

UNIVERSITA' DEGLI STUDI DI MILANO

Scuola di Dottorato in Scienze Biomediche Cliniche e Sperimentali

Dipartimento di Scienze Biomediche e Cliniche "Luigi Sacco"

Corso di Dottorato in

Malattie Infettive XXVIII Ciclo



FILODINAMICA E FILOGEOGRAFIA DEI PRINCIPALI SOTTOTIPI DI HCV NEI BALCANI OCCIDENTALI

Dottoranda
Chiara **SORRENTINO**

Matricola: R09994

TUTOR:

Prof. Gianguglielmo **ZEHENDER**

COORDINATORE DEL DOTTORATO:

Prof.ssa Antonella **D'ARMINIO MONFORTE**

Anno accademico: 2014-2015

INDICE

| | |
|---|-----------|
| 1.INTRODUZIONE..... | 4 |
| 1.1 STORIA DELL'EPATITE C | 4 |
| 1.2 IL VIRUS DELL'EPATITE C | 5 |
| 1.2.1 Classificazione e struttura del virus | 5 |
| 1.2.2 Il genoma del Virus | 6 |
| 1.2.3 Le Proteine strutturali..... | 8 |
| 1.2.4 Le proteine Non Strutturali | 10 |
| 1.2.5 Il ciclo vitale del virus..... | 12 |
| 1.3 LA VARIABILITA' VIRALE | 15 |
| 1.3.1 Genotipi e sottotipi di HCV..... | 17 |
| 1.3.2 Distribuzione geografica dei genotipi di HCV..... | 19 |
| 1.4 EPIDEMIOLOGIA DEL VIRUS HCV | 22 |
| 1.4.1 Modalità di trasmissione | 24 |
| 1.4.1.1 Trasmissione parenterale..... | 25 |
| 1.4.1.2 Altre sorgenti di trasmissione..... | 26 |
| 1.4.2 Genotipi e fattori di rischio associati..... | 26 |
| 1.5 INFEZIONE DI HCV IN ITALIA..... | 28 |
| 1.6 INFEZIONE DI HCV NEI BALCANI OCCIDENTALI..... | 29 |
| 1.7 STORIA NATURALE DELL'INFEZIONE DI HCV | 30 |
| 1.7.1 Epatite C acuta..... | 31 |
| 1.7.2 Epatite C cronica | 32 |
| 1.7.3 Cirrosi ed epatocarcinoma..... | 33 |
| 1.8 LA TERAPIA FARMACOLOGICA | 34 |
| 2. SCOPO DELLO STUDIO | 38 |
| 3. MATERIALI E METODI..... | 40 |
| 3.1 LA CASISTICA MONTENEGRINA | 40 |
| 3.2 LA CASISTICA ALBANESE | 41 |
| 3.3 METODI..... | 42 |
| 3.3.1 Modalità di raccolta, preparazione e conservazione dei campioni..... | 42 |
| 3.3.2 Estrazione di RNA virale da plasma | 42 |
| 3.3.3 Retrotrascrizione del gene NS5B | 43 |
| 3.3.4 Amplificazione della regione NS5B mediante NESTED-PCR..... | 44 |
| 3.3.5 Rilevamento dei prodotti di PCR..... | 46 |
| 3.3.6 Purificazione dei prodotti di PCR..... | 47 |
| 3.3.7 Sequenziamento della regione genomica amplificata | 48 |
| 3.3.7.1 La reazione di sequenza | 48 |
| 3.3.7.2 Purificazione dei prodotti di sequenza | 49 |
| 3.3.7.3 Sequenziamento della regione NS5B..... | 50 |
| 3.3.7.4 Analisi dei dati di sequenza..... | 51 |
| 3.3.8 Analisi filogenetica..... | 51 |
| 3.3.8.1 Analisi filogenetica delle sequenze..... | 52 |
| 3.3.8.2 Preparazione dei data set per l'analisi filogenetica..... | 53 |
| 3.3.9 Analisi filodinamica | 54 |
| 3.3.10 Analisi Filogeografica..... | 55 |
| 3.3.11 Analisi Statistica..... | 56 |
| 4- RISULTATI..... | 57 |
| 4.1 ANALISI FILOGENETICA DELLE SEQUENZE MONTENEGRINE | 57 |
| 4.1.1 Correlazione tra sottotipi di HCV e fattori di rischio nei soggetti montenegrini..... | 58 |

| | |
|--|-----------|
| 4.1.2 Caratterizzazione dei sottotipi di HCV per sesso e per età nei soggetti montenegrini | 60 |
| 4.2 ANALISI FILOGENETICA DELLE SEQUENZE ALBANESI | 62 |
| 4.2.1 Correlazione tra sottotipi di HCV e fattori di rischio nei soggetti albanesi | 64 |
| 4.2.2 Caratterizzazione dei sottotipi di HCV per sesso e per età nei soggetti albanesi | 65 |
| 4.3 STORIA EPIDEMIOLOGICA DEL SOTTOTIPO 1B..... | 67 |
| 4.3.1 Analisi filogeografica del sottotipo 1b | 67 |
| 4.3.2 Analisi filodinamica del sottotipo 1b..... | 70 |
| 4.4 STORIA EPIDEMIOLOGICA DEL SOTTOTIPO 3A | 71 |
| 4.4.1 Analisi filogeografica del sottotipo 3a | 71 |
| 4.4.2 Analisi filodinamica degli IVDUs 3a | 75 |
| 4.5 STORIA EPIDEMIOLOGICA DEL SOTTOTIPO 4D | 78 |
| 4.5.1 Analisi filogeografica del sottotipo 4d | 78 |
| 4.5.2 Analisi filodinamica del sottotipo 4d..... | 80 |
| 5. DISCUSSIONE | 81 |
| 6.BIBLIOGRAFIA..... | 87 |

1.INTRODUZIONE

1.1 STORIA DELL'EPATITE C

A partire dal 1970, con lo sviluppo di indagini sierologiche per l'identificazione dell'epatite A e dell'epatite B, venne identificato un ulteriore agente, trasmesso principalmente per via parenterale, responsabile di un vasto numero di epatiti post-trasfusionali che non erano dovute ai due virus già noti; vennero pertanto genericamente definite con il termine di epatiti “non-A e non-B” (NANB)¹.

Studi epidemiologici indicavano che le epatiti NANB rappresentavano, a livello mondiale, il 15-50% delle epatiti acute dell'adulto, definite come epatiti “acquisite in comunità”².

A causa del decorso spesso asintomatico dell'epatite NANB non era chiaro se si trattasse di un'infezione virale o se fosse una malattia epatica aspecifica contratta in seguito ad ospedalizzazione³. La possibilità dell'esistenza di un terzo tipo di epatite coinvolse un ampio numero di ricercatori nello sforzo di isolare il nuovo agente eziologico.

In questo scenario, un aiuto notevole arrivò dalla sperimentazione sugli scimpanzé, che si rivelarono essenziali per dimostrare che l'epatite NANB era causata da un agente trasmissibile^{4,5}, filtrabile⁶ e dotato di *envelope* lipidico⁷. Inoltre, gli scimpanzé offrirono anche un sistema *in vivo* per l'amplificazione del virus, essenziale per ottenere sufficiente materiale per i tentativi di identificazione.

Un altro contributo chiave per la conoscenza delle epatiti NANB furono le informazioni raccolte in studi prospettici su epatiti associate a trasfusioni. Questi studi misero in luce che l'epatite NANB costituiva la più rilevante complicazione post-trasfusionale e permisero anche di descriverne la storia naturale⁸.

La scoperta di questo agente richiese circa 15 anni dal momento del suo riconoscimento come entità clinica e venne realizzata da Micheal Houghton e collaboratori presso i laboratori della Chiron. Grazie all'avvento delle tecnologie del DNA ricombinante, fu possibile clonare il genoma del virus che fu chiamato **virus dell'epatite C (HCV)**. Il rapido clonaggio del genoma virale portò alla messa a punto di test sierologici per la determinazione degli anticorpi di HCV⁹, la cui applicazione nella pratica clinica ha

permesso di evidenziare in HCV il principale agente delle epatiti NANB post-trasfusionali¹⁰ e di quelle acquisite in comunità.

Fu il 21 Aprile del 1989 che Michael Houghton della *Chiron Corporation* di Emeryville e Daniel Bradley del *Centre of Disease Control* di Atlanta, annunciarono, in una celebre pubblicazione su *Science*, la scoperta del virus dell'epatite C (HCV)¹¹.

1.2 IL VIRUS DELL'EPATITE C

1.2.1 Classificazione e struttura del virus

HCV è stato classificato nella famiglia delle *Flaviviridae* nel genere *Hepacivirus*¹². La famiglia delle *Flaviviridae*¹³ comprende anche il genere dei *Flavivirus*, a cui appartengono virus trasmessi prevalentemente da artropodi come il virus della febbre gialla, il virus della Dengue, il virus dell'encefalite giapponese, il virus dell'encefalite trasmessa da zecche; il genere dei *Pestivirus* che comprende alcuni virus veterinari tra cui il virus della diarrea bovina; mentre il genere *Hepacivirus*, oltre ad HCV, include il recente virus *GBV-B* isolato nei tamarin mentre gli altri GB-like virus sono inclusi nel genere *Pegivirus* (GBV-A, GBV-C e GB-D)¹⁴ (Figura 1).

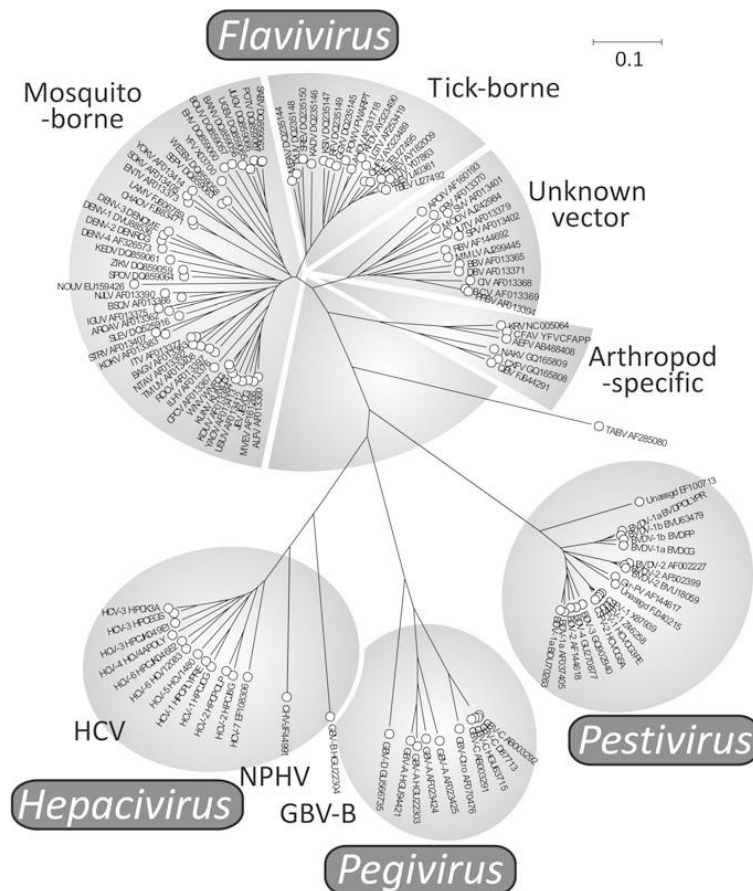


Figura 1: Classificazione della famiglia *Flaviviridae* sulla base del confronto del gene della Polimerasi (Simmonds 2013)

I membri della famiglia delle Flaviviridae, e in particolare HCV, sono tutti dotati di un *envelope* derivante dalla membrana della cellula ospite in cui sono inserite le glicoproteine virali (E1 e E2) che circondano un nucleocapside in cui è contenuto il genoma virale. Il genoma dei Flaviviridae è costituito da un singolo filamento di RNA a polarità positiva approssimativamente di 9.500-12.300 nucleotidi¹⁵ contenente un'unica *open reading frame* codificante una poliproteina di circa 3.000 aminoacidi. (Figura 2)

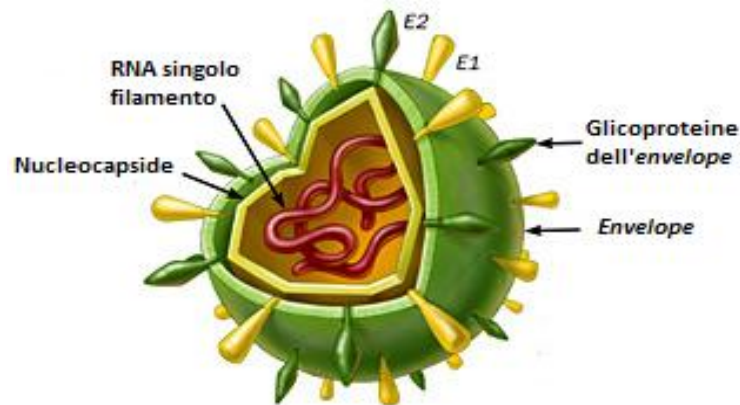


Figura 2: Morfologia del virus dell'epatite C.

1.2.2 Il genoma del Virus

I progressi compiuti dalla biologia molecolare su HCV sono stati così rapidi e massicci da permettere l'ottenimento dell'intera sequenza nucleotidica di HCV in un anno¹⁶. L'isolato originale (HCV-1) è costituito da una molecola di RNA a singola elica (ss-RNA) a polarità positiva di circa 9.600 ribonucleotidi contenuto in un capsido proteico circondato da un rivestimento lipidico^{17,18}. Quasi l'intero genoma è occupato da una lunga sequenza codificante, *open reading frame (ORF)*, che codifica per un'unica poliproteina di 3.011 aminoacidi (Figura 3). L'ORF dei diversi genotipi di HCV ha una lunghezza differente compresa tra 9.024-9.111 nucleotidi¹⁹.

Nel genoma virale si possono distinguere tre diverse regioni:

- una regione al 5' non tradotta (**5'UTR**);
- una regione centrale, che codifica per una singola poliproteina di circa 3.000 aminoacidi;
- una regione al 3' non tradotta (**3'UTR**).

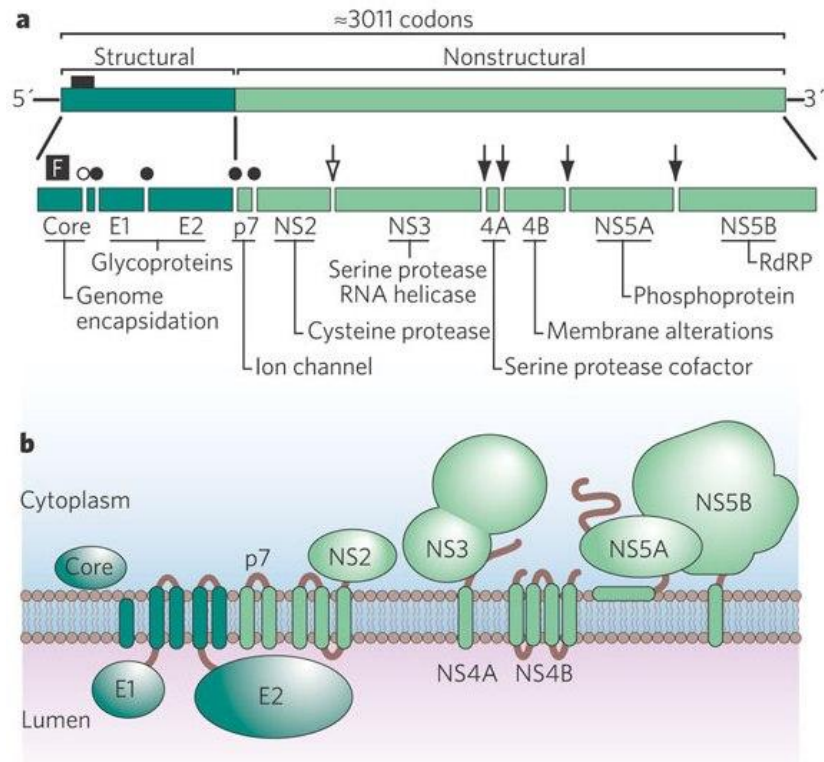


Figura 3: Struttura genomica del virus HCV. Parte a: La struttura del genoma. Parte b. La topologia delle proteine di HCV rispetto alla membrana intracellulare.

La **regione 5'UTR**, è lunga 341 nucleotidi ed è la regione più conservata di tutto il genoma; infatti è stata utilizzata per sviluppare test diagnostici per HCV-RNA. Il 5'UTR contiene 4 domini altamente strutturati nominati da I a IV, che contengono *stem-loop* e pseudo nodi. I domini II, III e IV insieme ai primi nucleotidi (12-30 nucleotidi) della regione del *core* costituiscono il sito di entrata dei ribosomi (*internal ribosome entry site, IRES*) che è essenziale per la sintesi delle proteine virali. IRES ha la capacità di formare un complesso di pre-replicazione stabile senza il bisogno dei fattori di inizio di traduzione classici²⁰. Probabilmente diverse proteine cellulari interagiscono con IRES aiutandola ad assumere la sua configurazione funzionale e regolandone l'attività. Alcuni studi riportano che l'efficienza traduzionale della regione 5'UTR varia nei diversi genotipi di HCV¹⁵.

La **regione 3'UTR** contiene circa 225 nucleotidi ed è divisa in tre regioni: una regione variabile di 30-40 nucleotidi, un lungo tratto di poly-U e una regione altamente conservata di 98 nucleotidi (*3'X region*) che include tre *stem-loop*, indispensabile per la replicazione di HCV *in vivo* e per la stabilizzazione dell'RNA di HCV. Qualunque mutazione in questa regione fa sì che non avvenga la replicazione virale²¹.

L'alto grado di conservazione della regione la rende un buon *target* per test diagnostici e terapie contro l'infezione di HCV¹⁵.

La **regione centrale** codifica per un'unica poliproteina che viene processata a livello co- e post-traduzionale da proteasi cellulari e virali a produrre 11 differenti prodotti proteici che si raggruppano in:

- **Proteine strutturali:** localizzate all’N-terminale, comprendono le proteine del *core* (C) e le proteine dell’*envelope* (E1 e E2),
- **Proteine non strutturali NS:** localizzate nella porzione rimanente della poliproteina, comprendono le proteine P7, NS2, NS3, NS4A, NS4B, NS5A e NS5B. (Figura 3)
- **Proteina F/ARF^a:** derivante da un *frameshift* nella regione del core.

La maturazione delle proteine strutturali del *core* e dell’*envelope* è mediato da peptidasi della cellula ospite mentre il processamento della regione non strutturale avviene tramite due proteasi virali.

1.2.3 Le Proteine strutturali

Il primo prodotto dell’ORF è la **proteina basica del core** (21 kDA).

I primi 191 aminoacidi costituiscono il principale componente del nucleocapside virale. La proteina del *core* modula diversi processi cellulari. E’ una proteina altamente conservata e immunogenica: infatti nella zona N-terminale contiene diversi epitopi per i linfociti B, usati comunemente per i test di *screening*. La proteina del *core* contiene alcuni domini distinti: un dominio N-terminale idrofilico (D1) e un dominio C-terminale idrofobico (D2). Il dominio D1 contiene una regione di localizzazione nucleare e di legame all’RNA ed è essenziale per l’assemblamento del nucleocapside. Il dominio D2 è responsabile dell’associazione della proteina del *core* con il reticolo endoplasmatico, la membrana mitocondriale e i *droplet* lipidici²².

Inoltre è stato dimostrato che la proteina del *core* è in grado di interagire con proteine cellulari e vari *pathway* importanti per il ciclo cellulare e può regolare attività di proto-oncogeni a livello trascrizionale e sopprimere l’apoptosi cellulare¹⁹. Infatti, è stato dimostrato da esperimenti condotti con geni *reporter*, può attivare diversi promotori cellulari, come ad esempio quelli dei geni *c-myc* e *c-fos*, e anche numerosi promotori virali (retrovirus e HBV)²³.

Queste funzioni del *core* potrebbero avere implicazioni negative su alcune funzioni cellulari e quindi contribuire alla patogenesi di alcune malattie, come ad esempio il carcinoma epatico²⁴.

Le **glicoproteine** (gp) **dell'*envelope***, E1 (33-35 KDa) e E2/NS1 (70-72 kDa) sono proteine transmembrana di tipo I altamente glicosilate, componenti essenziali dell'*envelope* virale e necessarie per i processi di fusione e entrata in cellula. Le proteine dell'*envelope* contengono domini idrofobici al C-terminale che possono essere importanti per l'ancoraggio alla membrana. La delezione di questa regione porta alla produzione di una forma solubile della proteina.

Le proteine E1 e E2 sono coinvolte nel legame con il recettore cellulare e nella fusione con la cellula ospite.

Grazie a sistemi di espressione ed esperimenti di immunoprecipitazione, è stato dimostrato che E1 e E2 interagiscono formando complessi eterodimerici stabili. La maggior parte di questi complessi eterodimerici sono formati da legami non-covalenti, mentre una parte minore di essi presenta ponti disolfuro¹⁵.

