizzare nella diagnostica clinica: è infatti in grado di genotipizzare correttamente pressoché tutti i campioni HCV considerati. L'attribuzione a livello di sottotipo è meno efficiente, ma la rilevanza clinica di ciò non è ancora stata stabilita. Quindi, può essere considerato come il metodo di scelta da adottare in laboratori di Microbiologia clinica per la genotipizzazione di HCV.

L'analisi di sequenza di HCV, rispetto al test LiPA, a fronte di una complessa ottimizzazione e standardizzazione, presenta ovviamente una migliore capacità di determinare i sottotipi in particolare per i genotipi 1, 2 e 4 ed è pertanto un metodo molto importante per lo studio epidemiologico degli isolati, dove è necessaria una informazione molto dettagliata su di essi. I nostri risultati sono in accordo con quanto riportato da altri autori (6, 8, 11) che considerano l'analisi della regione 5'UTR meno efficiente di altre per via dell'alto grado di conservazione tra i diversi genotipi di HCV, in specie del genotipo 1 e 2. Il metodo di tipizzazione basato sull'analisi diretta della sequenza della regione virale NS5B, regione a moderata variabilità genetica, presenta un maggiore potere di discriminazione e risulta pertanto idoneo per studi di tipo epidemiologico, e può pertanto essere applicato allo studio della diffusione nosocomiale di HCV.