Studi sulla regione C-terminale hanno suggerito che questa possa essere la regione di legame tra E1 e E2. Si ipotizza, inoltre, che questi legami a livello del reticolo endoplasmatico determinino la gemmazione virale.

E1 ed E2 contengono delle sequenze di ritenzione nel reticolo endoplasmatico che ne limitano la localizzazione a questo distretto cellulare²⁵. L'espressione di E1 e E2 sulla superficie cellulare è molto limitata e questo può spiegare come le cellule infettate possano evadere l'attacco da parte del sistema immune.

Le regioni E1 e E2 presentano, inoltre, il più alto tasso di mutazione a livello nucleotidico e aminoacidico, tanto da suggerire che tali regioni siano sottoposte alla pressione selettiva del sistema immune dell'ospite.

Un grado particolarmente elevato di variazione aminoacidica è stata osservata nella regione N-terminale di E2, denominata HVR1 (regione ipervariabile1). Essa è localizzata tra gli aminoacidi 384-410 della poliproteina. Numerosi dati indicano che E2 forma parte dell'*envelope* virale e che HVR1 è presente sulla superficie virale e contiene epitopi immunogenici neutralizzanti. La regione HVR1 varia enormemente, sia all'interno di un singolo isolato, generando una forma di quasispecie virale, sia tra genotipi diversi¹⁵.

Per queste sue caratteristiche di elevata eterogeneità, HVR1 non può essere usata per la genotipizzazione virale. Studi della struttura secondaria mostrano che HVR-1 è astrutturato, in accordo con una sequenza che deve poter tollerare un alto grado di

variabilità e in contrasto con altre regioni della poliproteina che mostrano una precisa struttura secondaria e terziaria.

Una seconda regione ipervariabile collocata all'estremità 3' di HVR-1 e denominata HVR-2 è stata descritta solo per ceppi di genotipo 1b²⁶. Essa è lunga solo 7 aminoacidi e la sua funzione non è ancora nota.

La presenza di rapidi cambiamenti in questa regione permette il meccanismo mediante il quale HCV evade la sorveglianza immunitaria dell'ospite ed è in grado di stabilire e mantenere un'infezione persistente²⁷.

1.2.4 Le proteine Non Strutturali

P7 è un piccolo polipeptide di 63 aminoacidi e rappresenta una proteina integrale di membrana²⁸. Si compone di due domini transmembrana organizzati in α -eliche, collegate da un *loop* citoplasmatico. p7 sembra essere essenziale, perché mutazioni o delezioni nel *loop* citoplasmatico diminuiscono l'infettività della trasfezione intraepatica dell'HCV cDNA in scimpanzé²⁹. Studi in vitro hanno suggerito che p7 appartiene alla famiglia delle viroporine e potrebbe fungere da canali ionici del calcio³⁰. Tuttavia, questi risultati restano da confermare *in vivo*.

La maggior parte delle proteine NS sono richieste per la replicazione dell'RNA virale. Il taglio tra E2 e NS2, mediato da una peptidasi cellulare, avviene relativamente tardi nel processo di formazione di HCV. La maturazione delle proteine non strutturali è realizzata da 2 enzimi di origine virale che sono la proteasi NS2-3 e NS3-4A.

La proteina NS2 ed il dominio N-terminale di NS3 costituiscono la **proteasi NS2-3** che ha azione autoproteolitica in quanto catalizza il taglio a livello della giunzione NS2-3; si trova tra i residui 827-1207 della poliproteina. La proteasi richiede lo zinco e può funzionare in *trans* anche se inefficientemente¹⁵. NS2 perde la sua proprietà di proteasi dopo l'autoclivaggio da NS3 e viene degradata nel proteasoma.

NS3 è una proteina multifunzionale che nella porzione N-terminale di circa 189 aminoacidi codifica per una serin-proteasi (p70). NS4A è il cofattore dell'attività proteasica.

La **proteasi NS3-4A** è responsabile del taglio alle giunzioni NS3/4A, NS4A/B, NS4B/5A e NS5A/B. Per questa ragione, NS3 è indispensabile per la crescita e il processamento virale. L'attività serin-proteasica è stata ben caratterizzata e gioca un

ruolo cruciale nel processamento di HCV. Dunque la serin-proteasi rappresenta un'importante *target* nella terapia antivirale.

La regione N-terminale di NS3 interagisce con la proteina NS4A creando un complesso stabile tra le due proteine. Sebbene NS4A non sia essenziale per tutte le funzioni di NS3, è una parte integrante del complesso proteasico, stabilizza la proteina in un complesso attivo e permette il posizionamento in membrana dove ha luogo il processo di clivaggio¹⁵.

Nella parte C-terminale di NS3 invece risiede l'attività NTPasica/RNA elicastica, essenziale per la traduzione e replicazione del genoma di HCV. Probabilmente l'attività elicastica di NS3 è indispensabile per la separazione dei filamenti di RNA durante la replicazione virale, durante la trascrizione, per risolvere eventuali strutture secondarie e per facilitare l'entrata dei ribosomi durante la traduzione. L'attività NTPasica fornisce energia per separare gli acidi nucleici, idrolizzando nucleosidi trifosfati¹⁵.

La **regione NS4** codifica per due proteine denominate NS4A e NS4B.

NS4A è il cofattore essenziale di NS3.

NS4B è una proteina idrofobica di 27 kDa la cui funzione è rimasta sconosciuta per molto tempo. Una delle funzioni di NS4B è servire come membrana ancoraggio per il complesso di replicazione³¹. Recentemente è stato scoperto che NS4B è implicata nell'attività della RNA polimerasi virale³² e in vari *pathways* di trasduzione del segnale dell'ospite.

La **regione NS5** codifica per due proteine; NS5A e NS5B.

NS5A è una proteina altamente fosforilata da chinasi cellulari in molti residui di serina. Il ruolo di NS5A nella replicazione di HCV non è ancora ben definito, ma come già visto per altri virus a RNA, le fosfoproteine sono in genere regolatori della replicazione. NS5A si associa con *raft* lipidici derivati da membrane intracellulari. Questa interazione sembra essere essenziale per la formazione del complesso replicazione di HCV in connessione con *raft* lipidici. Inoltre, NS5A può interagire direttamente con NS5B, ma il meccanismo con cui NS5A modula l'attività RdRp non è stato chiarito³³.

NS5B codifica per RNA polimerasi RNA-dipendente (RdRp), in grado di legare e copiare diversi stampi a RNA. Proprio per la sua funzione può rappresentare il bersaglio principale per bloccare la replicazione virale¹⁵.

La proteina NS5B possiede un'attività di RNA polimerasi RNA-dipendente (RdRp)³⁴ richiesta per replicazione del genoma virale. L'enzima è in grado di copiare il genoma completo di HCV *in vitro* in assenza di ulteriori fattori virali o cellulari. La proteina

NS5B, come altri RdRp finora conosciuti, contiene 6 motivi conservati denominati A-F. Grazie ad una analisi mutazionale, sono stati identificati quattro motivi cruciali per l'attività RdRp (A, B, C, D). Sostituzioni di residui conservati all'interno di questi motivi ha ridotto notevolmente le attività enzimatiche della polimerasi; una singola sostituzione nel motivo D aumenta l'attività RdRp di circa il 50%. Alcuni studi indicano che l'eliminazione di residui amminoacidici ai terminali, in particolare l'N-terminale, sono importanti per l'attività RdRp ma non vincolanti per l'RNA³⁵.

La struttura tridimensionale mostra che la proteina NS5B possiede una tipica struttura "a mano destra" con siti catalitici presenti alla base del dominio a palmo, circondato da domini dita e pollice abbastanza distinguibili. Questi domini circondano il sito attivo e creano un canale per il legame con l'ssRNA stampo.

Tramite saggi di legame ed esperimenti di competizione è stato dimostrato che l'enzima purificato si lega preferibilmente a sequenze al 3' di HCV.

Infine, è stata trovata una attività *terminal nucleotidyl transferasi* (TNTasi) associata all'enzima purificato³⁴.

Proteina Frameshift: la proteina F o ARFP (proteina alternativa *reading frame*) viene generata come risultato di un *frameshift* ribosomiale -2 / + 1 nella regione codificante del core N-terminale della poliproteina HCV. Sono stati rilevati anticorpi anti-peptidi dalla proteina F in pazienti con infezione cronica, suggerendo che la proteina viene prodotta durante l'infezione. Tuttavia, l'esatto meccanismo di traduzione che disciplina la frequenza e la resa della proteina F durante le varie fasi di infezione da HCV sono completamente sconosciuti. Così, il ruolo della proteina F nella ciclo di HCV rimane enigmatico, ma si pensa sia coinvolta nel fenomeno della persistenza virale³⁶.

1.2.5 Il ciclo vitale del virus

HCV è in grado di replicare principalmente in cellule epatiche, ma la replicazione virale può avvenire anche in cellule mononucleate del sangue periferico (PBMC), se pur con minor efficienza.

La dinamica della replicazione di HCV può essere dedotta dall'elevata velocità di produzione del virus e dalla emergenza dei mutanti. Assumendo che circa il 10% degli epatociti del fegato sia infetto, che il fegato contenga circa 2×10^{11} epatociti e che vengono prodotte circa 10^{12} particelle virali al giorno, si deduce che ogni epatocita produce circa 50 particelle virali al giorno³⁷.

Il ciclo replicativo di HCV (Figura 4) può essere riassunto nei seguenti punti:

1. Attacco ed entrata nella cellula ospite: è la prima tappa del ciclo replicativo del virus in cui è necessaria un'interazione specifica tra un recettore sulla superficie cellulare e una proteina di superficie virale. La particella virale, inizialmente, si attacca ai glicosamminoglicani e, possibilmente, al recettore delle LDL (LDL-R). Dopo l'attacco, la particella utilizza 4 fattori di entrata: il CD81, membro della famiglia delle tetraspanine, proteine espresse sulla superficie di diverse cellule tra cui gli epatociti, SR-BI, recettore "scavenger" di classe B di tipo I che lega le lipoproteine e, più di recente, Claudin1 e Occludin, proteine che compongono le giunzioni *tight* tra cellule, per le fasi di post *binding*.

La glicoproteina che si pensa sia responsabile dell'inizio dell'attacco del virus alla cellula ospite è la glicoproteina dell'*envelope* E2. Il ruolo di E1 è meno chiaro, ma la presenza di un gruppo di aminoacidi idrofobici, chiamati peptidi di fusione, suggerirebbe che anche E1 sia coinvolta con E2 nella fusione cellulare³⁸. Il virione entra mediante un processo di endocitosi clatrina-dipendente. Dopo la fusione cellulare il genoma è rilasciato nel citosol³⁹.

2. Traduzione dell'RNA e processamento della poliproteina virale: L'ss-RNA virale è completamente equipaggiato per fungere da RNA messaggero, quindi una volta entrato nel citoplasma viene immediatamente tradotto. La traduzione dell'RNA virale è mediata da IRES. La poliproteina viene processata co- e post-traduzionalmente da proteasi cellulari e virali a livello del reticolo endoplasmico rugoso. Le proteine di HCV così prodotte, formano un complesso replicasico altamente stabile e ordinato associato alla membrana intracellulare.

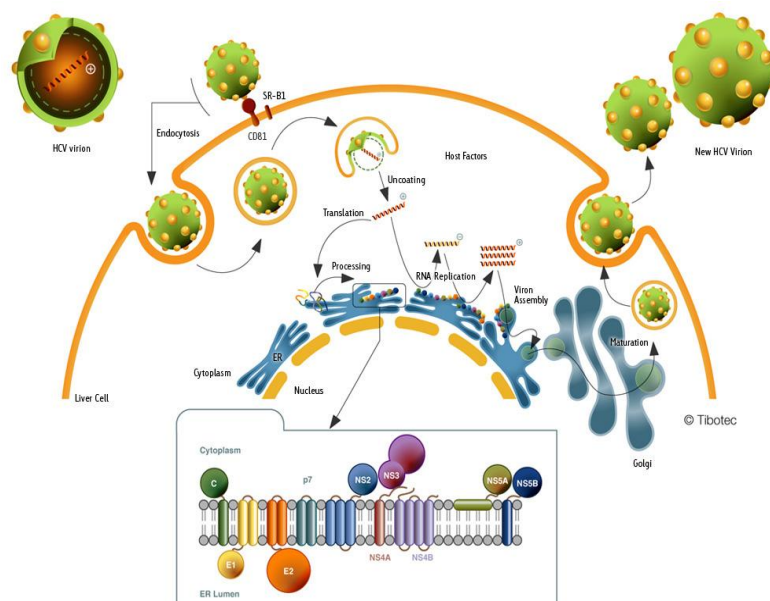


Figura 4: Ciclo replicativo del virus HCV.

3.Replicazione dell'RNA virale: Il complesso replicasico formatosi nella fase precedente contiene già tutte le proteine di HCV ed è una caratteristica tipica dei virus a RNA *plus-strand*. Questo processo, tipico dei *Poliovirus* e dei *Flavivirus*, consente la produzione delle proteine e dell'RNA virali in un singolo compartimento, avvicinando strettamente e concentrando proteine a diversa funzione.

I singoli passaggi di questa fase sono poco chiari. Probabilmente, una volta che le proteine virali hanno raggiunto un livello sufficiente, avviene uno *switch* e una parte dell'RNA passa dal processo di traduzione a quello di replicazione. Nella prima parte di questa fase, le proteine virali si legano all'RNA genomico e lo copiano in un RNA antigenomico o a polarità negativa. Non è chiaro se l'RNA genomico e antigenomico formino un intermedio a ds-RNA o se si appaiano tra loro solo con poche basi. In ogni caso l'antigenomico viene poi copiato a sua volta in RNA a polarità positiva. Il rapporto di equilibrio genomico-antigenomico è 10:1.

La progenie di molecole di RNA positive può subire tre diverse destinazioni:

- RNA genomico impaccato in particelle infettive rilasciate in circolo;
- mRNA per la produzione di proteine di HCV;
- RNA per iniziare nuovi cicli di replicazione.

È ovvio che l'enzima RdRp, prodotto da NS5B, è la chiave che catalizza la sintesi di RNA positivo e negativo. Una questione ancora non risolta è se NS5B legghi specificatamente solo HCV-RNA o RNA in modo aspecifico.

4. Assemblaggio e rilascio in circolo del virione: La formazione delle particelle dovrebbe iniziare con l'associazione della proteina del *core* con l'RNA genomico. Questo legame sarebbe selettivo per l'RNA a polarità positiva ed inoltre inibirebbe la traduzione da parte di IRES instaurando un meccanismo di passaggio da traduzione/replicazione ad assemblaggio. Altre proteine coinvolte sono P7, NS2 e NS5A. Le proteine dell'*envelope*, invece, sono trattenute nel reticolo endoplasmico e quindi il nucleocapside virale acquisisce le proteine dell'*envelope* mediante gemmazione dalle membrane del reticolo stesso. Il rilascio del virione sembra che avvenga mediante il normale *pathway* secretorio della cellula⁴⁰.

1.3 LA VARIABILITA' VIRALE

Molti organismi che causano malattie infettive, in particolare virus a RNA, hanno un alto tasso di variabilità nucleotidica che permette loro di generare mutazioni e vari adattamenti⁴¹.

Secondo la “teoria neutrale” la maggior parte dei cambiamenti di sequenza che avvengono all'interno e tra le specie non hanno un effetto significativo sul fenotipo.

La maggior parte delle sostituzioni nucleotidiche che si accumulano nel tempo sono sinonime, probabilmente perché meno deleterie per il fenotipo virale di quelle che producono sostituzioni aminoacidiche (dette non-sinonime). L'accumulo di sostituzioni sinonime fornisce un “orologio molecolare” che rende possibile stimare i tempi di divergenza delle varianti di HCV in una popolazione. Cambiamenti nucleotidici nelle sequenze codificanti e non codificanti, con minimo o alcun effetto sulla *fitness* di un organismo, vengono comunque accumulate nella popolazione. In questo modo membri di una stessa specie, ma geograficamente isolati tra loro, possono differire molto geneticamente, rimanendo invece invariati nella morfologia e nei comportamenti.

Un possibile esempio di cambiamento adattativo in HCV è la rapida evoluzione della regione ipervariabile delle glicoproteine dell'*envelope* di E2 al fine di prevenire il riconoscimento da parte degli anticorpi neutralizzanti indotti dall'infezione.

Al contrario, variazioni di sequenza neutrali che non alterano consistentemente il fenotipo virale sono alla base della maggior parte della diversità che si osserva tra le popolazioni di HCV separate geograficamente⁴¹.

Come per gli altri virus a RNA, anche HCV è caratterizzato da un elevato grado di variabilità intrinseca, dovuto al fatto che, durante la replicazione dell'RNA del virus, errori sono introdotti in maniera random nel genoma dall'RNA polimerasi virale. La polimerasi virale manca, infatti, dell'attività di *proofreading* o di correzione delle bozze (attività esonucleasica 3'-5') per cui non corregge gli errori introdotti nella replicazione. Alcuni di questi inserimenti errati sono letali per il virus e non vengono trasmessi, altri sono silenti e portano a cambiamenti nella sequenza nucleotidica senza però cambiare la codifica dei corrispondenti aminoacidi. Infine un terzo tipo di mutazione determina un cambio nella sequenza aminoacidica. Queste ultime mutazioni non influenzano la virulenza ma in alcuni casi possono conferire un vantaggio al virus che quindi può replicare più efficientemente o eludere la risposta immunitaria dell'ospite, generando varianti *escape mutant*¹⁵.

Il tasso di incorporazione di nucleotidi sbagliati da parte della RNA polimerasi di HCV è di circa 10^{-3} - 10^{-4} sostituzioni di basi per sito genomico per anno⁴². Il tasso di mutazione di HCV e degli altri virus a RNA, è quindi molto elevato, 10^3 - 10^6 volte più grande di quello dei virus a DNA.

Molti virus riescono a tollerare la sostituzione di più della metà dei loro nucleotidi senza perdere capacità replicativa. L'alto grado di variabilità e la loro tolleranza alla diversità genetica generano la coesistenza di popolazioni di genomi virali strettamente correlati ma distinti in uno stesso paziente. Questo tipo di popolazione virale è definita *quasispecie*. Le varianti presenti nelle quasispecie non sono solo mutanti generati dagli errori del processo di replicazione, ma sono piuttosto il risultato di una selezione competitiva basata sulla capacità replicativa di tali mutanti in un dato microambiente. Così genomi che hanno una capacità maggiore di produrre prole, ossia dotati di maggiore *fitness*, rispetto agli altri, saranno più rappresentati nella popolazione. In ogni momento della storia naturale dell'infezione la distribuzione delle quasispecie rappresenta la popolazione ottimale che ha raggiunto un equilibrio, sebbene transiente, con l'ospite in cui si trova.

Inoltre, altra fonte di variabilità deriva dalla ricombinazione che è un evento importante per i virus a RNA; in HCV l'evento ricombinativo, invece, è molto scarso; poiché le cellule HCV-infette sono refrattarie ad una successiva infezione, meccanismo alla base del fenomeno della ricombinazione. Tuttavia negli ultimi anni sono stati identificati alcuni ricombinanti naturali inter-genotipo o inter-sottotipo che stanno a dimostrare che tale fenomeno può avvenire durante la replicazione dell'RNA. Per rilevare tali eventi una cellula deve essere infettata da due o più varianti del virus nello stesso individuo. Il primo lavoro che attesta la presenza di ceppi ricombinanti è stato pubblicato a St Pietroburgo (Russia) nel 2002. Gli autori di questo studio hanno identificato un inter-genotipo ricombinante tra il sottotipo 2k e il sottotipo 1b. Il *breakpoint* di ricombinazione è stato mappato il gene NS2, attorno alla posizione 3.175. Ciò ha portato alla proposta di designazione del nuovo ceppo come RF1_2k/1b, in analogia con la nomenclatura utilizzata per forme circolanti ricombinanti di HIV. Otto differenti ricombinanti inter-genotipo sono stati descritti finora; essi interessano tutti i genotipi tranne il 4.

La ricombinazione in HCV è ancora un evento raro e il numero di casi ben documentati è ancora molto basso. Vi sono difficoltà a diversi livelli per la produzione e rivelazione di ceppi ricombinanti di questo virus e, anche se vi è ancora molta incertezza sulla sua

incidenza reale, è probabile che i livelli di ricombinazione in HCV vengano sottovalutati⁴³. Molte analisi genotipiche di routine per HCV, infatti, sono basate su piccole porzioni genomiche che non sono sottoposte a sequenziamento.

Per tutte queste ragioni discusse, la variabilità genetica è una delle caratteristiche più importanti di HCV ed è distribuita lungo tutto il genoma. Le regioni codificanti NS3 e NS5B mostrano un grado di eterogeneità paragonabile a quello della media dell'intero genoma. Le regioni 5' e 3' UTR e il *core* sono molto conservate, mentre le glicoproteine E1 e E2 sono più variabili, con E2 contenente HVR1 ad elevatissima variabilità.

Variazioni genetiche all'interno di regioni ben conservate del genoma sono state studiate approfonditamente e usate come base per la classificazione genotipica degli isolati.¹⁹.

Comparando le sequenze di campioni provenienti da pazienti cronicamente infetti è stato stimato un tasso di cambiamento di sequenza pari a 1.4×10^{-3} variazioni nucleotidiche per sito per anno in tutto il genoma e pari a 5 e 7.9×10^{-4} cambiamenti per sito per anno nella regione NS5 e E1 rispettivamente. Nella regione codificante del genoma i cambiamenti avvengono preferenzialmente in siti sinonimi e rappresentano l'accumulo di cambiamenti fenotipici neutrali che avvengono con tasso costante nel tempo. Per questa ragione si possono usare per stimare il tempo di diffusione di HCV in una catena di trasmissione⁴⁴.

1.3.1 Genotipi e sottotipi di HCV

L'analisi filogenetica, basata sull'intero genoma di HCV, ha permesso di identificare 7 genotipi principali che possono essere ulteriormente suddivisi in vari sottotipi⁴⁵.

Confermata l'eterogeneità genetica del virus HCV, furono utilizzati diversi sistemi di classificazione delle varianti generando molta confusione in letteratura. Durante la Seconda Conferenza Internazionale di HCV e Virus Correlati fu così proposta una nomenclatura consenso per classificare i diversi genotipi e sottotipi.

Un documento di consenso, pubblicato nel 2005, propose la classificazione dell'infezione da HCV attraverso metodi filogenetici in 7 genotipi maggiori contrassegnati con numeri arabi attribuiti secondo l'ordine della scoperta e un numero variabile di sottotipi strettamente connessi geneticamente (ed epidemiologicamente), contrassegnati con lettere minuscole in ordine alfabetico e cronologico (Figura 5). I genotipi differiscono l'uno dall'altro dal 31% al 33% a livello nucleotidico. Ad oggi,

sono stati identificati nel mondo più di 67 sottotipi confermati e 20 provvisori che, a loro volta, differiscono l'un l'altro del 15%. Il numero di nuovi sottotipi è ancora in aumento per tutti genotipi escluso il 5. Questi genotipi sono la dimostrazione dell'eterogeneità e della continua evoluzione di HCV⁴⁵.

Negli ultimi anni, con la messa a punto di nuovi test commerciali per l'identificazione genotipica di HCV, un consistente numero di studi ha concentrato l'attenzione sull'epidemiologia del virus e sulle possibili interazioni tra i genotipi e la patogenesi o la terapia di HCV.

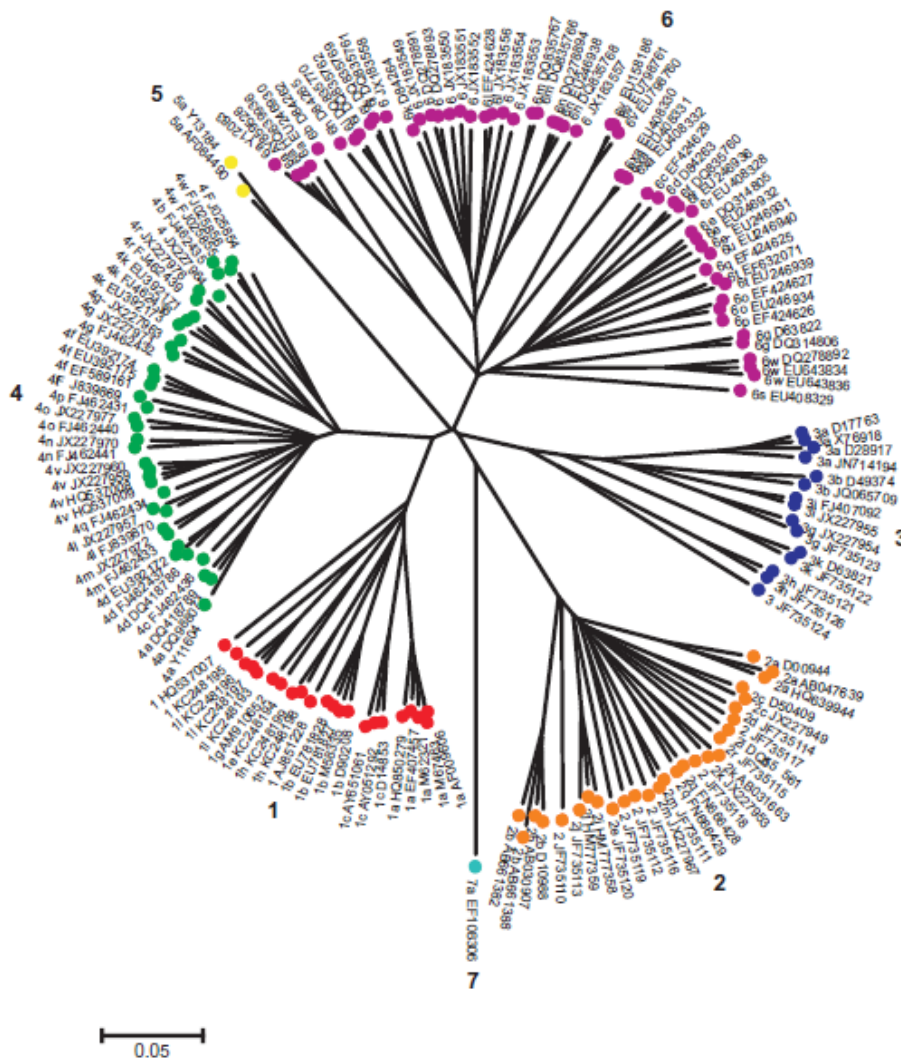


Figura 5: Albero filogenetico per la classificazione di HCV in 7 distinti genotipi, contrassegnati con numeri arabi nell'ordine di scoperta.

1.3.2 Distribuzione geografica dei genotipi di HCV

La distribuzione su scala mondiale dei 7 diversi genotipi di HCV risulta differente e complessa. (Figura 6)

Probabilmente la diversa distribuzione dei genotipi riflette la differente storia epidemiologica del virus nel tempo. Studi filogenetici ed epidemiologici hanno dimostrato che un'elevata variabilità genetica tra ceppi che sono presenti in una limitata area geografica può essere acquisita solo durante un'evoluzione a lungo termine; al contrario ceppi che mostrano una bassa distanza genetica e risultano largamente diffusi possono avere una più breve storia e un alto tasso di trasmissione.

HCV mostra entrambi gli schemi evolutivi.

Alcuni genotipi di HCV sono, così, considerati ceppi "endemic" perché circolano da molto tempo in regioni ristrette del mondo e mostrano un'elevata variabilità genetica. Essi sono il genotipo 1 e 2 in Africa Occidentale, il 3 ritrovato in Asia meridionale, il genotipo 4 in Africa centrale e Medio Oriente.

Mentre parecchi sottotipi, specialmente 1a, 1b, 3a, 2a, 2b e 2c, sono ampiamente distribuiti nel mondo, sostengono una larga proporzione di infezioni ma mostrano una bassa variabilità. Essi sono chiamati sottotipi "epidemic" perché si sono diffusi rapidamente nello scorso secolo attraverso sangue e siringhe infette, uso di droghe per via iniettiva e altre modalità.

Oltre a queste tipologie di diffusione endemica o epidemica, alcuni genotipi/sottotipi di HCV si sono disseminati rapidamente in una limitata area geografica e vengono, per questa ragione, denominati "*local epidemics*" come il sottotipi 4a in Egitto o il 5a in Belgio⁴⁶.

Genotipo 1: è distribuito in tutto il mondo ed è il responsabile della maggioranza dei casi in America, Europa, Australia e Giappone. I sottotipi 1a e 1b sono i sottotipi più comuni negli Stati Uniti e in Europa. L'1a è predominante in Europa occidentale e in Nord America; l'1b è il più comune in Europa e Asia.

Genotipo 2: è prevalente in alcuni paesi dell'Africa. Il sottotipo 2b è comune nel Nord America, Europa e Giappone, il 2c è molto comune nell'Europa Occidentale e Meridionale e in un'area ristretta nel Nord Italia.

Genotipo 3: Il sottotipo 3a è particolarmente prevalente nei tossicodipendenti in Europa, USA, Australia e Sud Asia ed è il genotipo più comune in Thailandia; 3b è

confinato in Nepal, Tailandia e in Bangladesh; i non-3a sono presenti in India e Pakistan; 3c-3f sono presenti solo in Nepal.

Genotipo 4: E' prevalente nel Nord Africa e Medio Oriente. Il 4a è il principale sottotipo in Egitto mentre il 4c è prevalente nell'Africa centrale.

Genotipo 5: Sembra essere confinato in sud Africa ma si è diffuso ampiamente in Belgio e Francia.

Genotipo 6: Insieme ai suoi numerosi sottotipi è localizzato principalmente nel sud-est asiatico.

Genotipo 7: è stato recentemente identificato in Canada e Belgio in pazienti immigrati dall'Africa centrale ^{47,48}.

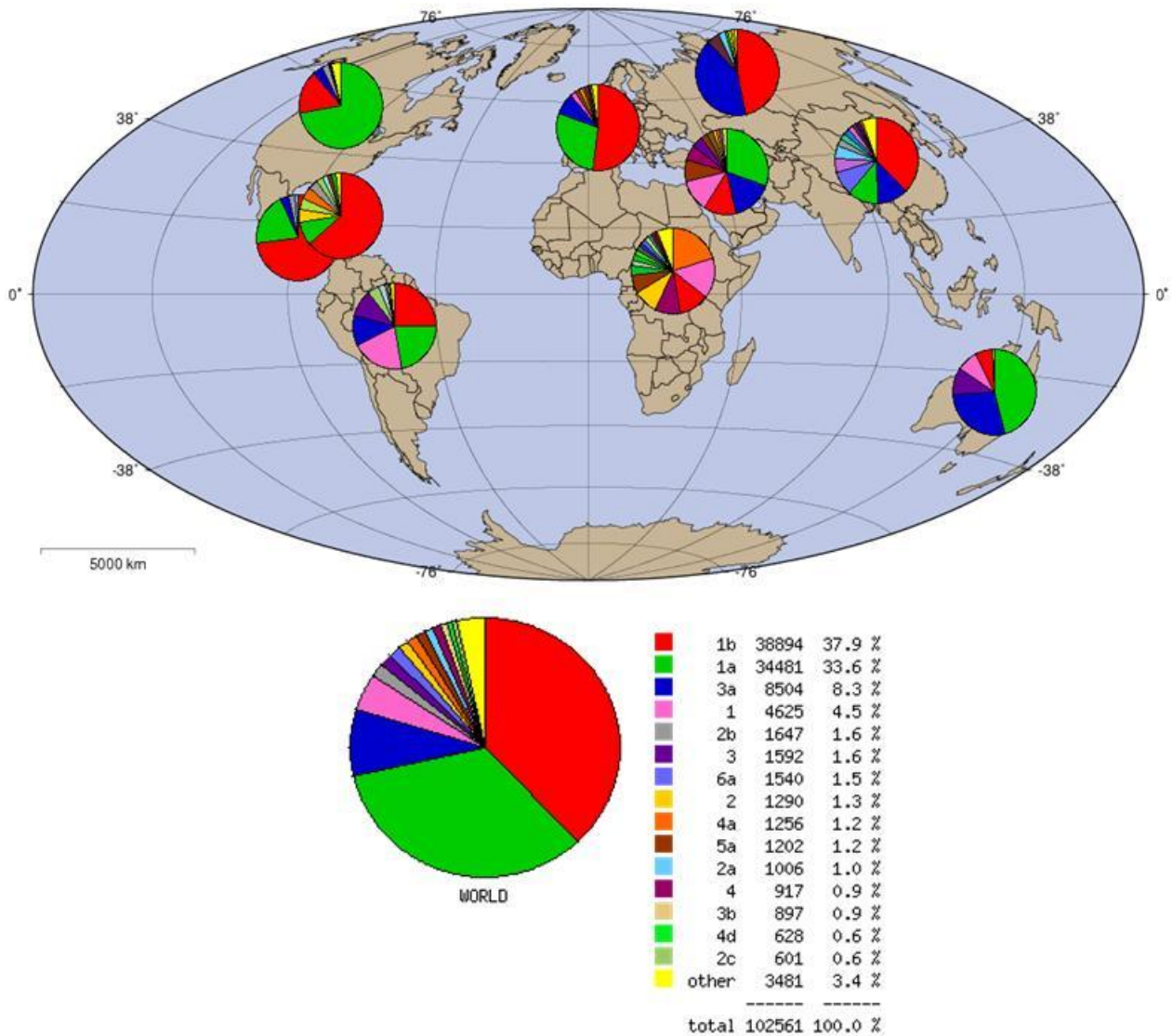


Figura 6: Distribuzione dei genotipi e sottotipi di HCV in accordo alle sequenze sottomesse nel database Los Alamos (<http://hcv.lanl.gov/content/index>).

La distribuzione geografica e la diversità dei genotipi di HCV può fornire indizi circa l'origine storica di HCV⁴⁹.

L'analisi del genotipo può fornire anche informazioni sulle diverse ondate epidemiologiche avvenute nel tempo. Alcuni studi europei, ad esempio, hanno documentato l'emergenza dei sottotipi 1a e 3a in soggetti giovani e una diminuzione del 2a/c e 1b, riportando l'elevata diffusione dei primi tra i tossicodipendenti e la loro recente diffusione da questo *reservoir* alla popolazione generale⁴⁸.

Infine, usando l'orologio molecolare, si può verificare che il sottotipo 1b è originato da un antecessore comune circa 76 anni fa; mentre altri sottotipi, come 1a e 3a, presenti in Europa, in base alla bassa divergenza dei siti sinonimi, sarebbero derivati da un comune antecessore circa 50 anni fa. Al contrario 2c, che presenta elevate divergenze tra le diverse varianti, si sarebbe originato almeno 150-200 anni fa. I genotipi 1b e 2, infatti, sono responsabili delle infezioni in individui tra i 50-70 anni prevalentemente nell'est (1b) e nel sud dell'Europa (1b e 2), mentre 1a e 3a sono presenti in popolazioni più giovani e principalmente tossicodipendenti.

Se il tasso di mutazione si fosse mantenuto per lunghi periodi, sarebbe stato possibile calcolare i tempi di divergenza dei diversi sottotipi o genotipi di HCV. La variabilità così estesa di HCV rende invece difficile la ricostruzione evolutiva. In particolare risulta complicato stimare le distanze tra varianti per l'elevata frequenza di sostituzioni multiple, spesso non identificabili, ad ogni sito nucleotidico.

Le stime minime vanno da 600 anni per il tempo di divergenza dei sottotipi di HCV a 1.000-2.000 anni per la divergenza dei 6 diversi genotipi. Sicuramente si tratta di una sottostima, ma è un'indicazione preliminare della grande distanza delle popolazioni virali attuali. La mancanza di contatti tra i diversi continenti, almeno fino a tempi relativamente recenti, non giustifica l'attuale distribuzione geografica di HCV. In mancanza di viaggi ad ampio raggio sembra difficile trovare un meccanismo con cui l'antecessore comune di tutte le varianti di HCV si sia diffuso in tutte le popolazioni e differenziato in genotipi regione-specifici in un periodo di 2.000 anni.

Inoltre non c'è alcuna evidenza di come HCV o una sua variante abbia infettato i primati. Non si è sicuri, come per HIV, che vi possa essere stata una trasmissione cross-specie che ha portato a infettare la popolazione umana⁵⁰.

1.4 EPIDEMIOLOGIA DEL VIRUS HCV

Fin dalla sua scoperta, il virus dell'epatite C è stato riconosciuto come la maggiore causa di malattie epatiche croniche in tutto il mondo. Da una stima dell'Organizzazione Mondiale della Sanità (OMS) circa 130-170 milioni di persone, approssimativamente il 3% della popolazione mondiale, sono cronicamente infette da HCV ed a rischio di sviluppare cirrosi epatica e/o cancro del fegato. Inoltre si stima che 3-4 milioni di persone sono infette con HCV ogni anno e che, attualmente 15 milioni sono infette nei paesi europei⁵¹.

Poiché di molti paesi non sono disponibili dati, questa stima si basa sulle medie ponderate per regioni anziché per singoli paesi⁵².

Sebbene HCV è endemico in tutto il mondo, vi è un grande grado di variabilità geografica nella sua distribuzione. I paesi con i più alti tassi di prevalenza si trovano in Africa e in Asia dove l'Egitto, Cameroon, Arabia Saudita, Iraq e Siria contano la maggior parte dei casi e la prevalenza oscilla tra il 2% al 15%; aree a bassa prevalenza includono le nazioni industrializzate del Nord America, Europa settentrionale e occidentale e l'Australia (<2%) (Figura 7)^{53,54}.

Nazioni molto popolate del mondo sviluppato con tassi relativamente bassi di sieroprevalenza di HCV sono: Germania (0,6%), Canada (0,8%), Francia (1,1%), e Australia (1,1%).

Bassi tassi di sieroprevalenza, anche se leggermente superiori, sono stati riportati negli Stati Uniti (1,8%), in Giappone (1,5-2,3%) e Italia (2,2%)⁵⁵.

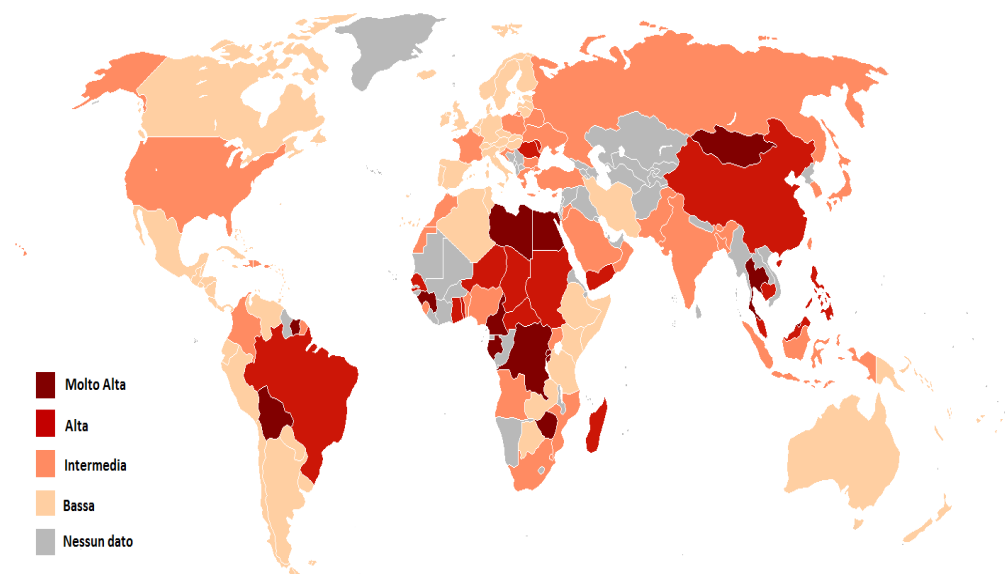


Figura 7: Prevalenza di HCV nel mondo.

Sulla base di studi effettuati su donatori di sangue, la prevalenza più bassa (0.01-0.1%) è stata riportata nel Regno Unito e in Scandinavia; prevalenze leggermente superiori (0.2-0.5%) si trovano nell'Europa occidentale, nel Nord America (USA) soprattutto tra i Caucasicci e Ispanici, in molte aree dell'America Centrale e del Sud, in Australia e in regioni limitate dell'Africa, incluso il Sud Africa.

Prevalenze medio-alte (1-5%) sono state riscontrate in Brasile, nell'Europa dell'est, nel Mediterraneo, nel Medio Oriente, nell'India sub continentale, in aree dell'Africa e dell'Asia. Cina, India, Egitto, Pakistan e Indonesia mostrano la metà dei soggetti infetti da HCV nel mondo.

In generale ai paesi in via di sviluppo è associato un maggiore peso relativo alle infezioni da HCV ma presentano anche maggiori limitazioni nella sorveglianza dell'infezione; dati provenienti da paesi africani, asiatici e sud americani sono mancanti. L'epidemiologia di HCV in Europa è eterogenea: Europa settentrionale e occidentale riportano basse prevalenze (<1%) mentre Europa meridionale e orientale riportano una prevalenza da intermedia ad alta (>2,5%). In particolare, la completezza delle informazioni è limitata in molti paesi dell'area balcanica e del Mediterraneo, specialmente per alcuni gruppi a rischio⁴⁸.

Paesi quali Stati Uniti, Australia, Turchia, Spagna, Italia e Giappone appartengono a regioni del mondo con una prevalenza media complessiva dell'infezione di HCV simile tra loro (1-1,9%) ma hanno differenti profili di prevalenza in correlazione all'età. Negli Stati Uniti la prevalenza è più alta tra le persone di 30-40 anni mentre è più bassa tra persone con meno di 20 anni e con più di 50 anni. Questo profilo, tipico degli Stati Uniti e dell'Australia, indica che la trasmissione è avvenuta negli ultimi 30-40 anni e prevalentemente tra i giovani adulti⁵⁶.