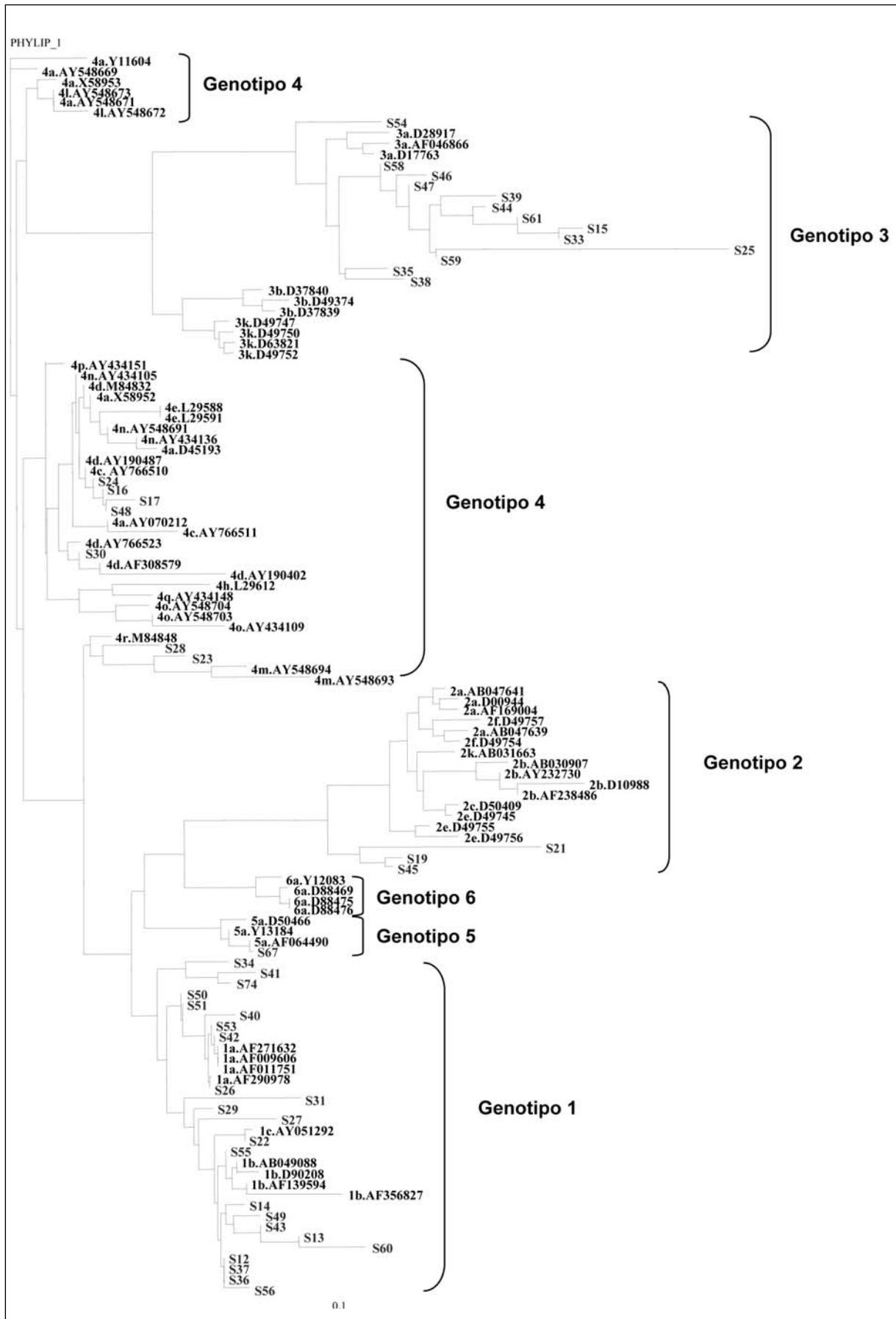


Figura I. Analisi filogenetica di 47 isolati e 71 sequenze di riferimento della regione 5'UTR del virus dell'epatite C (neighborjoining method).