In altri paesi come Turchia, Spagna, Italia, Giappone e Cina la prevalenza età-specifica aumenta con l'età. La prevalenza, infatti, aumenta tra i soggetti più anziani che rappresentano la fascia più colpita e che coincide con un'infezione contratta nel lontano passato (30-50 anni fa)⁵⁷.

La più alta prevalenza nel mondo è presente in Egitto e aumenta gradualmente con l'età e in tutte le fasce. Questo profilo indica un alto rischio di infezione nel passato mantenutosi costante nel tempo.

Questa prevalenza età-specifica dell'infezione da HCV fornisce informazioni che aiutano a comprendere come l'entità della malattia possa cambiare nel tempo.

Determinare l'incidenza di infezione da HCV (cioè, il tasso di infezioni di nuova acquisizione) è difficile, perché la maggior parte delle infezioni acute sono asintomatiche, i test disponibili non distinguono infezione acuta da infezione cronica o risolta, e la maggior parte dei paesi non raccolgono sistematicamente dati sui casi di malattia acuta. Anche nei paesi con consolidati sistemi di sorveglianza e sistemi di segnalazione delle malattie acute si sottovaluta l'incidenza di infezione da HCV⁴⁹.

In conclusione, il virus dell'epatite C è una delle maggiori cause di epatite nel mondo e una delle potenziali cause di mortalità e morbidità nel mondo, rappresentando, in questo modo, un grande problema per la Sanità Pubblica. La complessità e l'incertezza legata alla distribuzione geografica dell'infezione di HCV, alla determinazione dei fattori di rischio associati e alla valutazione dei cofattori che accelerano la sua progressione, sottolineano le difficoltà nella prevenzione e controllo globale di HCV⁵⁵.

1.4.1 Modalità di trasmissione

Durante gli ultimi dieci anni l'introduzione di pratiche mediche più sicure ed una migliore educazione e informazione, hanno contribuito a far diminuire l'incidenza delle epatiti C acute, dovuta a una maggior sicurezza nelle trasfusioni di sangue e al miglioramento delle condizioni sanitarie, pur essendoci una continua espansione dell'uso di droghe per via endovenosa e immigrazione di persone che vivono in aree ad elevata prevalenza del virus. La maggior parte delle differenze epidemiologiche geografiche e temporali possono essere correlate all'importanza relativa dei diversi fattori di rischio nel tempo.

I fattori di rischio associati con l'acquisizione dell'infezione sono stati determinati da studi di coorte (prospettivi) o caso-controllo (retrospettivi) su persone con infezione acuta e includono: trasfusioni di sangue proveniente da trapianti d'organo e da donatori infetti, uso droghe iniettabili, iniezioni non sicure, esposizione lavorativa al sangue, nascita da madre infetta, rapporti sessuali con partner infetti e rapporti con partner multipli⁵².

Generalmente la via di trasmissione prevalente di HCV è quella parenterale: prima del 1990 la via più comune di infezione era la trasfusione di sangue, quella attuale è rappresentata dalla tossicodipendenza. Condizioni di vita socio-economiche povere⁵⁸, basso livello di scolarità, rapporti sessuali ad alto rischio, elevato numero di partner sessuali, sono tutti fattori legati ad un maggior rischio di contrarre l'infezione.

Esistono casi di infezione in cui il fattore di rischio non è chiaramente definito e che rappresentano infezioni parenterali inapparenti “acquisite in comunità”, presupponendo una via di trasmissione diversa da quella “percutanea aperta”.

In alcune parti del mondo inoltre, alcune procedure di cosmesi o “pratiche culturali” rappresentano fattori importanti nella trasmissione dell’infezione.

1.4.1.1 Trasmissione parenterale

La trasfusione di sangue è una modalità più frequente di infezione di HCV. Nei paesi sviluppati del mondo le numerose misure di controllo, attuate negli ultimi decenni, hanno permesso una progressiva riduzione del rischio di trasmissione dell’infezione di HCV tramite trasfusione. Queste misure implicano l’adozione di un sistema di volontari e di *screening* delle donazioni. La selezione sempre più rigorosa dei donatori di sangue, intrapresa per la prevenzione dell’AIDS, ha determinato l’esclusione dei soggetti esposti a specifici fattori di rischio e ha comportato, negli ultimi anni, a una diminuzione dell’incidenza di infezione tramite trasmissione parenterale⁵². In seguito all’introduzione del test ELISA per la ricerca di anticorpi anti-HCV nei donatori, il rischio di epatite post-trasfusionale è diventato praticamente nullo⁵⁹.

I tossicodipendenti costituiscono un importante serbatoio del virus HCV in tutti i paesi industrializzati⁶⁰. In paesi come gli USA e l’Australia, dove la maggior parte delle infezioni è stata contratta in giovane età, l’uso illegale delle droghe endovenose è stato il principale modo di trasmissione di HCV durante gli ultimi 30 anni ed è responsabile del 68-80% delle infezioni nei due Paesi. L’infezione da HCV viene acquisita con questa modalità più rapidamente rispetto ad altre infezioni virali, così dopo 6-12 mesi di tossicodipendenza circa l’80% risulta infetto. In Paesi dove l’incidenza di epatite C è dovuta ad infezioni contratte nel lontano passato, l’uso di droghe endovenose sembra aver giocato un ruolo minore; tale modalità è associata alle infezioni più recenti.

In diversi paesi europei l’uso di droghe è stato identificato come il fattore di rischio dominante per HCV all’interno dei loro confini. Ad esempio in Italia uso di droghe è stato riportato come fattore di rischio tra casi di epatite C dal 1994 al 1996 tra pazienti di età compresa tra i 14 e 24 anni. Pochi dati sono disponibili riguardo la prevalenza dei tossicodipendenti nei paesi in via di sviluppo⁵².

1.4.1.2 Altre sorgenti di trasmissione

Altre modalità di trasmissione sono rappresentate dall'esposizione professionale, perinatale e sessuale ma tali modalità hanno, tuttavia, frequenze minori rispetto all'esposizione percutanea. La trasmissione professionale è confinata ai lavoratori in ambito sanitario che hanno subito una contaminazione con materiale infetto; la percentuale di acquisizione dell'infezione in queste circostanze è pari allo 0.3%.

La trasmissione perinatale di HCV avviene principalmente durante le manovre del parto. L'HCV-RNA è stato, infatti, trovato sia nel fluido puerperale che nel sangue mestruale. La percentuale di trasmissione perinatale di bambini nati da madri infette è pari al 2.7-8.4% senza differenze geografiche e temporali.

Diversi studi hanno focalizzato l'attenzione sull'efficacia relativa alla via di trasmissione sessuale. A tal proposito si è ricercata la presenza di HCV-RNA in differenti liquidi biologici a livello di gruppi di soggetti a rischio quali prostitute, omosessuali, individui sessualmente promiscui e coniugi di soggetti anti-HCV positivi. La trasmissione sessuale ha evidenziato una scarsissima capacità di trasmissione del virus dell'epatite C attraverso questa via che è risultata meno efficiente di quella di altri virus sessualmente trasmessi.

Ci sono altre modalità di trasmissione plausibili dovute a una grande varietà di attività umane che comportano l'esposizione percutanea al sangue o a fluidi corporei come procedure cosmetiche o religiose, pratiche culturali come tatuaggi e piercing⁵².

1.4.2 Genotipi e fattori di rischio associati

Dall'identificazione del genoma del virus dell'epatite C, molti studi sono stati condotti per evidenziare un'associazione tra genotipo e via di trasmissione.

Diversi lavori presenti in letteratura hanno evidenziato un'associazione tra il genotipo 1b ed il 2a/2c ed infezione a carattere sporadico o la trasfusione di sangue, mentre i genotipi 1a, 3a e 4, invece, sono risultati associati alla tossicodipendenza⁶¹.

In accordo con quanto descritto in letteratura, il genotipo 1 è il più frequentemente riscontrato nei pazienti cronicamente infetti; in particolare i ceppi 1a e 1b sono i più diffusi a livello mondiale, seguiti dal genotipo 3 e dal 2 con le varianti 2a, 2b e 2c.

Considerando, inoltre, la suddivisione in classi di età, i genotipi 1b e 2a/2c sono risultati più frequenti nei soggetti più anziani, mentre i genotipi 3a e 1a sono più diffusi negli

individui più giovani. È stato notato anche l'emergere del genotipo 4c/4d è passata dall'1,6% nella classe di età più anziana al 7,2% nella classe di età più giovane.

Le indagini svolte, al fine di ricercare una possibile associazione tra i fattori di rischio e genotipi, hanno evidenziato che le due principali vie di trasmissione sono quella trasfusionale (26,3%) e la tossicodipendenza per via endovenosa (21,7%). Esposizione professionale, rapporti sessuali e/o familiari con soggetti a rischio, interventi chirurgici e tatuaggi sono fattori di rischio di minor entità con percentuali sempre inferiori al 6%.

La stratificazione dei genotipi in relazione ai fattori di rischio ha confermato l'esistenza di una associazione tra genotipi e modalità di infezione, in quanto nei pazienti con precedenti terapie trasfusionali i genotipi presenti sono soprattutto l'1b e il 2a/2c (69,5% e 23,9%).

Per quanto riguarda la distribuzione nel tempo dei genotipi, mettendo in relazione il fattore di rischio con l'età dei soggetti infettati, nei trasfusi non esiste differenza tra le prevalenze dei genotipi 1b e 2a/2c, indicando ciò una costante circolazione nel tempo di questi genotipi.

In realtà, però, considerando al posto dell'età anagrafica dei pazienti il momento dell'infezione, è stato invece riscontrato un aumento statisticamente significativo della presenza del 2a/2c nei soggetti con trasfusioni più recenti rispetto a quelle avvenute prima del 1971, con concomitante diminuzione dell'1b. Il momento della trasfusione è, quindi, un parametro più preciso da utilizzare rispetto all'età anagrafica.

Si può ipotizzare, quindi, che in tempi più lontani i principali genotipi fossero l'1b ed il 2a/2c la cui trasmissione avveniva attraverso la trasfusione di sangue. Attraverso vie di trasmissione non identificate questi genotipi potevano circolare anche tra i soggetti non trasfusi. L'introduzione dello *screening* su tutte le donazioni di sangue per anti-HCV ha interrotto la trasmissione dei genotipi 1b e 2a/2c attraverso questa via. Contemporaneamente, l'aumento della circolazione dei genotipi 1a e 3a soprattutto tra i tossicodipendenti per via endovenosa ha portato a un cambiamento della distribuzione dei vari genotipi attualmente osservati. I genotipi 1a e 3a possono essersi, quindi, trasmessi anche a soggetti giovani non tossicodipendenti attraverso vie di trasmissione non identificate⁶¹.

1.5 INFEZIONE DI HCV IN ITALIA

Le informazioni riguardanti i dati di incidenza dell'epatite C in Italia sono forniti annualmente dal Sistema Nazionale di sorveglianza per le epatiti virali acute (SEIEVA).

Si stima che HCV costituisce circa il 18% di tutti i casi di epatite acuta notificata.

La prevalenza di HCV in Italia è stata rilevata attraverso indagini sui donatori di sangue e sulla popolazione generale, mostrando una differente distribuzione geografica di HCV tra le regioni del Nord, in cui il tasso di prevalenza sembra essere minore, rispetto alle regioni del Sud. Questo fenomeno riflette le diverse condizioni socio-sanitarie esistenti negli anni passati e mette in evidenza l'importanza delle vie di trasmissione non trasfusionali⁶².

Per quanto riguarda i fattori di rischio, secondo i dati ufficiali più recenti (SEIEVA 2014), l'4,6% dei casi di epatite acuta avevano ricevuto donazioni di sangue o di emoderivati, il 25,3% erano tossicodipendenti per via endovenosa, il 21,6% erano stati sottoposti a interventi chirurgici, l'8% erano stati ospedalizzati nei 6 mesi precedenti l'esordio della malattia acuta, il 21,6% avevano ricevuto cure odontoiatriche, il 20,3% aveva un convivente con epatite C, il 22,5% era stato sottoposto a altre esposizioni parenterali come ad esempio piercing, tatuaggi, agopuntura, rasatura dal barbiere ecc (Tabella 1).

Nel nostro Paese la percentuale di soggetti infettati dall'HCV è superiore al 3% in soggetti nati prima del 1950 e progressivamente aumenta con l'età avanzata mentre è notevolmente minore nelle generazioni più giovani⁶³. In Italia il sottotipo 1b ha mostrato la più alta prevalenza in tutte le aree geografiche e gruppi di età, seguita dai sottotipi 2c (rilevati soprattutto nella popolazione anziana nel Sud Italia), 4a/d, 3a (rilevata esclusivamente in adulti) e 1a. Questi risultati confermano che l'Italia centrale e meridionale sono aree iperendemiche.

La sieroprevalenza aumenta con l'età, a seguito di un gradiente nord-sud, ed è notevolmente aumentata tra le fasce d'età 15-30 e 31-45.

L'alta prevalenza osservata negli adulti di età superiore ai 30 anni è da attribuirsi principalmente ad un aumento della prevalenza del sottotipo 1b, ma anche dei sottotipi 2c, 3 e 4. Dati di prevalenza età specifici, approfonditi nello studio, e la caratterizzazione molecolare del virus suggeriscono che, in Italia, coesistono due modelli di trasmissione: uno caratterizzato da infezioni dovute a sottotipi 1b e 2c, soprattutto negli adulti di età superiore ai 60 anni, e l'altra da sottotipi 1a, 3 e 4,

soprattutto nella fascia di età 31-60 anni, coerente con l'uso di droghe per via endovenosa⁶⁴.

Tabella 1: Casi notificati di Epatite C con fattore di rischio identificato (fonte, SEIEVA 2014).

| Fattore di rischio parenterale o sessuale | Fascia di età | | | | | | | | | | | |
|---|---------------|-------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|-----------|--------|
| | 0-14 | | 15-24 | | 25-34 | | 35-54 | | 55+ | | TOTALE | |
| | N. | (%) | N. | (%) | N. | (%) | N. | (%) | N. | (%) | N. | (%) |
| Trasfusione sanguie | 0 | (0,0) | 0 | (0,0) | 0 | (0,0) | 3 | (8,1) | 1 | (2,9) | 4 | (4,6) |
| Interventi chirurgici | 0 | (0,0) | 0 | (0,0) | 0 | (0,0) | 9 | (23,7) | 10 | (29,4) | 19 | (21,6) |
| Ospedalizzazione | 0 | (0,0) | 0 | (0,0) | 1 | (12,5) | 2 | (5,6) | 5 | (14,7) | 8 | (9,2) |
| Altre esposizioni parenterali** | 0 | (0,0) | 3 | (33,3) | 2 | (25,0) | 8 | (21,6) | 7 | (20,0) | 20 | (22,5) |
| Terapia odontoiatrica | 0 | (0,0) | 3 | (33,3) | 1 | (14,3) | 8 | (21,6) | 7 | (20,0) | 19 | (21,6) |
| Uso di droghe E.V. | 0 | (0,0) | 8 | (80,0) | 6 | (60,0) | 9 | (24,3) | 0 | (0,0) | 23 | (25,3) |
| Convivente tossicodipendente | 0 | (0,0) | 1 | (12,5) | 0 | (0,0) | 1 | (2,9) | 0 | (0,0) | 2 | (2,4) |
| Contatto con itterico nei 6 mesi | 0 | (0,0) | 3 | (33,3) | 1 | (16,7) | 5 | (15,6) | 0 | (0,0) | 9 | (11,4) |
| Partner sessuali (>1 nell'ultimo anno) | 0 | (0,0) | 2 | (22,2) | 1 | (25,0) | 3 | (12,0) | 0 | (0,0) | 6 | (11,8) |
| Convivente di soggetto HCV+ | 0 | (0,0) | 4 | (50,0) | 1 | (12,5) | 6 | (23,1) | 3 | (11,1) | 14 | (20,3) |
| TOTALE CASI*** | 0 | | 10 | | 12 | | 42 | | 35 | | 99 | |

* I casi possono avere più di un fattore di rischio

** Piercing, tatuaggi, agopuntura, manicure/pedicure, rasatura dal barbiere

*** Per alcuni casi l'informazione relativa ad alcuni fattori di rischio non è disponibile

1.6 INFEZIONE DI HCV NEI BALCANI OCCIDENTALI

Pochi sono i dati disponibili sull'infezione del virus HCV nella regione Balcanica soprattutto in regioni come Albania e il Montenegro.

Il Montenegro è uno stato europeo situato nella penisola balcanica e che si affaccia sul Mar Adriatico. Fino al 2 giugno 2006 la Repubblica del Montenegro è stata unita alla Repubblica di Serbia con il nome di **Serbia e Montenegro**. Dal 3 giugno 2006 il Montenegro è diventato uno stato indipendente, proclamato a seguito del referendum sull'indipendenza del 21 maggio 2006. Per questa ragione, in letteratura, non si trovano informazioni relative al solo Montenegro.

Mancano completamente informazioni relative all'infezione di HCV, sulla sua circolazione e prevalenza in Montenegro.

In un lavoro di Svirtlih *et al* è stata analizzata la prevalenza dei genotipi del virus HCV in Serbia e Montenegro e la loro influenza su alcune caratteristiche cliniche dei pazienti. Sono stati studiati 164 pazienti con infezione cronica di HCV. I più comuni genotipi riscontrati nello studio sono stati il genotipo 1 (57.9%), in particolare 1b (54.9%). Nel

genotipo 2 erano compresi i sottotipi 2a, 2b, 2c; mentre per il genotipi 3 era presente solo il sottotipo 3a.

La prevalenza dei genotipi in questo studio è simile alla distribuzione genotipica presente in altre parti d'Europa.

In Serbia e Montenegro, i genotipi 1b e 3a sono risultati predominanti in pazienti con infezione cronica di HCV. Il sottotipo 1b era caratteristico di pazienti anziani con età superiore a 40 anni. Il sottotipo 3a era comune tra i tossicodipendenti con età inferiore ai 40 anni, suggerendo un aumento di questa modalità di trasmissione. L'associazione del sottotipo 1b con la malattia avanzata, con un'alta carica virale e con un'alta attività istologica ha suggerito una precoce infezione con questo genotipo e un aumento della patogenicità⁶⁵.

Per quanto riguarda l'Albania, presenta un elevato livello endemico di HBV nella popolazione generale ma pochi sono gli studi condotti su HCV, suggerendo una limitata diffusione del virus (circa 1% nei rifugiati) e un livello del 5-10% tra i tossicodipendenti (dati dell'Istituto di Sanità Pubblica).

1.7 STORIA NATURALE DELL'INFEZIONE DI HCV

Dopo un'incubazione di 3 o 12 settimane, l'infezione da HCV è di solito seguita da un'epatite clinicamente asintomatica o paucisintomatica che, solo raramente, è sintomatica. L'epatite fulminante è poco comune.

Attualmente HCV è responsabile di circa il 20% dei casi di epatite acute e del 70% di quelle croniche. E' una delle maggiori cause di cirrosi e di carcinoma primitivo del fegato, e le malattie terminali del fegato correlate all'infezione sono una delle principali motivazioni di trapianto epatico.

L'infezione post-trasfusionale, per la quale è possibile definire con precisione il momento del contagio, ha permesso di stabilire la storia naturale dell'epatite C. Il quadro clinico è a lungo asintomatico e pertanto solo l'accurato *follow-up* biochimico e sierologico del paziente può consentire di rilevare la malattia. Nonostante la progressione iniziale asintomatica, l'infezione da virus dell'epatite C causa non solo uno spettro di patologie epatiche con diverso andamento e severità, ma anche un numero di manifestazioni extraepatiche principalmente correlate alla stimolazione cronica del sistema immunitario e ad una risposta autoimmunitaria virus indotta⁶⁶.

L'infezione da HCV si evolve a velocità diversa e in modi diversi nei diversi pazienti. I principali punti di diramazione includono una progressione da infezione acuta a infezione a lungo termine, da malattia a lungo termine alla cirrosi e da cirrosi a epatocarcinoma⁶⁷ (Figura 8).

1.7.1 Epatite C acuta

Il periodo di incubazione e la gravità della fase acuta possono dipendere dall'importanza dell'inoculo. Il periodo medio di incubazione è 7-8 settimane, ma può variare ampiamente (2-26 settimane). Sintomi prodromici sono rari. L'epatite acuta è itterica solo in una piccola parte di casi (20%) e non itterica con pochi o nessun sintomo nella maggior parte delle infezioni (80%). I sintomi sono generalmente aspecifici: malessere, nausea, dolore al quadrante superiore destro del fegato, urine scure e itterizia. La diagnosi clinica di epatite acuta risulta quindi difficile e può essere effettuata solo con la ricerca di *marker* virali. La forma severa dell'epatite acuta è rara e l'esistenza di epatite C fulminante è controversa. Quando è visibile clinicamente, la malattia dura generalmente 2-12 settimane.

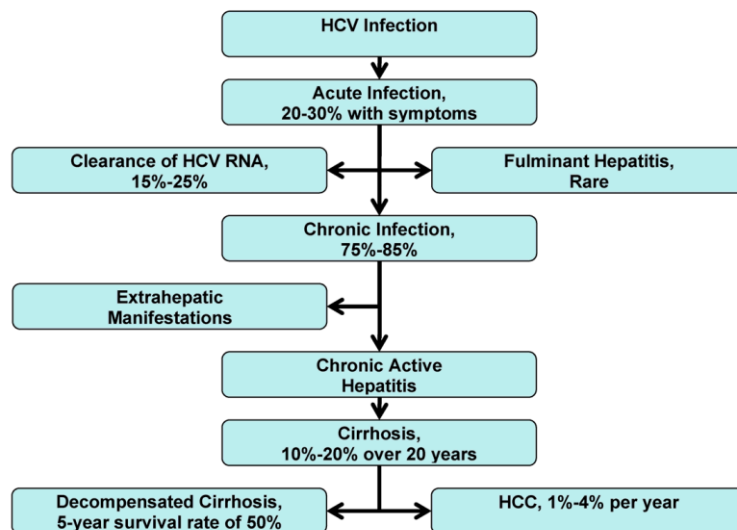


Figura 8: Storia naturale dell'infezione da HCV.

Il primo *marker* dell'infezione sono livelli di HCV-RNA rilevabili nel siero mediante PCR subito dopo la prima settimana dall'esposizione e il successivo aumento a 10^6 - 10^8 genomi/mL. Gli anticorpi diventano rilevabili in fase acuta nella maggior parte dei casi, ma altri la sierconversione è ritardata di alcune settimane. Le alanino aminotrasferasi sieriche (ALT) iniziano ad aumentare poco prima che appaiano i sintomi clinici.

Possono essere raggiunti picchi 10 volte più alti della norma, anche se in genere invece si riscontrano aumenti lievi o moderati (Figura 9).

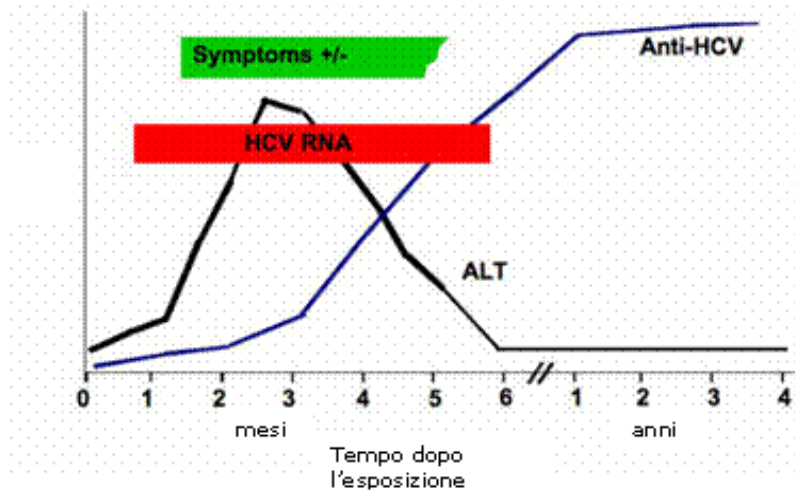


Figura 9: Andamento sierologico dei marcatori in corso di infezione acuta da HCV.

In pazienti che risolvono l'infezione, le ALT ritornano normali e l'HCV-RNA non è più rilevabile a livello plasmatico; gli anticorpi diminuiscono progressivamente anche se rimangono evidenziabili per molti anni. Non è chiaro se negli epatociti e nelle altre cellule l'infezione venga eradicata o rimanga a livelli molto bassi.

La completa risoluzione dell'epatite acuta non è un evento comune. Sembra che il tasso di risoluzione sia intorno al 15% e il livello di cronicizzazione varia dal 75-85%.

Gli studi, basati solo sulle ALT, riportavano bassi livelli di cronicizzazione, quelli più recenti, basati sulla PCR, stimano l'85% di cronicizzazione⁶⁶.

1.7.2 Epatite C cronica

L'epatite C cronica può essere definita come malattia necrotico-infiammatoria del fegato caratterizzata dalla persistenza nel tempo di necrosi epatocitaria e di infiammazione istologicamente dimostrabili. Studi prospettici indicano che circa il 75-85% di individui sviluppano un'infezione a lungo termine, definita dalla persistenza di HCV-RNA nel sangue per oltre 6 mesi. La diagnosi viene condotta attraverso parametri biochimici (elevate ALT), virologici (presenza di HCV-RNA nel siero) ed istologici. Esistono due diversi *pattern* di epatite C cronica caratterizzati da:

- **ALT normali:** circa il 25% (range 10-40%) dei soggetti con epatite C cronica ha ALT normali anche se l'HCV-RNA è presente nel siero a livelli evidenziabili. Questi pazienti sono per lo più asintomatici e le loro caratteristiche non sono diverse da quelli con ALT elevate. Sono stati denominati *carrier* sani, anche tale termine non

è del tutto corretto poiché, a volte, presentano anomalie istologiche alla biopsia epatica. Diversi studi indicano che il 25% di questi soggetti ha un'istologia epatica normale, il 54% ha una lieve epatite cronica e il 21% una epatite moderata. Normalmente la fibrosi è assente o minima e la cirrosi è presente in meno dell'1% di questi pazienti.

– **ALT elevate:** circa il 75% dei soggetti con epatite cronica ha ALT elevate. La gravità della malattia epatica può variare considerevolmente. A seconda delle lesioni istologiche epatiche, si può distinguere l'epatite cronica lieve e quella moderata-severa. Questa distinzione è importante per la prognosi della malattia e per il trattamento terapeutico. L'epatite cronica lieve viene diagnosticata in seguito ad identificazione di lesioni epatiche minori durante la biopsia epatica; definita da valori di fibrosi 0-1.

L'epatite cronica moderata-severa è definita così per la presenza di lesioni necro-infiammatorie e/o fibrosi estesa all'atto della biopsia epatica. Lo *score* della fibrosi è in questo caso 3-4. Questi pazienti rappresentano circa il 50% dei soggetti con epatite C cronica e ALT elevate e sono difficili da distinguere da quelli con epatite cronica lieve⁶⁶.

1.7.3 Cirrosi ed epatocarcinoma

La cirrosi è la complicazione più seria dell'epatite C cronica. L'infezione di solito progredisce molto lentamente e quindi la cirrosi si sviluppa generalmente in tempi lunghi (2-3 decenni). Il 20-30% dei soggetti con epatite C cronica sviluppa cirrosi dopo 10-20 anni. L'infezione cronica da HCV di solito rimane asintomatica per decenni e può anche rimanere tale, almeno inizialmente, anche se il 10-15% degli individui infettati vanno incontro allo sviluppo progressivo di fibrosi e cirrosi.

I fattori che aumentano il rischio di malattia epatica progressiva comprendono l'età più avanzata al momento dell'infezione, il sesso maschile, la razza caucasica, la coinfezione HIV, la concomitante epatite B o schistosomiasi, il consumo di alcool, il sovraccarico di ferro e malattie del fegato.

Fattori virologici come il titolo virale e il genotipo non sembrano incidere in modo significativo sulla prognosi dell'infezione. Studi recenti ipotizzano, invece, il ruolo della variabilità genetica di HCV nella progressione dell'infezione anche se l'associazione non è ben definita⁶⁸.

I segni clinici della cirrosi appaiono solo durante lo stadio terminale e includono affaticamento, debolezza muscolare, ritenzione di liquidi, bruciori e prurito. Gli esami clinici e biochimici possono avere valore predittivo di cirrosi ma non sono completamente attendibili. La diagnosi più attendibile di cirrosi viene fatta mediante la biopsia epatica.

La sopravvivenza per cirrosi durante i primi 5 anni di malattia è del 90% e dopo 10 anni si riduce all'80%. Però il rischio di decompensazione epatica è del 30% ogni anno. La cirrosi in fase terminale è una delle maggiori indicazioni per il trapianto di fegato: il 30% dei trapianti è dovuto a tale causa.

Le manifestazioni extraepatiche meno comuni dell'infezione cronica da HCV, che possono diventare clinicamente significative sono, ad esempio, crioglobulinemia mista, Linfoma *non-Hodgkin*, glomerulonefrite, artrite sieronegativa, cheratocongiuntivite secca, *lichen planus*, varie neuropatie, disturbi cognitivi, eporfiria cutanea tarda.

Il carcinoma epatico si manifesta generalmente in soggetti con cirrosi. Tra questi l'incidenza di epatocarcinoma è elevata (3-10% l'anno). Tale incidenza è del 3-5% nei Paesi occidentali mentre risulta più elevata in Asia. Normalmente il carcinoma epatico si presenta nelle cirrosi compensate ed è clinicamente silente per lungo tempo. Le terapie non sono efficaci e il trapianto d'organo può essere effettuato solo in un piccolo numero di pazienti⁶⁷.

1.8 LA TERAPIA FARMACOLOGICA

La decisione di intraprendere una terapia farmacologica si basa sull'analisi di numerose variabili come la persistenza (>6 mesi) di ALT elevate, positività a anti-HCV e RNA nel siero, fibrosi, stato infiammazione del fegato.

La terapia standard consiste in somministrazioni subcutanee di 3 milioni di unità di **Interferon α , IFN- α (o pegilato)** 3 volte a settimana e in una somministrazione orale di 1000-1200 mg di **Ribavarina (RBV)** al giorno. IFN- α e ribavarina operano nella stessa direzione attraverso un meccanismo antivirale e immunomodulatorio⁶⁹.

Sia come monoterapia che in combinazione con altri agenti antivirali, costituisce lo scheletro degli strumenti usati per combattere l'infezione. L'IFN- α costituisce una famiglia di proteine naturali con pesi molecolari di 15-21 Kdal, descritte per la prima volta nel 1957 e così definite per la loro capacità di "interferire" con la replicazione virale. Sono state utilizzate per il trattamento delle epatiti B, C e D. Vengono secrete

dalle cellule dei mammiferi ed esercitano un ampio spettro di azioni antivirali, antiproliferative, e immunomodulatorie. IFN α appartiene ad una famiglia di 25 citochine, che differiscono nella loro composizione aminoacidica e sono soggette a diversi livelli di glicosilazione. A livello commerciale sono disponibili 4 forme di IFN α ciascuna con uguale efficacia, usate a livello mondiale per curare l'infezione da HCV. Tre di queste forme sono prodotte mediante tecnica del DNA ricombinante (tutte approvate dalla *Food and Drug Administration*). L'unica forma naturale (IFN linfoblastoide/leucocitico) è usata estensivamente in Europa ma non è disponibile negli USA⁷⁰. La ribavarina è un analogo della guanosina con un'ampia attività antivirale. La monoterapia con ribavarina ha un'attività limitata per l'infezione da HCV⁷¹.

L'efficacia dell'IFN- α è monitorata in base alla risposta biochimica (normalizzazione delle ALT sieriche) e virologica (diminuzione di HCV-RNA nel siero). Fattori che favoriscono un buon esito della terapia sono: bassi livelli di viremia ($< 2 \times 10^6$ copie/mL), genotipo non di tipo 1 e assenza di cirrosi⁷². I differenti genotipi e sottotipi variano nella risposta al trattamento con IFN- α o IFN- α /RBV. Solo il 10-20% degli individui cronicamente infetti con genotipo 1 trattati con IFN- α in monoterapia e il 40-50% in combinazione con RBV mostrano una completa e permanente risoluzione dell'infezione. I genotipi 2 e 3, considerati più "facili" da trattare, vengono curati con solo 24 settimane di terapia, con maggiore probabilità di guarigione (70-80% dei casi nel genotipo 3 e oltre 50% nel genotipo 2)⁷³.

Lievi effetti collaterali sono inevitabili durante la terapia e si manifestano in oltre il 90% dei trattati. Molti pazienti presentano sintomi simil-influenzali subito dopo le prime ore dall'iniezione come affaticamento ($>25\%$), mal di testa ($>35\%$), febbre e mialgia (30%).

La caratterizzazione molecolare del genoma di HCV ha reso possibile l'identificazione di *target* molecolari per lo sviluppo di farmaci antiretrovirali. Specifici inibitori virali di enzimi prodotti da HCV offrono nuove speranze per efficaci terapie combinate, analoghe a quelle attualmente adottate per HIV, per il controllo dell'infezione da HCV e delle sue complicazioni. Ad esempio l'RNA polimerasi RNA-dipendente (RpRd) di HCV, enzima indispensabile per la replicazione virale, rappresenta un buon *target* molecolare.

Recentemente è nata la necessità di sviluppare nuovi agenti antivirali sia dalle industrie chimiche che da sorgenti naturali⁷³.

Questi farmaci, comunemente chiamati **Agenti Antivirali Diretti (DAA)** includono nuovi inibitori della replicazione virale, della elicasi, della NS3-4A serin proteasi, della polimerasi RdRp sono stati sintetizzati e alcuni di questi sono in differenti fasi di sperimentazione clinica.

Per quanto riguarda gli inibitori della proteasi NS3/4A, essi inibiscono l'attività enzimatica della proteasi mimando il prodotto finale della reazione proteolitica.

Il **Telaprevir** è un inibitore selettivo della proteasi NS3/4A somministrato oralmente ogni 8 ore; il **Boceprevir** è anch'esso un inibitore lineare somministrato tre volte al giorno. È importante notare che telaprevir e boceprevir sono stati valutati con studi differenti perciò un loro confronto è complicato. Entrambi sono attivi contro il genotipo 1.

A causa della loro bassa barriera genetica non possono essere usati in monoterapia ma devono essere somministrati in combinazione con altri composti antivirali più potenti e privi di cross-resistenza. La combinazione tra IFN α pegilato e ribavarina con telaprevir/boceprevir è stata studiata in studi clinici in fase II e III. Così nel 2011 il risultato di questi studi di fase III hanno portato all'approvazione sia del telaprevir sia del boceprevir in combinazione con IFN α pegilato e ribavarina dalla *U.S. Food and Drug Administration (FDA)*.

La durata del trattamento deve essere adattata alla gravità della malattia al fegato, alla risposta virologica all'IFN α pegilato con ribavarina in pazienti precedentemente trattati e alla risposta virologica durante la terapia. Sono stati registrati tassi di risposta virologica del 66-75% in pazienti non precedentemente trattati con terapia standard, il valore ottenuto è nettamente superiore rispetto a quello ottenuto solo con somministrazione di IFN α pegilato e ribavarina standard.

Oltre alla complessità del trattamento, le nuove terapie triple sollevarono, tuttavia, una serie di questioni legate alla loro limitazione al tipo di genotipo, elevati e ulteriori effetti collaterali, selezione di varianti resistenti ai farmaci e problemi di costi⁷⁴.

Per tutti questi motivi, durante il 2015, la FDA e EMA hanno approvato 6-7 nuovi DAA includendo l'inibitore della polimerasi, **Sofosbuvir**; l'inibitore della proteasi di seconda generazione, **Simeprevir**; inibitore dell'NS5A, **Daclatasvir**; la combinazione di Sofosbuvir più inibitore dell'NS5A **Ledipasvir** e i regimi 3D di **Paritaprevir**, **Ombitasvir** e **Dasabuvir**.

Scopo della terapia attuale di HCV è eradicare la malattia nel paziente evitando che essa proceda verso la cirrosi e la sue complicanze, e con queste nuove classi di farmaci ciò si

può ottenere grazie alla loro elevata tollerabilità, pangenotipicità, diminuita tossicità e un ottimo successo terapeutico.

Questi notevoli passi avanti nella clinica dell'epatite C sono unici nella moderna medicina. Per la prima volta epatite cronica può essere curata nel 90% dei pazienti con 3 mesi di terapia. Ma i costi molto elevati di tali terapie hanno limitato accessibilità a tali cure⁷⁵.

2. SCOPO DELLO STUDIO

Il virus dell'epatite C (HCV), con una prevalenza mondiale del 3% circa (ovvero 130-170 milioni di persone infette) e più di 15 milioni di persone infette nei paesi europei, costituisce, ancora oggi, un problema rilevante per la Sanità Pubblica e una delle maggiori cause di malattia epatica cronica, di cirrosi e di carcinoma epatico. Esistono tuttavia differenze geografiche molto marcate sui dati di prevalenza di HCV nel mondo, con proporzioni che variano dall'1.8% nel nord America al 9.6-13.6% nel nord Africa.

Il livello endemico di queste infezioni nel Mediterraneo è diminuito in modo significativo in alcuni Paesi quali Italia, Grecia, Spagna e Francia, mentre in altri Paesi come Albania, Turchia e Bulgaria esistono ancora livelli di alta endemicità.

Montenegro e Albania sono due paesi che si affacciano sul Mediterraneo. Il Montenegro è confinante con la Serbia e l'Albania e nel 2006 si è reso autonomo rispetto alla Serbia e alle altre nazioni facenti parte della ex-Yugoslavia; mentre l'Albania è confinante con il Montenegro, la Grecia e il Kosovo. Riguardo l'infezione da HCV in Montenegro e Albania, al momento attuale, i dati di prevalenza e di distribuzione genotipica sono scarsi.

Scopo del presente lavoro di tesi è stato quello di ricostruire la storia epidemiologica del virus HCV, utilizzando un approccio filogenetico.

Per tale ragione lo studio ha avuto come obiettivo specifico l'analisi delle sequenze nucleotidiche del gene NS5B di HCV ottenute da campioni di siero raccolti da pazienti affetti da epatite cronica di tipo C, provenienti dal Montenegro e dall'Albania allo scopo di:

- determinare il genotipo/sottotipo di HCV infettante sulla base dell'approccio filogenetico in modo da studiarne distribuzione, prevalenza, origini e dinamiche evolutive;
- effettuare uno studio filogeografico nell'Europa dell'Est e nell'area del Mediterraneo analizzando i dati di sequenza ottenuti nel presente studio e confrontandoli con quelli di sequenze, reperibili nelle principali banche dati, provenienti da isolati di differenti Paesi.

La ricostruzione filogenetica e filogeografica dei genotipi di HCV può contribuire, infatti, a trarre rilevanti informazioni sull'origine e la diffusione di HCV in Europa e in tutta l'area del Mediterraneo almeno nei tempi più recenti e permettere di spiegare la

diversa distribuzione geografica dei sottotipi migliorando le conoscenze sul virus e sulla sua circolazione nelle popolazioni.

3. MATERIALI E METODI

3.1 LA CASISTICA MONTENEGRINA

Questo lavoro di tesi è stato condotto su 162 campioni di sieri ottenuti da un pari numero di pazienti viventi in Montenegro, raccolti presso l'Istituto di Sanità Pubblica di Podgorica, tra il 2007 e il 2014. Tutti i pazienti erano affetti da epatite cronica di tipo C ed erano risultati positivi per anticorpi anti-HCV e HCV-RNA plasmatico. In relazione al gruppo a rischio, 98 pazienti erano tossicodipendenti (IVDUs *intra-venus drug users*) (60.5%), 37 soggetti dichiaravano diversi fattori di esposizione al virus tra cui, in particolare, le infezioni iatrogene che comprendevano gli interventi chirurgici, le trasfusioni di sangue e la dialisi che nel loro complesso costituiscono l'81.1% (n=30) di questo gruppo a rischio, mentre per 27 non erano riscontrabili fattori di rischio noti.

I dati relativi a tale popolazione sono riportati nella tabella 2.

Tabella 2: Caratteristiche della popolazione generale montenegrina.

| | Popolazione in studio n (%) |
|--|------------------------------------|
| Età mediana (range), anni | 53 (24 - 77) |
| Sesso | |
| Maschi | 123 (75.9) |
| Femmine | 39 (24.1) |
| Fattori di rischio | |
| Uso di droghe per via iniettiva (IVDUs) | 98 (60.5) |
| Esposizione iatrogena: | 30 (18.5) |
| - interventi chirurgici | 15 (50.0) |
| - dialisi | 9 (30.0) |
| - trasfusione | 6 (20.0) |
| Esposizione professionale (HCW) | 3 (1.9) |
| Trasmissione familiare | 2 (1.2) |
| Prostituzione | 1 (0.6) |
| Viaggi | 1 (0.6) |
| Non noti | 27 (16.7) |

Dei 162 pazienti 123 erano maschi (75.9%) e 39 erano femmine (24.1%) con un'età compresa tra 24 e 77 anni (età mediana 53) e appartenenti a diverse fasce d'età (Figura 10).