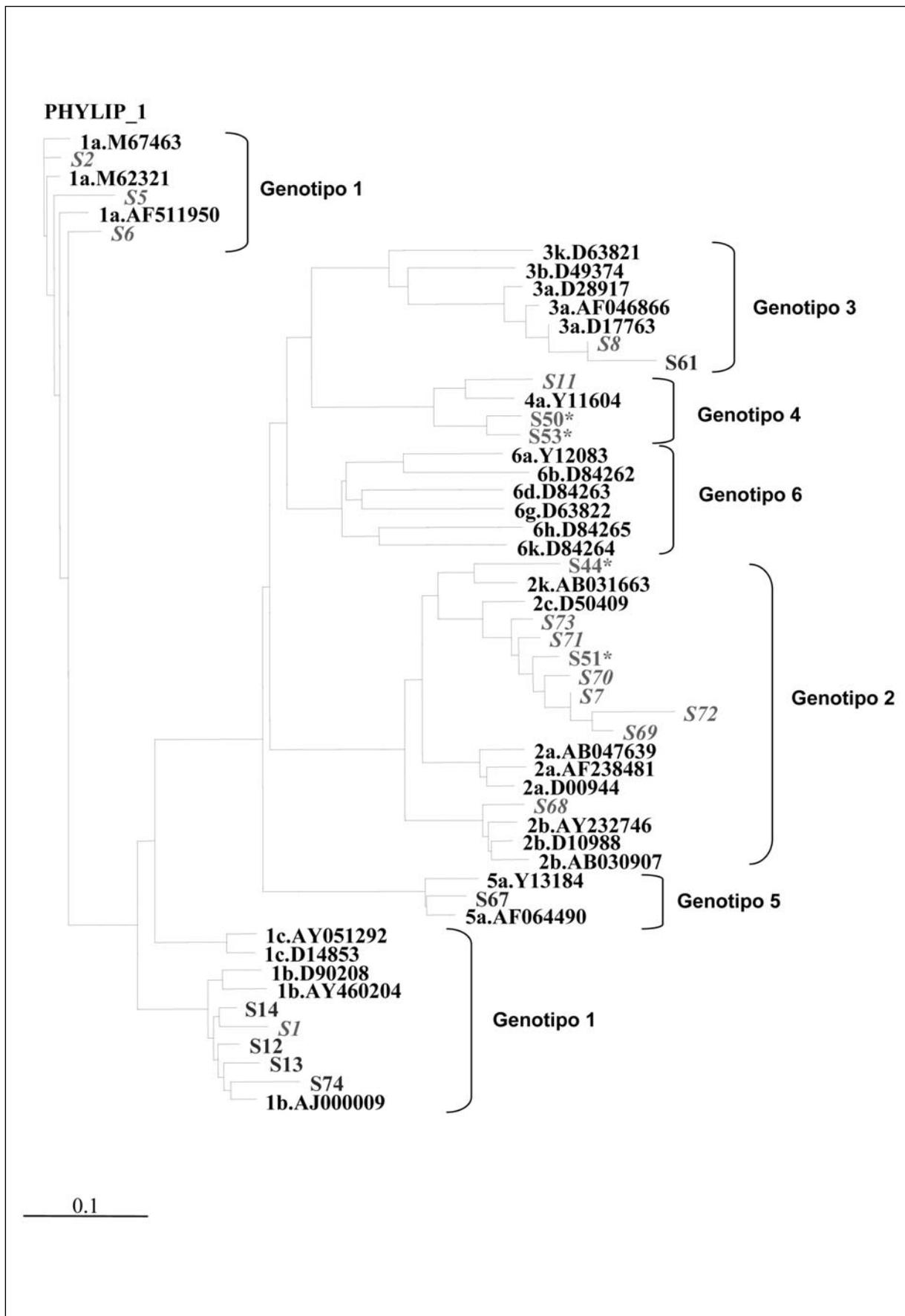


Figura I. Analisi filogenetica di 23 campioni e 30 sequenze di riferimento della regione NS5B del virus dell'epatite C (neighbor-joining method). In corsivo 13 isolati non sequenziati con 5'UTR Seq e con l'asterisco (*) 4 isolati discordanti per il test LiPA e quello 5'UTR Seq.

BIBLIOGRAFIA

1. Alberti A, Chemello L, Noventa F, Cavalletto L, and Salvo GD. Therapy of hepatitis C: re-treatment with alpha interferon. *Hepatology*. 1997; 26: 137S-142S.
2. Alter MJ. Epidemiology of hepatitis C. *Hepatology*. 1997; 26(Suppl 1): 62S-65S.
3. Bukh J, Miller RH, and Purcell RH. Genetic heterogeneity of hepatitis C virus: quasispecies and genotypes. *Semin. Liver Dis.* 1995; 15:41-63.
4. Choo QL, Kuo AJ, Weiner L, Overby R, Bradley DW, and Houghton M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989; 244:359-362.
5. Hmaied F, Ben Mamou M, Saune-Sandres K, et al. Hepatitis C virus infection among dialysis patients in Tunisia: incidence and molecular evidence for nosocomial transmission. *J Med Virol.* 2006; 78: 185-91.
6. Le Pogam S, Dubois F, Christen R, Raby C, Cavicchini A, Goudeau A. Comparison of DNA enzyme immunoassay and line probe assays (Inno-LiPA HCV I and II) for hepatitis C virus genotyping. *J Clin Microbiol.* 1998; 36: 1461-3.
7. Negro F, Krawczynski K, Quadri R, et al. Detection of Genomic - and Minus-Strand of Hepatitis C Virus RNA in the Liver of Chronic Hepatitis C Patients by Strand-Specific Semiquantitative Reverse-Transcriptase Polymerase Chain Reaction. *Hepatology.* 1999; 29: 536-542.
8. Sandres-Saune K, Deny P, Pasquier C, Thibaut V, Duverlie G, Izopet J. Determining hepatitis C genotype by analyzing the sequence of the NS5b region. *J Virol Methods.* 2003; 109: 187-93.
9. Simmonds P. Variability of hepatitis C virus. *Hepatology.* 1995; 21: 570-583.
10. Simmonds P, Bukh J, Combet C, et al. Consensus Proposals for a Unified System of Nomenclature of Hepatitis C Virus Genotypes. *Hepatology.* 2005; 42 : 962-973.
11. Stuyver L, Wyseur A, van Arnhem W, et al. Hepatitis C virus genotyping by means of 5'-UR/core line probe assays and molecular analysis of untypeable samples. *Virus Res.* 1995; 38: 137-57.
12. Tamalet C, Colson P, Tissot-Dupont H, et al. Genomic and phylogenetic analysis of hepatitis C virus isolates: a survey of 535 strains circulating in southern France. *J Med Virol.* 2003; 71: 391-398.
13. Zein NN. Clinical significance of hepatitis C virus genotypes. *Clin Microbiol Rev.* 2000; 13: 223-225.

Tiziano Alice
Laboratorio di Microbiologia
Ospedale Molinette
Corso Bramante 88/90 - 10126 Torino
E-mail: tizall@yahoo.it