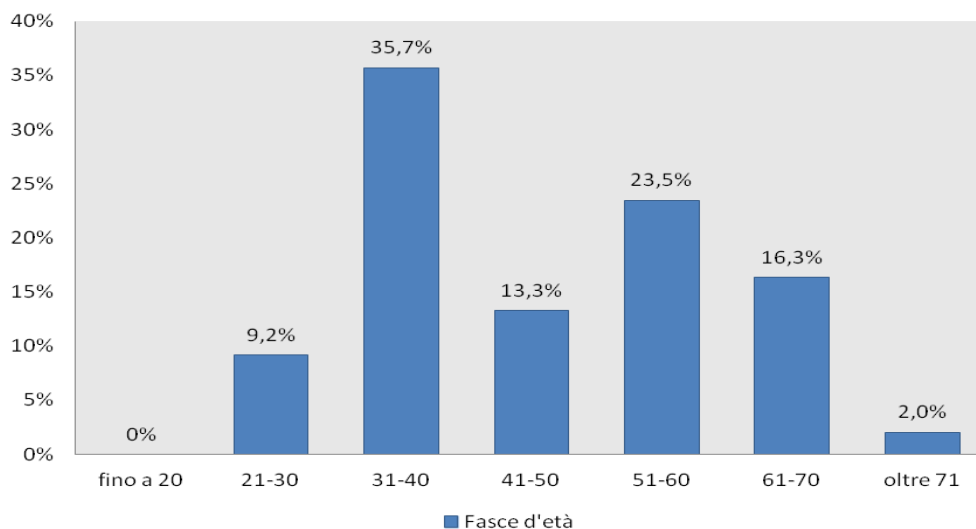


Figura 10: Distribuzione dei soggetti per fasce d'età.

3.2 LA CASISTICA ALBANESE

Sono stati inclusi nell'analisi 67 campioni di siero di pazienti albanesi provenienti dall'Istituto di Sanità Pubblica e dal Centro Nazionale Trasfusioni di sangue di Tirana, arruolati tra 2005 e 2007 e tra il 2008 e 2014, affetti da epatite C e positivi al test ELISA.

I pazienti avevano un'età compresa tra 18 e 55 anni (età mediana=37 anni) e comprendevano 12 tossicodipendenti (IVDUs) (17.9%) mentre per 55 non era noto il fattore d'esposizione (popolazione generale) (82.1%). I dati relativi alla popolazione albanese in studio sono riportati nella tabella seguente:

Tabella 3: Caratteristiche della popolazione albanese.

| | Popolazione in studio n (%) |
|-----------------------------------|-----------------------------|
| Età' mediana (range), anni | 37 (18-55) |
| Sesso | |
| Maschi | 47 (70.1) |
| Femmine | 13 (19.4) |
| Non noti | 7 (10.4) |
| Gruppi a rischio | |
| IVDUs | 12 (17.9) |
| Popolazione generale | 55 (82.1) |

3.3 METODI

3.3.1 Modalità di raccolta, preparazione e conservazione dei campioni

Il prelievo del sangue periferico e lo stoccaggio del siero sono stati realizzati presso l'Istituto di Sanità Pubblica in Montenegro e in quello di Tirana.

I prelievi di sangue periferico sono stati effettuati in sterilità mediante sistema *vacutainer*, utilizzando provette dotate di cuscino separatore, senza anticoagulante.

Il sangue (7 ml) è stato quindi centrifugato a 1900 rpm per 10 minuti a 4°C. Il siero così ottenuto è stato suddiviso in aliquote da 1 ml e quindi congelato a -20°C.

Successivamente, i campioni di siero sono stati trasportati in ghiaccio secco presso il Dipartimento di Scienze Biomediche e Cliniche, Ospedale L. Sacco, Laboratorio di Malattie Infettive e conservati a -80° C fino al momento dell'utilizzo.

3.3.2 Estrazione di RNA virale da plasma

L'estrazione dell'RNA è stata condotta a partire da siero di pazienti infetti mediante il *kit NucleoMag 96 Virus (Machery-Nagel, Germany)* e l'estrattore automatico *King Fisher mL Magnetic Particle Processor (Thermo Fusher Scientific, Inc., NYSE:TMO)*.

La procedura di estrazione si basa sul reversibile assorbimento degli acidi nucleici tramite biglie magnetiche sottoposte ad opportune condizioni di reazione e la loro successiva purificazione per rimuovere i contaminanti e i sali utilizzando adeguate soluzioni tampone.

Le fasi della procedura di estrazione, eseguita sui campioni in esame, sono le seguenti: 200 µl di siero sono aliquotati in una provetta sterile da 1.5 ml con un ugual volume di tampone di lisi (*Lysis Buffer MVI*), 20 µl di Proteinasi K (20 mg/ml) e 4 µl di RNA *carrier*. Dopo aver miscelato, si incuba nel termoblocco per 10 minuti a 56°C in agitazione. La proteinasi K digerirà le proteine plasmatiche, mentre il *Lysis Buffer*, contenente Guanidina tiocianato, contribuirà alla denaturazione delle proteine.

Dopo la lisi, si aggiungono, ad ogni campione, 30 µl di *Resuspended v-Beads* e 600 µl di *Binding Buffer MV2*. Contemporaneamente si preparano le *strip*, una per ogni campione, in cui si dispensano rispettivamente:

- il campione lisato e le *beads* nel pozzetto A
- 500 µl di *Wash Buffer MV3* nel pozzetto B
- 500 µl di *Wash Buffer MV4* nel pozzetto C
- 550 µl di *Wash Buffer MV5* nel pozzetto D

- 50 µl di acqua *RNasi free* nel pozzetto E

La piastra con le *strip* viene inserita nel vano dell'estrattore.

L'RNA nel pozzetto A viene legato alle *beads* e l'attrazione delle stesse ai magneti permette l'agitazione del campione e la sua purificazione mediante i *buffer* di lavaggio presenti negli altri pozzetti. Infine l'eluato, recuperato alla fine della procedura, viene raccolto nel pozzetto E e conservato a -80°C fino al momento dell'utilizzo. Per escludere la possibilità di contaminazioni tra i diversi campioni, ogni estrazione di RNA dei pazienti in studio è avvenuta in parallelo con identico materiale biologico proveniente da donatori sani (controlli negativi di estrazione). Inoltre, per escludere l'eventuale presenza di inibitori o un errore durante il procedimento di estrazione, è stato sempre utilizzato anche un controllo positivo noto.

3.3.3 Retrotrascrizione del gene NS5B

Utilizzando i reagenti contenuti nel modulo per la retrotrascrizione, si prepara la *mix* per la RT-PCR (*SuperScript III First-Strand Synthesis SuperMix for qRT-PCR*, Invitrogen):

| Reagente | Volume (per campione) |
|---------------|--------------------------|
| RT Mix 2x | 10 µl |
| RT Enzyme Mix | 2 µl |
| RNA | 8 µl |
| Totale | 20 µl |

Il funzionamento del modulo è basato sulla trascrittasi inversa del virus della leucemia murina di *Maloney* (*Maloney Murine Leukemia Virus*, MuLV), che estende un *primer* specifico per l'RNA virale contenuto nel buffer di reazione, sintetizzando una molecola di cDNA a doppio filamento.

Il kit contiene dei *random primers* che permettono così di retrotrascrivere tutto il genoma di HCV che può essere poi sottoposto ad ulteriore amplificazione per la regione di interesse.

Preparata la *mix* di reazione si inseriscono i campioni in un termociclatore, impostato con il seguente programma:

| T°C | t | Processo |
|------------|----------|-----------------------|
| 25°C | 10' | Innesco reazione |
| 50°C | 30' | Retrotrascrizione |
| 85°C | 5' | Inattivazione MuLV RT |
| 4°C | ∞ | Conservazione |

Al termine della reazione si aggiunge 1 µl di RNasi H di E.Coli e si incuba a 37°C per venti minuti per eliminare l'RNA in eccesso. Il cDNA così ottenuto viene poi sottoposto a due successive amplificazioni.

3.3.4 Amplificazione della regione NS5B mediante NESTED-PCR

In seguito alla retrotrascrizione si effettua l'amplificazione della regione desiderata tramite *nested-PCR*. Tale tecnica è una variazione della PCR classica e prevede due reazioni di amplificazione successive utilizzando due diverse coppie di *primers*.

Oltre ai campioni estratti ed al controllo negativo si aggiunge anche un controllo positivo, precedentemente estratto.

Nel presente lavoro è stata utilizzata una *nested-PCR* che ha previsto l'amplificazione di un frammento della regione genomica NS5B di circa 376 nt codificante per la polimerasi virale. Per l'amplificazione dei tre frammenti è stato utilizzato il kit: *GoTaq® Hot Start Polymerase (Promega, Madison, WI, USA)*.

Questa reazione di nested-PCR si avvale di due successivi cicli (step) di amplificazione. I *primers* utilizzati sono stati precedentemente descritti in letteratura⁷⁶.

Nel primo ciclo di amplificazione si utilizza una singola coppia di *primer* più esterni (ENO2/ENO4) che ci permette di ottenere un lungo frammento di circa 395 bp, presente nella regione NS5B.

ENO2_NS5BF1 (posizione 8245-8275)

5'TGGG(GC)TT(CT)(GT)C(GC)TATGA(CY)C(CT)CG(AC)TG(CT)TTTGA-3'

ENO4_NS5BR1 (posizione 8616-8640)

5'A(AG)TACCT(AG)GTCATAGCCTCCGTGAA 3'

Nella seguente PCR viene preparata una *mix* usando i corrispettivi *primer* esterni:

| Reagente | Volume (per campione) |
|-----------------------------|-----------------------|
| 5x | 10,0 μ L |
| dNTPs | 4,0 μ L |
| MgCl ₂ | 3,0 μ L |
| ENO2 (50 pmoli/ μ l) | 0,5 μ L |
| ENO4 (50 pmoli/ μ l) | 0,5 μ L |
| Go Taq Hot Start | 0,2 μ L |
| H ₂ O | 21,8 μ L |
| Volume finale | 40 μ L |

A questa *mix* si aggiungono 10 μ L di cDNA virale estratto ed il tutto viene sottoposto ad amplificazione in un termociclatore così impostato:

| T | t | n° cicli | Processo |
|------|----------|----------|---------------------------------------|
| 94°C | 2' | 1 | Attivazione Taq |
| 94°C | 15'' | | Denaturazione DNA |
| 55°C | 30'' | 25 | <i>Annealing</i> dei <i>primer</i> |
| 72°C | 45'' | | Estensione |
| 72°C | 7' | 1 | Estensione finale |
| 4°C | ∞ | | Conservazione |

Il secondo ciclo di amplificazione prevede poi l'uso di una coppia di *primer* (S3/A5) interni al frammento già amplificato. Il prodotto del secondo *step* è costituito da un frammento di 376pb.

NS5B S3 forward (position 8256-8281)

5'-TATGATACCCGCTGCTTTGACTCCAC-3'

NS5B A5 reverse (position 8611-8632)

5'- GTCATAGCCTCCGTGAAGGCTC-3.

Per i *primers* interni viene preparata la seguente *mix* di reazione:

| Reagente | Volume (per campione) |
|------------------------|--------------------------|
| 5x | 10,0 μ L |
| dNTP | 4,0 μ L |
| MgCl ₂ | 3,0 μ L |
| S3 (50 pmoli/ μ l) | 0,5 μ L |
| A5 (50 pmoli/ μ l) | 0,5 μ L |
| Go TAQ Hot Start | 0,2 μ L |
| H ₂ O | 26,8 μ L |
| Volume finale | 45 μ L |

Per il secondo *step* vengono utilizzati 5 μ l del prodotto del primo *step* per ciascuna *mix*.

L'amplificazione viene condotta nel termociclatore alle seguenti condizioni:

| T | t | n° cicli | Processo |
|------|----------|----------|-----------------------------|
| 94°C | 2' | 1 | Attivazione Taq |
| 94°C | 15'' | 25 | Denaturazione DNA |
| 55°C | 30'' | | <i>Annealing dei primer</i> |
| 72°C | 45'' | | Estensione |
| 72°C | 7' | 1 | Estensione finale |
| 4°C | ∞ | | Conservazione |

3.3.5 Rilevamento dei prodotti di PCR

Il rilevamento dei prodotti di amplificazione è stato ottenuto mediante corsa elettroforetica su gel di agarosio al 1,5% in tampone TAE 1X.

La migrazione del DNA su gel di agarosio dipende dalla lunghezza del frammento di DNA analizzato rispetto alla concentrazione dell'agarosio stesso.

Le bande prodotte in seguito alla migrazione del DNA vengono esaminate esponendo il gel alla luce ultravioletta; per questo motivo nel gel stesso viene incorporato un

colorante fluorescente, etidio bromuro (0,01 $\mu\text{g}/\mu\text{l}$), che si intercala tra le basi dell'acido nucleico. È possibile riconoscere l'amplificato di nostro interesse tramite il confronto con un opportuno standard di riferimento o *marker* a concentrazione nota, che viene caricato in ogni corsa elettroforetica contemporaneamente ai campioni in esame.

Il confronto con il *marker* di riferimento consente anche l'eventuale quantificazione dell'amplificato in esame in base all'intensità della banda

3.3.6 Purificazione dei prodotti di PCR

La purificazione dei prodotti di PCR è eseguita impiegando il kit di purificazione *QIAquick* (Qiagen). La purificazione dei prodotti di PCR permette di rimuovere efficacemente tamponi di reazione, enzimi, *primer* e dNTPs (marcati o non). Un limite di questo kit è rappresentato dalle dimensioni del frammento, che devono essere comprese tra 100 bp e 10 Kb.

Prima di iniziare è necessario porre attenzione ai seguenti punti:

- il buffer PE, con il quale viene eluito il DNA dopo la purificazione, deve essere addizionato di etanolo al 96-100%,
- tutte le centrifugazioni previste dal protocollo devono essere impostate ad una velocità di 13.000 rpm,

La purificazione avviene tramite la seguente procedura:

- al secondo step di PCR si aggiunge un volume 5 volte maggiore di buffer BP, si miscela attentamente con la pipetta fino ad ottenere una soluzione omogenea,
- si posiziona una colonnina *QIAquick* in una provetta di raccolta da 2 mL,
- si inserisce il campione nella colonnina *QIAquick* e si centrifuga per 30-60 secondi (13000 rpm),
- si scarta l'eluito e si riposiziona la colonna nella medesima provetta,
- per lavare, si aggiungono 700 μl di buffer PE nella colonnina e si centrifuga per 30-60 secondi,
- si elimina l'eluito e si riposiziona la colonnina nella medesima provetta,
- si centrifuga ulteriormente per 1 minuto,
- si posiziona la colonnina in una nuova provetta da 1.5mL,
- per eluire il DNA, si aggiungono 50 μL di buffer EB o di acqua distillata direttamente sopra la membrana della colonnina e si centrifuga per 1 minuto.
- si eliminano le colonnine. Il DNA purificato deve essere conservato a 4°C.

3.3.7 Sequenziamento della regione genomica amplificata

3.3.7.1 La reazione di sequenza

Il sequenziamento dei prodotti di PCR è condotto con la tecnica del *cycle sequencing*; in questa variante della PCR, il DNA viene amplificato con una DNA polimerasi termostabile a partire da un unico *primer* di sequenza, in presenza di dideossiribonucleotidi (ddNTPs) marcati con fluorocromi.

Per il sequenziamento dei campioni è stato utilizzato il Kit “ABI PRISM BigDyeTM Terminator Cycle Sequencing Ready Reaction”.

Il Kit è caratterizzato dalla presenza di 4 dideossiribonucleotidi terminatori marcati con quattro diversi fluorocromi, le molecole R6G, R110, TAMRA, ROX.

I dideossiribonucleotidi sono definiti terminatori in quanto mancano del gruppo idrossilico in 3', coinvolto nella formazione del legame fosfodiesterico tra nucleotidi che costituiscono le molecole di DNA. Pertanto, la loro incorporazione al posto dei normali deossiribonucleotidi (dNTPs) nella catena nascente del DNA, determina l'arresto della reazione di polimerizzazione.

Il Kit contiene una miscela di diversi dideossinucleotidi terminatori marcati, denominata *Terminator Ready Reaction Mix*, composta da: ddATP marcato 1,58 mM, ddTTP marcato 94,74 mM, ddGTP marcato 0,42 Mm, ddCTP marcato 47,37 mM, dTTP 15,79 mM, Tris-HCl 68,42 mM pH 9.1, (NH₄)₂SO₄ 4,21 mM, enzima AmpliTaq FS DNA-polymerase 0,42 unità/mL.

Il rapporto di concentrazione tra ddNTPs e dNTPs è calcolato in modo che la polimerasi termini la catena nascente in ogni posizione in cui il nucleotide può essere inserito, formando così una serie di frammenti sfasati e marcati in 3', che condividono tutti l'estremità del *primer*.

A differenza di una normale PCR, necessita l'utilizzo di un solo *primer*, determinando l'accumulo di prodotti di reazione in maniera lineare e non esponenziale, al termine della reazione ci sarà la sovrabbondanza di un filamento rispetto all'altro in modo tale che la riassociazione tra i filamenti complementari non possa avvenire.

Altra componente peculiare è costituita dall'enzima termostabile *AmpliTaq[®] FS DNA-polymerase* che ha la caratteristica di svolgere l'attività di polimerizzazione del DNA ad alte temperature, minimizzando i problemi causati dalla formazione di strutture secondarie e riducendo la formazione di legami tra *primer* e sequenze non specifiche.

Inoltre, l'enzima resiste alla denaturazione quando la miscela di sequenziamento è incubata alla temperatura di 96°C per denaturare il DNA templato.

Inoltre l'AmpliTaQ DNA polimerasi FS è una variante della consueta DNA polimerasi di *Thermophilus aquaticus*, in quanto contiene una mutazione puntiforme a livello di un sito attivo; tale mutazione comporta una ridotta capacità di discriminazione nei confronti dei dideossinucleotidi, che vengono perciò incorporati con maggior facilità.

La reazione di sequenza è stata condotta secondo il protocollo indicato dalla ditta produttrice, effettuando su ogni campione 2 diverse reazioni di sequenza, una in presenza del *primer forward* (senso S3) e l'altra con il *primer reverse* (antisenso A5), per poter verificare l'attendibilità del successivo sequenziamento.

La miscela di reazione si compone di 3,2 µl di *primer* (1pmole/ µl), 1,1 µl di BigDye (miscela di reazione), 1-2 µl di DNA purificato e diluito, stimato in base all'intensità rilevata su gel, e acqua fino ad arrivare ad un volume finale di 10 µl.

La reazione di sequenza viene condotta utilizzando un termociclatore alle seguenti condizioni:

| <i>T</i> ° | <i>n</i> °cicli | <i>t</i> | <i>Processo</i> |
|------------|-----------------|----------|-------------------------------|
| 96°C | 1 | 1' | Attivazione <i>BigDye</i> |
| 96°C | 25 | 10" | Denaturazione DNA |
| 50°C | | 5" | Legame dei <i>primers</i> |
| 60°C | | 4' | Estensione dei <i>primers</i> |
| 4°C | | ∞ | Conservazione |

3.3.7.2 Purificazione dei prodotti di sequenza

La purificazione dei prodotti di sequenza si effettua utilizzando la precipitazione in etanolo con sodio acetato.

I campioni da purificare vengono miscelati con 50 µl di etanolo al 100% e 2 µL di sodio acetato (2 M, pH = 4), centrifugati per 30 minuti a 14.000 rpm

Rimosso il surnatante, si aggiungono 250 µl di etanolo al 70% e si centrifuga per 5 minuti a 14.000 rpm.

Si elimina l'etanolo, i *pellet* sono portati a secco incubando a 95°C per 3 minuti circa.

I prodotti di sequenza così ottenuti sono purificati da possibili contaminanti, *primers*, dNTPs non incorporati e la DNA-polimerasi in eccesso che potrebbero disturbare nelle fasi successive, in questo modo sono pronti per procedere al sequenziamento.

3.3.7.3 Sequenziamento della regione NS5B

Per il sequenziamento è stato utilizzato il sequenziatore automatico *ABI Prism 3100* (*Applied Biosystems*), basato sulla tecnologia dell'elettroforesi capillare. La parte più importante di questo sistema è una unità intercambiabile composta da sedici capillari in silice fusa; questi, una volta riempiti di polimero, permettono la separazione elettroforetica dei frammenti di DNA marcati (Figura 11).

Dopo aver precedentemente risospeso il campione in 10 μL di acqua bidistillata, 2 μL si aggiungono a 8 μL di una soluzione denaturante di formammide.

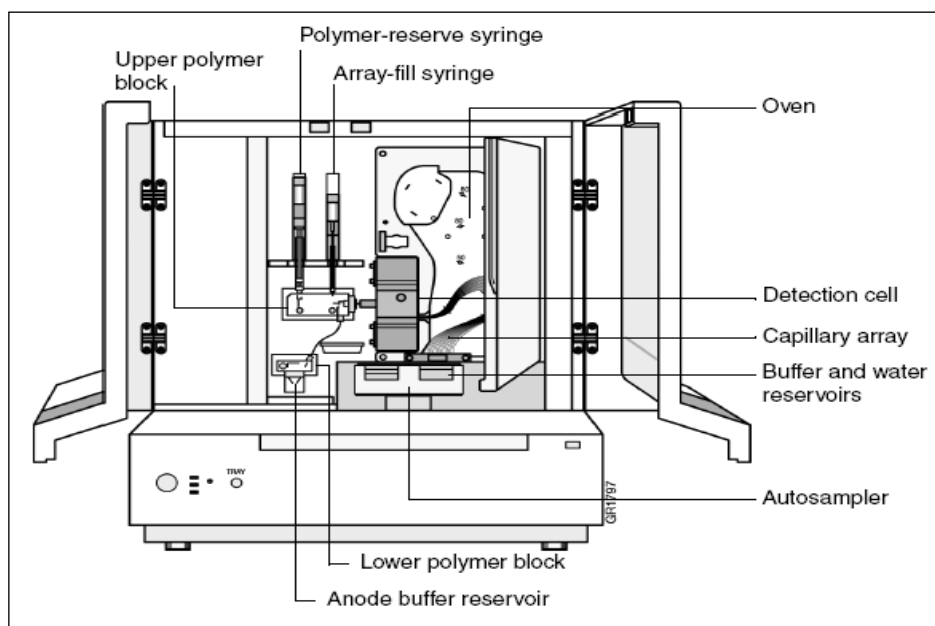


Figura 11: Sequenziatore automatico *ABI Prism 3100*.

I campioni vengono caricati in piastre da novantasei pozzetti, da cui sono prelevati, in gruppi di sedici, e spinti tramite iniezione elettrocinetica nei capillari; qui un circuito ad alto voltaggio fa migrare i frammenti di DNA, carichi negativamente, verso l'anodo.

La presenza del polimero all'interno dei capillari induce un rallentamento nella corsa delle molecole di DNA, proporzionale alla loro lunghezza; in tal modo queste si separano secondo la loro dimensione e, spinte attraverso una camera di conta, vengono colpite da un laser che ne eccita i marcatori. La fluorescenza viene rilevata da un

dispositivo ad accoppiamento di carica (CCD) che trasforma i segnali luminosi in forma digitale, trasmettendo l'elettroferogramma ottenuto ad un *Personal Computer*, collegato direttamente all'apparecchio. Ogni corsa ha una durata di 1,4 ore impostando il sequenziatore con i seguenti valori: voltaggio 12,2 KV, corrente 67-77 mA, laser 15 mW e temperatura 50°C.

3.3.7.4 Analisi dei dati di sequenza

L'analisi dei dati di sequenza si esegue utilizzando il programma BioEdit (ver. 7.2.5)⁷⁷. Le sequenze della regione NS5B dei campioni vengono allineate tra loro insieme ad un numero variabile di sequenze di riferimento, ottenute dalle principali banche dati, quali *GenBank*, presso il *National Center for Biotechnology Information* (USA). Le estremità vengono eliminate al fine di ottenere frammenti della medesima lunghezza e l'allineamento viene effettuato utilizzando l'algoritmo *ClustalX* contenuto nel programma, intervenendo con un *editing* manuale quando necessario. In questo modo vengono costruiti i *data set* di sequenze utilizzati per le successive analisi di tipizzazione.

3.3.8 Analisi filogenetica

L'analisi filogenetica è volta a studiare le distanze evolutive esistenti tra le sequenze in esame, tali distanze sono rappresentate attraverso alberi filogenetici, che costituiscono dei grafici bidimensionali composti da nodi connessi tra loro tramite rami. I nodi terminali rappresentano i *taxa* attuali, mentre i nodi interni rappresentano quelli ancestrali.

Le relazioni filogenetiche vengono definite attraverso la topologia dell'albero: la distribuzione dei nodi e la lunghezza dei rami forniscono indicazioni sulla divergenza evolutiva esistente tra i diversi *taxa*, permettendo di individuare gruppi monofiletici e linee evolutive.

L'analisi filogenetica rappresenta, oggi, l'unico metodo affidabile per la tipizzazione delle sequenze, siano esse sottotipi puri o ricombinanti; permette inoltre di esaminare le relazioni genetiche all'interno di un gruppo di sequenze o tra virus provenienti da diverse aree geografiche, fornendo indicazioni epidemiologiche sull'origine e la propagazione dei virus in una popolazione.

3.3.8.1 Analisi filogenetica delle sequenze

L'analisi filogenetica è stata eseguita su sequenze lunghe circa 318 nucleotidi e comprendenti la regione NS5B. Per la caratterizzazione del genotipo degli isolati montenegrini e albanesi, abbiamo utilizzato 67 sequenze di riferimento, specifiche per ogni sottotipo di HCV.

Dopo l'analisi mediante *Bioedit* il *data set* viene analizzato utilizzando i programmi SEQBOOT, DNADIST, NEIGHBOR e CONSENSE contenuti nel pacchetto *software* PHYLIP⁷⁸ (ver. 3.67).

In particolare per la scelta del modello di calcolo delle distanze genetiche, abbiamo effettuato preventivamente la stima del rapporto transizioni (mutazioni quali purina-purina o pirimidina-pirimidina) rispetto alle trasversioni (mutazioni del tipo pirimidina-purina o purina-pirimidina) con il programma *JmodelTest v.2.1.7*⁷⁹ utilizzato per calcolare le distanze genetiche tra le sequenze sottomesse mediante opportuni modelli evolutivisti di differente complessità.

Le distanze genetiche sono state espresse mediante medie (ed errore standard) in termini di n. sostituzioni/sito.

La filogenesi delle sequenze in studio è stata ricostruita utilizzando due differenti approcci: il metodo *Neighbor-Joining* (NJ), incluso nel pacchetto *Phylip*, e il metodo *Maximum-Likelihood* (ML), utilizzando un nuovo algoritmo "hill-climbing", incluso nel server *Phyml 3.0* (<http://www.atgc-montpellier.fr/phyml/>).

L'algoritmo *Neighbor-Joining*, consente di generare alberi senza radice, basandosi sulle matrici di distanza con un'elevata efficienza e rapidità, dando una lunghezza dei rami proporzionale alla distanza genetica tra le unità tassonomiche.

Nell'approccio *Maximum-Likelihood* viene calcolata la probabilità di osservare i dati, dato un albero filogenetico e un modello che descrive il processo di evoluzione, e tra tutti gli alberi testati, viene scelto come albero ML finale quello che presenta il maggior valore di probabilità.

Il software *Figtree v.1.4.2*, disponibile gratuitamente sul web (<http://tree.bio.ed.ac.uk/software/figtree/>), è stato usato per la riproduzione grafica degli alberi.

La significatività degli alberi è stata stimata mediante il metodo *Bootstrap*, che consiste in un ricampionamento del multiallineamento estraendone a caso i siti. Sono stati

considerati significativi i nodi interni supportati da un *bootstrap* maggiore del 70% (dopo 1000 replicati in NJ o 200 in Phylm).

La caratterizzazione molecolare del genotipo/sottotipo di HCV è stata effettuata mediante l'analisi filogenetica, confrontando ciascuna sequenza del gene NS5B in studio con opportune sequenze di riferimento.

3.3.8.2 Preparazione dei data set per l'analisi filogenetica

Gli alberi filogenetici, i parametri dei modelli, i valori di *evolutionary rate* (ER) e i modelli di crescita di popolazione sono stati stimati mediante un metodo Bayesiano, basato sull'algoritmo *Markov Chain Monte Carlo* (MCMC), incluso nel software Beast v. 1.8.0⁸⁰.

A tal proposito abbiamo usato tre *data set* differenti composti da 271 sequenze di sottotipo 3a, 227 sequenze del sottotipo 1b e 106 sequenze di sottotipo 4d.

L'analisi filogenetica può essere condotta sulla base di un modello temporale definito orologio molecolare^{81, 82, 83} (*Molecular Clock*). Tale modello si basa sull'assunto che le mutazioni si accumulino nel tempo in modo pressoché costante e permette di delineare le relazioni intercorrenti tra le distanze genetiche e il tempo, permettendo al tempo stesso di datare gli eventi di divergenza evolutiva⁸⁴.

I modelli più semplici prevedono un orologio molecolare più rigido (*strict clock model*), nel quale si assume che tutti i *lineage* evolvano alla stessa velocità; diversamente, nei modelli più complessi (*relaxed clock model*) le velocità evolutive variano nel tempo e tra i vari *lineage*, fornendo degli intervalli di variabilità intorno ad un valore medio della velocità evolutiva. In questo studio sono stati testati sia il modello *strict* che il modello *relaxed*, con una distribuzione log-normale non correlata. Tre modelli demografici parametrici (*constant population size*, *exponential*, *logistic growth*) e i modelli non parametrici *Bayesian skyline* e *GMRF Bayesian Skyride* sono stati confrontati con un test denominato *Bayes Factor* (BF). Sulla base del *Bayes Factor* (BF) ottenuto è stato scelto un orologio molecolare di tipo *rilassato*, in cui ogni *lineage* ha un *rate* evolutivo caratteristico e non correlato con quello dei *lineage* contigui. In accordo con Kass e Raftery, la forza dell'evidenza contro H0 viene valutata come segue: $2\ln BF < 2$ nessuna evidenza; tra 2 e 6 debole evidenza; tra 6 e 10 forte evidenza e >10 evidenza molto forte. Un valore di $2\ln BF$ negativo è indice di evidenza in favore di H0. Solo valori ≥ 6 sono considerati significativi⁸⁵.

Per l'analisi dei nostri *data set* è stato selezionato un modello di sostituzione nucleotidica opportuno; in particolare per il *data set* 1b il modello selezionato è stato GTR+G+I (*General-Time Reversible + Gamma distribution rate + proportion of Invariable sites*) mentre per i *data set* 4d e 3a è stato HKY+I+G (Hasegawa-Kishino-Yano)⁸⁶.

Le catene di Markov (MCMC) sono state fatte correre per almeno 50 milioni di generazioni e campionate ogni 5000 steps.

L'analisi ottenuta tramite Beast deve poi essere validata dal software TRACER⁸⁷, il quale verifica la stima dell'effettiva dimensione del campione (ESS, *effective sample size*). L'incertezza delle stime ottenute è stata indicata con intervalli di *Highest Posterior Density* del 95% (HPD).

L'albero datato mostrato è stato ottenuto selezionando l'albero con il più alto prodotto di probabilità a posteriori (*maximum clade credibility*), mediante il programma *Tree Annotator*. In particolare *clades* con una probabilità posteriore ≥ 0.7 vengono considerati significativi.

Le sequenze virali vengono datate in base al momento del prelievo e i rami dell'albero filogenetico hanno una lunghezza proporzionale al tempo. Questo tipo di approccio stima la distanza temporale dei rami ed il tempo dell'ancestore comune più recente (TMRCA, *time of the most recent common ancestor*).

Per quanto riguarda il *data set* 1b mancavano le date di campionamento delle sequenze e per questa ragione abbiamo utilizzato un ER già stimato in precedenza su un *data set* indipendente, secondo un approccio ormai consolidato (Huè *et al*). In particolare il valore di ER utilizzato corrisponde ad una stima ottenuta dal nostro gruppo di ricerca su una serie di campioni provenienti dalla Turchia, precedentemente pubblicato⁸⁸. È stata utilizzata una stima dell'ER pari a 0.84×10^{-3} *substitution/site/year* con un intervallo di credibilità del 95% tra $0.16-1.5 \times 10^{-3}$.

3.3.9 Analisi filodinamica

Al fine di delineare la storia epidemiologica/evolutiva del virus HCV nei due paesi in studio abbiamo selezionato dai *data set* di partenza 46 sequenze montenegrine e 8 sequenze albanesi di sottotipo 3a che sapevamo fossero IVDUs; le sequenze del *data set* 1b e 4d che rappresentano, rispettivamente, i sottotipi di HCV più frequenti nella popolazione.

La teoria della coalescenza si basa sulla correlazione tra il tempo in cui due *taxa* o due *lineage* hanno condiviso uno stesso ancestore e le dimensioni della popolazione in esame, rappresentando così un ottimo strumento per valutare le modificazioni di una popolazione virale nel tempo.

Per tracciare l'evoluzione dei *lineage* virali, da cui estrapolare gli eventi di coalescenza, è necessario disegnare uno *skyline plot*, che mostra le dimensioni della popolazione virale (numero delle infezioni effettive N sulle ordinate) e il suo andamento su scala temporale (tempo t sulle ascisse).

Questo tipo di approccio è alla base dei metodi Bayesiani e del programma Beast, il quale consente di stimare contemporaneamente la tipologia dell'albero filogenetico, la velocità evolutiva, la popolazione effettiva e i tMRCA dei nodi interni, a partire da un allineamento multiplo.

L'analisi Bayesiana è stata eseguita assumendo lo stesso modello di sostituzione e orologio molecolare impiegato per il *data set* completo.

Sono stati testati due modelli demografici parametrici (*constant population size*, *exponential growth*) e il modello non parametrico *Bayesian skyline*. Per l'analisi filodinamica condotta in questo studio, dopo aver testato tutti i modelli evolutivi, è stato scelto, sulla base del *Bayes Factor* (BF) ottenuto, il modello *relaxed skyline* (RelSky).

3.3.10 Analisi Filogeografica

L'approccio filogeografico si basa sull'esistenza di una relazione tra l'evoluzione molecolare di un organismo e la sua distribuzione spaziale. Questa relazione richiede normalmente milioni di anni di evoluzione perché possa essere verificata, ma nel caso di virus ad RNA, caratterizzati da una elevatissima velocità evolutiva e dalla produzione di milioni di generazioni virali in un tempo brevissimo, è possibile osservare tale relazione anche nell'intervallo di pochi anni che possono intercorrere tra due campionamenti. Sono attualmente disponibili diversi approcci filogenetici per inferire la distribuzione spaziale di un organismo sulla base della filogenesi dei suoi geni; in questo lavoro è stato utilizzato un approccio Bayesiano, implementato nel programma Beast 1.8.0. Tale approccio consente di ricostruire la località più probabile dei nodi interni esclusivamente sulla base della conoscenza del luogo di isolamento delle sequenze incluse nell'analisi. Il modello utilizzato è detto CTMC (*Continuous-Time Markov Chain*), e consente di raggruppare le sequenze in gruppi separati corrispondenti a località definite. La filogenesi di un virus, ricostruita in scala spazio-temporale,

permette quindi di stimare la sua diffusione geografica, in relazione al tempo reale, e di ricostruire quindi la sua storia epidemiologica. Gli alberi ottenuti vengono rappresentati con il *Maximum Clade Credibility (MCC) tree*, cioè l'albero con la più alta probabilità posteriore dopo un *burn-in* del 10%, usando il programma Tree Annotator, incluso nel pacchetto Beast. La località più probabile per ogni nodo viene evidenziata contrassegnando le ramificazioni dell'albero con diversi colori.

L'analisi filogeografica è stata condotta utilizzando in parallelo i tre *data set* delle sequenze 1b, 3a e 4d della nostra casistica in aggiunta a sequenze di riferimento estratte dai database *Los Alamos* e *GenBank* per le quali era disponibile l'anno del campionamento ed era nota l'area geografica.

Inoltre per ricostruzione filogeografica le sequenze 1b sono state sottoposte ad analisi mediante il programma *Spread* che riproduce visivamente la migrazione virale nelle regioni studiate.

3.3.11 Analisi Statistica

La significatività statistica è stata valutata mediante test OpenEpi. Il livello di significatività per l'introduzione delle variabili nel modello finale è stato fissato a $p=0,01$ ($p < 0,05$).

4- RISULTATI

4.1 ANALISI FILOGENETICA DELLE SEQUENZE MONTENEGRINE

Nella figura 12 è rappresentato l'albero di genotipizzazione ottenuto dall'analisi delle 162 sequenze del Montenegro e dei ceppi riferimento.

Tale albero ha mostrato 5 clade significativi corrispondenti a 5 sottotipi virali: 1a, 1b, 2a/c, 3a, 4d. I valori di *bootstrap* dei 5 clade sono compresi tra 800-1000.

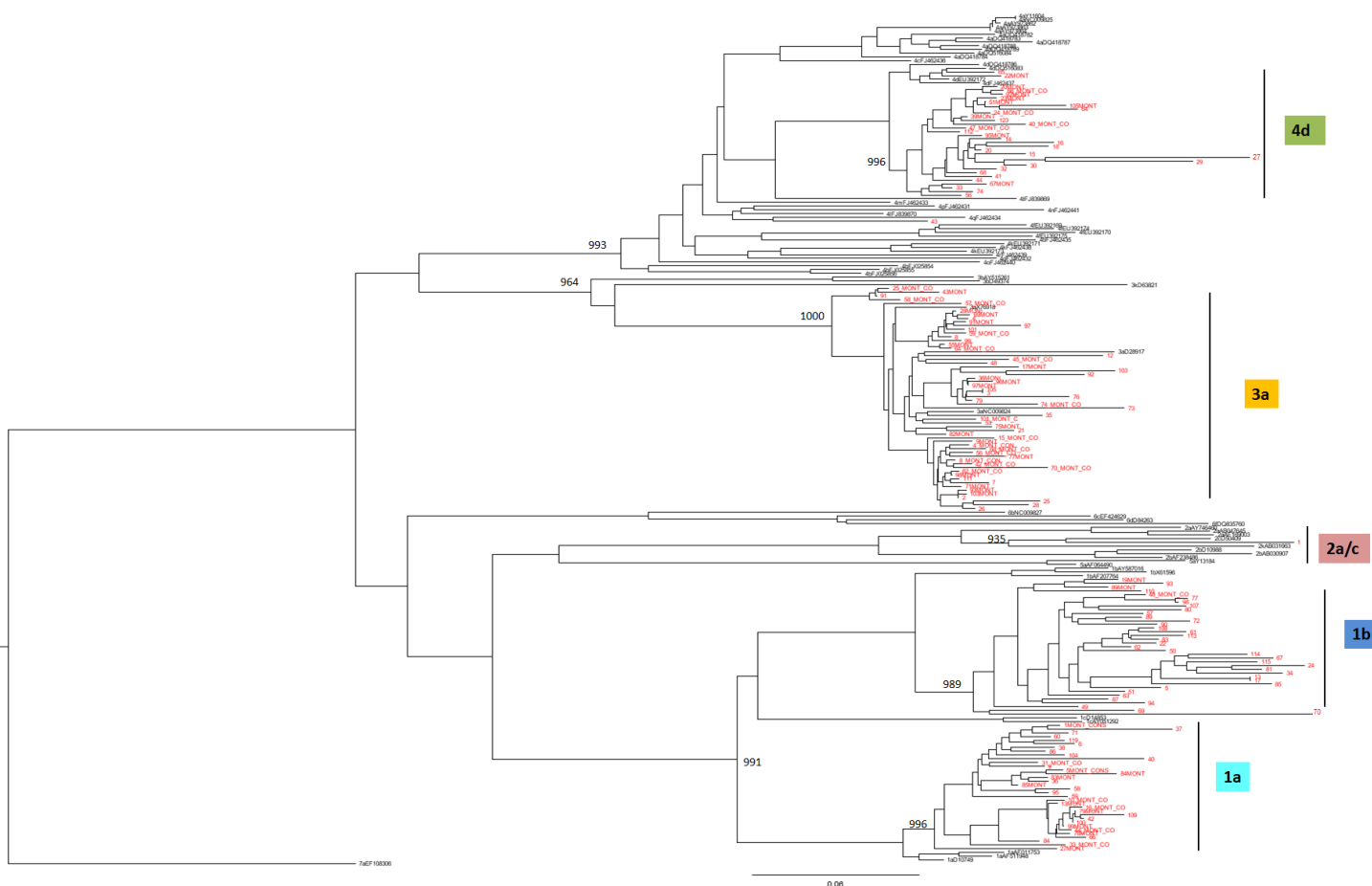


Figura 12: Albero NJ delle 162 sequenze montenegrine (indicate in rosso) allineate con 67 sequenze di riferimento rappresentative dei principali genotipi e sottotipi di HCV. I numeri sui rami indicano i valori di *bootstrap*. I clade comprendenti gli isolati montenegrini sono stati evidenziati in colore: 1a, azzurro; 1b, blu; 2a/c, rosa; 3a, arancione; 4d, verde

Sulla base di tale albero, 37 soggetti sono risultati di sottotipo 1b (22.8%), 34 di sottotipo 1a (21.0%), 57 di sottotipo 3a (35.2%), 32 di sottotipo 4d (19.8%), 1 di sottotipo 2c (0.6%), per un campione di genotipo 4 non è stata possibile l'assegnazione del sottotipo virale (0.6%).

4.1.1 Correlazione tra sottotipi di HCV e fattori di rischio nei soggetti montenegrini

Considerando i fattori di rischio nella popolazione montenegrina è stata analizzata la correlazione con i sottotipi virali identificati.

Tra i 64 pazienti non IVDUs (comprendenti 37 soggetti con fattore di rischio noto e 27 non noto), il sottotipo di HCV più frequente è risultato l'1b, presente nel 45% (n=29) dei pazienti, seguito dal sottotipo 4d presente nel 19% (n=12), dal sottotipo 3a nel 17% (n=11), dall'1a nel 16% (n=10), dal sottotipo 2c presente nell'1% (n=1), di un campione non era noto il sottotipo (n=1) (Figura 13).

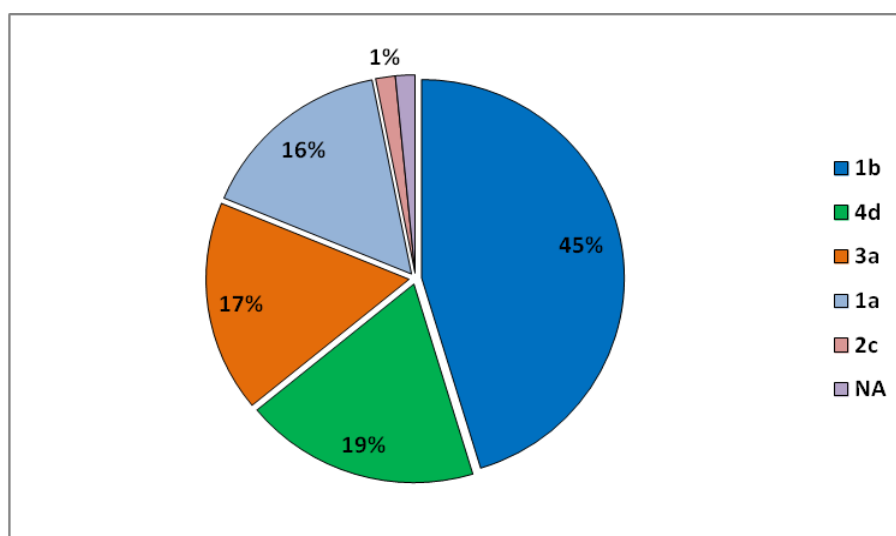


Figura 13: Distribuzione dei sottotipi di HCV nei soggetti non tossicodipendenti.

Il sottotipo di HCV più frequente nei 98 pazienti IVDUs, invece, è risultato essere il 3a, presente nel 47% (n=46) dei pazienti, seguito dal sottotipo 1a presente nel 25% (n=24), dal sottotipo 4d presente nel 20% (n=20) e dal sottotipo 1b presente nel 8% (n=8) (Figura 14).

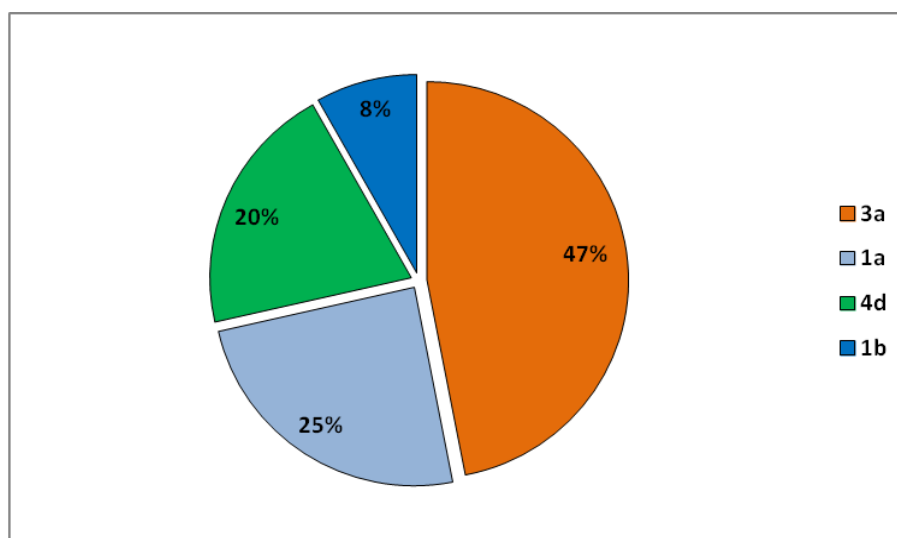


Figura 14: Distribuzione dei sottotipi di HCV negli IVUDs.

Per quanto riguarda i fattori di rischio noti nella popolazione, il genotipo 3a è risultato significativamente associato, nel 47%, a soggetti che facevano uso di droghe per via iniettiva ($p=0.0001$) e nel 6.6% in soggetti con fattore di rischio iatrogeno ($p=0.006$).

Il genotipo 1a è risultato più frequentemente presente tra i tossicodipendenti ma non in maniera significativa rispetto agli altri gruppi (24.5%). Al contrario, il sottotipo 1b è risultato significativamente associato alla trasmissione iatrogena (53.3%, $p=0.012$). Infine il sottotipo 4d è risultato equamente distribuito tra i vari gruppi a rischio. I dati sono riassunti nella tabella 4.

Tabella 4: Distribuzione sottotipi e fattori di rischio.

| | SOTTOTIPI n (%) | | | | | |
|-----------------------------------|-----------------|-----------|----------|-----------|-----------|----------|
| | 1a | 1b | 2c | 3a | 4d | NA |
| FATTORI DI RISCHIO (n) | | | | | | |
| IVUDs (98) | 24 (24.5) | 8 (8) | - | 46 (47) | 20 (20) | - |
| Esposizione iatrogena (30) | 4 (13.3) | 16 (53.3) | 1 (3.3) | 2 (6.6) | 6 (20) | 1 (3.3) |
| HCW (3) | 1 (33.3) | 2 (66.6) | - | - | - | - |
| Prostituzione (1) | - | 1 (100) | - | - | - | - |
| Trasmissione familiare (2) | - | 2 (100) | - | - | - | - |
| Viaggi (1) | 1 (100) | - | - | - | - | - |
| Non noto (27) | 4 (14.8) | 8 (29.6) | - | 9 (33.3) | 6 (22.2) | - |
| Totale | 34 | 37 | 1 | 57 | 32 | 1 |

4.1.2 Caratterizzazione dei sottotipi di HCV per sesso e per età nei soggetti montenegrini

Escludendo i soggetti tossicodipendenti che presentavano caratteristiche peculiari (netta prevalenza del genere maschile e giovane età), è stata valutata la distribuzione dei sottotipi in relazione al sesso e all'età nella restante popolazione. Per quanto riguarda il sesso, i sottotipi 1a, 1b, 3a, e 4d risultano maggiormente presenti nei maschi rispetto alle femmine, invece per il sottotipo 2c è presente solo una femmina (Figura 15).

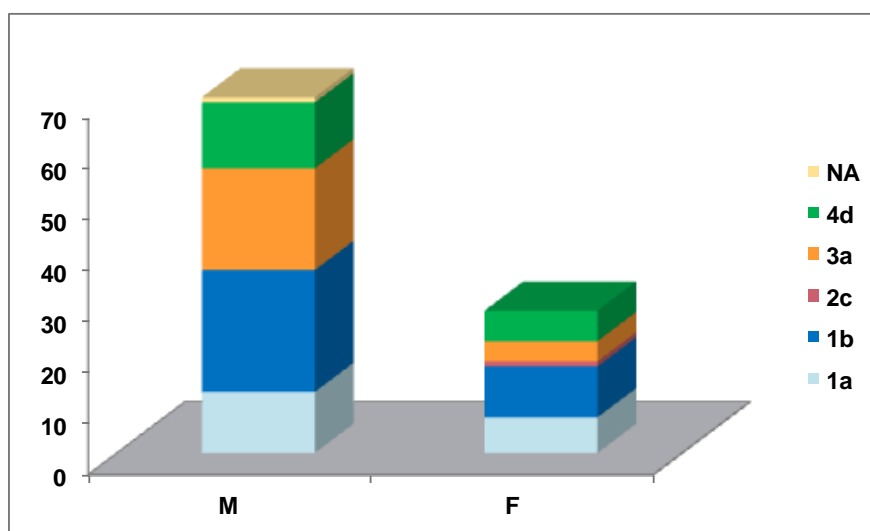


Figura 15: Distribuzione tra i sessi.

Riguardo l'età dei pazienti, la popolazione è stata suddivisa in 4 gruppi: 24-34 anni, 35-45 anni, 46-57 anni, 58-77 anni. Da tale analisi si è potuto osservare come il sottotipo 1b risulti maggiormente presente nella fascia d'età più avanzata 58-77 anni con un trend crescente dalla fasce più giovani a quelle più anziane (58.7%, $p=0.005$). I valori sono riportati nella tabella 5 e nella figura 16.

Tabella 5: Distribuzione sottotipi e fasce d'età.

| | SOTTOTIPI n (%) | | | | |
|---------------------|-----------------|-----------|----------|----------|---------|
| | 1a | 1b | 3a | 4d | 2c |
| FASCE D'ETA' | | | | | |
| 24-34 anni | 3 (30) | 2 (6.9) | 3 (27.3) | - | - |
| 35-45 anni | 1 (10) | 5 (17.2) | 1 (9) | 2 (16.6) | - |
| 46-56 anni | 4 (40) | 5 (17.2) | 5 (45.5) | 5 (50) | 1 (100) |
| 57-77 anni | 2 (20) | 17 (58.7) | 2 (18.2) | 4 (33.4) | - |
| Totale | 10 | 29 | 11 | 12 | 1 |

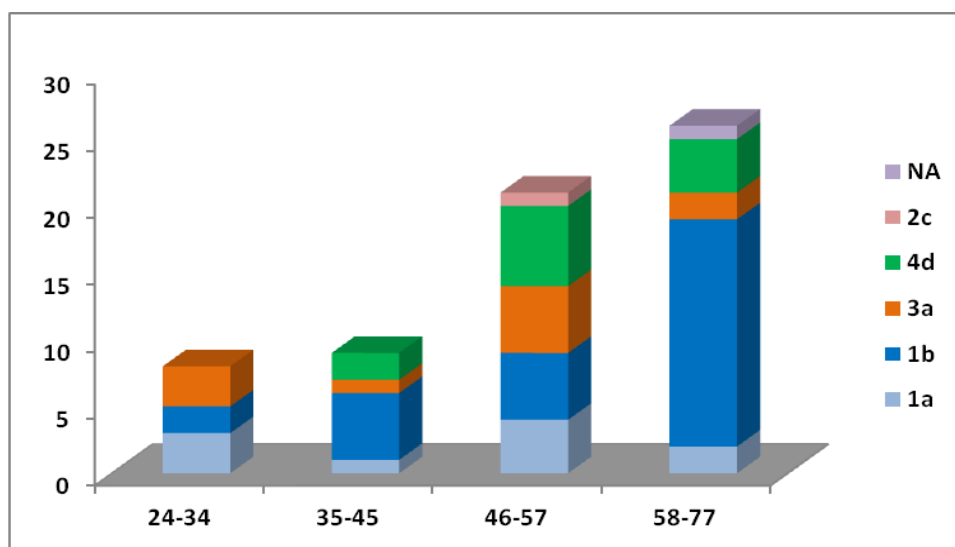


Figura 16: Distribuzione sottotipi e fasce d'età.

4.2 ANALISI FILOGENETICA DELLE SEQUENZE ALBANESE

La figura 17 rappresenta l'albero di genotipizzazione ottenuto dall'analisi delle 67 sequenze albanesi e dei ceppi di riferimento. Tale albero ha mostrato 5 clade significativi corrispondenti 5 sottotipi virali: 1a, 1b, 2a/c, 3a, 4. I valori di *bootstrap* dei 5 clade sono compresi tra 800-1000.

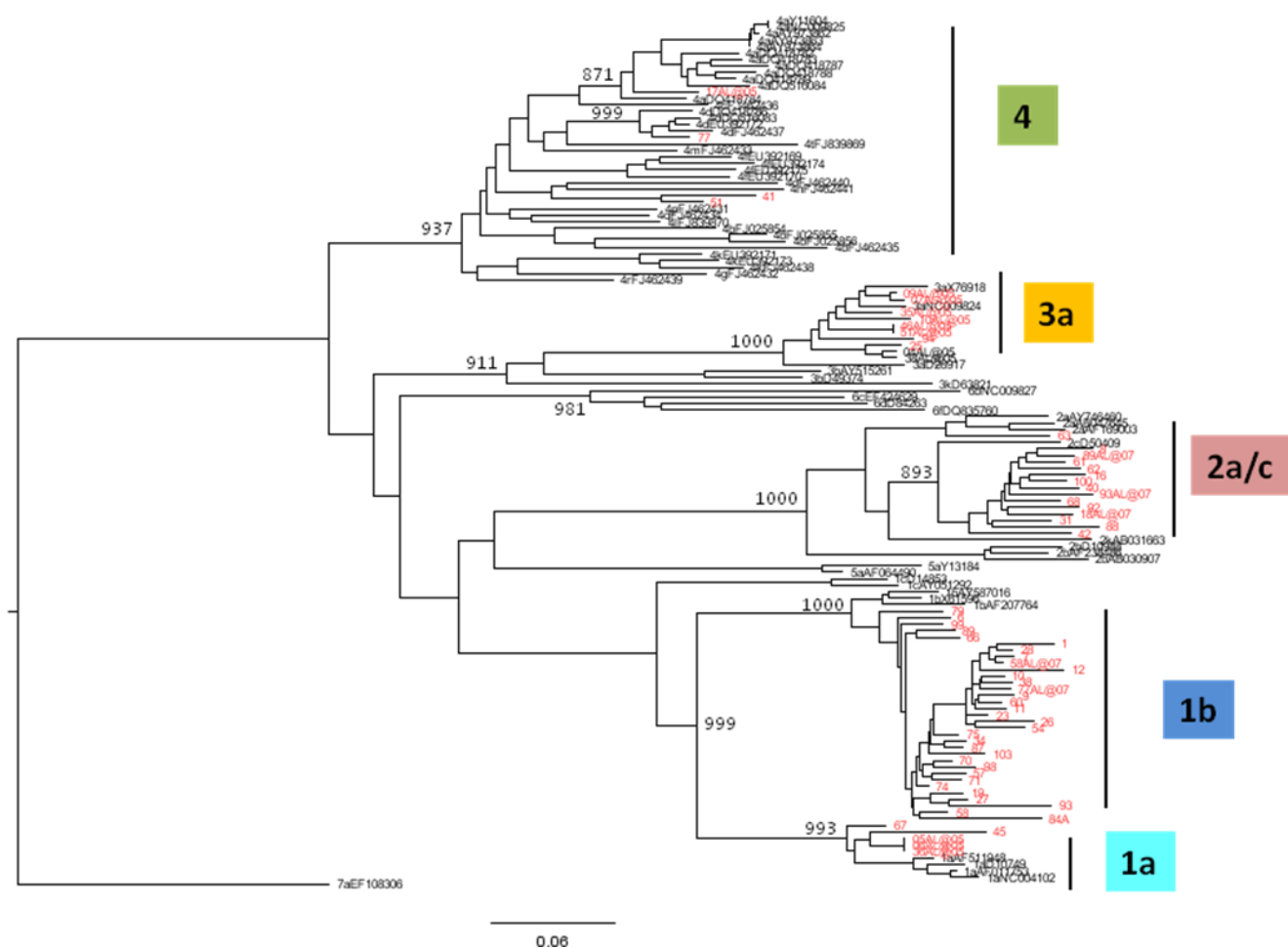


Figura 17: Albero NJ delle 67 sequenze albanesi (in rosso) allineate con 67 sequenze di riferimento rappresentative dei principali genotipi e sottotipi di HCV. I numeri sui rami indicano i valori di *bootstrap*. I clade comprendenti gli isolati albanesi sono stati evidenziati in colore: 1a, azzurro; 1b, blu; 2a/c, rosa; 3a, arancione; 4, verde.

Sulla base di tale albero, 33 soggetti sono risultati di sottotipo 1b (49.2%), 15 di sottotipo 2a/c (22.4%), 10 di sottotipo 3a (14.9%) 5 di sottotipo 1a (7.5%) e 4 di genotipo 4 (6%).

Un'analisi separata è stata condotta per 2 sequenze che non risultavano correttamente assegnate sulla base del precedente albero e sono risultate sottotipo 4o (Figura 18)

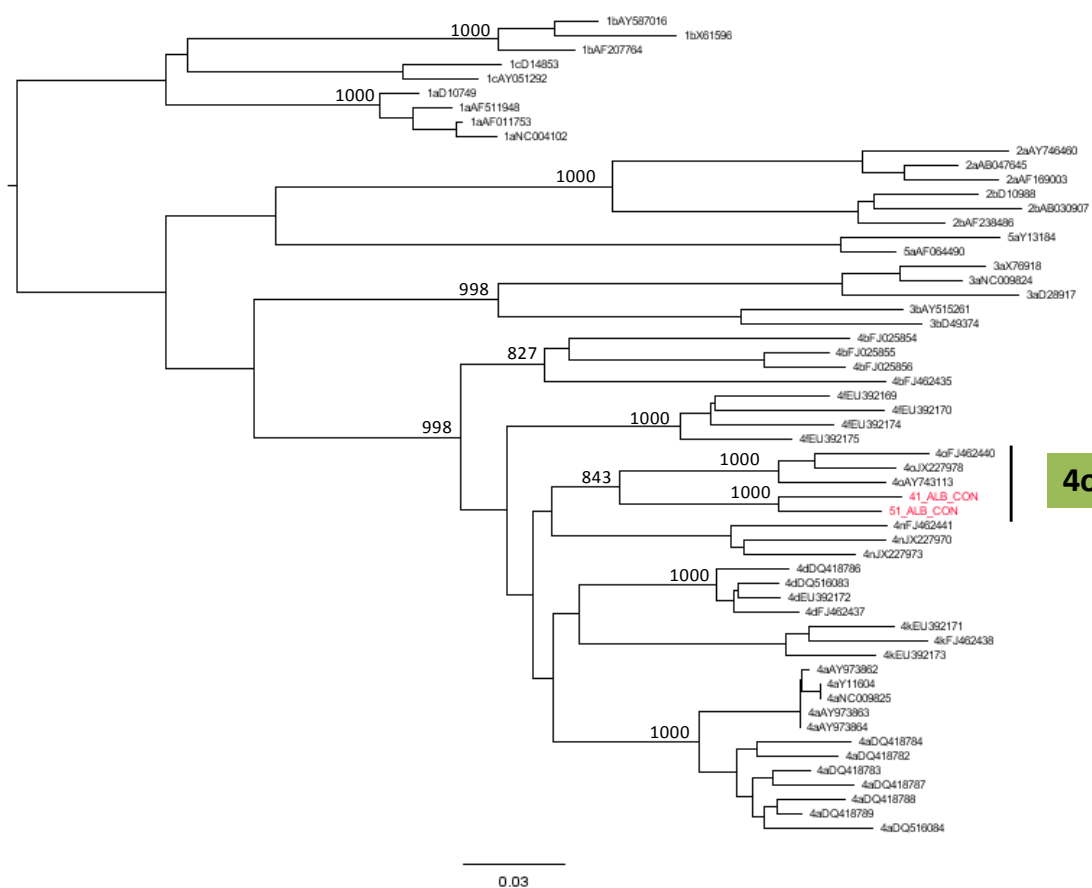


Figura 18: Albero NJ di 2 sequenze albanesi di sottotipo 4o (in rosso) allineate con le sequenze di riferimento rappresentative dei principali genotipi e sottotipi di HCV. I numeri sui rami indicano i valori di *bootstrap*.

4.2.1 Correlazione tra sottotipi di HCV e fattori di rischio nei soggetti albanesi

Considerando la popolazione albanese è stata analizzata l'associazione tra i sottotipi nei tossicodipendenti e nella popolazione generale.

Dei 55 pazienti albanesi, appartenenti alla popolazione generale, 33 sono risultati di sottotipo 1b (60%), 15 di sottotipo 2a/c (in particolare 14 di sottotipo 2c e 1 di sottotipo 2a) (27%), 2 di sottotipo 3a e 1a (4%) e 3 di genotipo 4 (5%), di cui uno di sottotipo 4d e due di sottotipo 4o (Figura 19).

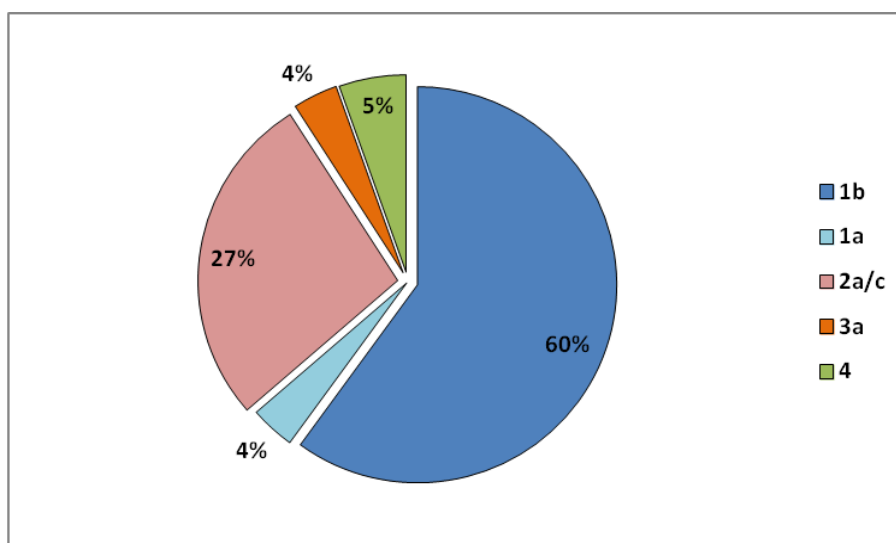


Figura 19: Distribuzione sottotipi nella popolazione generale albanese.

Per quanto riguarda il gruppo dei 12 IVDUs, 8 soggetti sono risultati di sottotipo 3a (67%), 3 di sottotipo 1a (25%) e 1 di sottotipo 4a (8%) (Figura 20). I dati sono riassunti nella tabella 6.

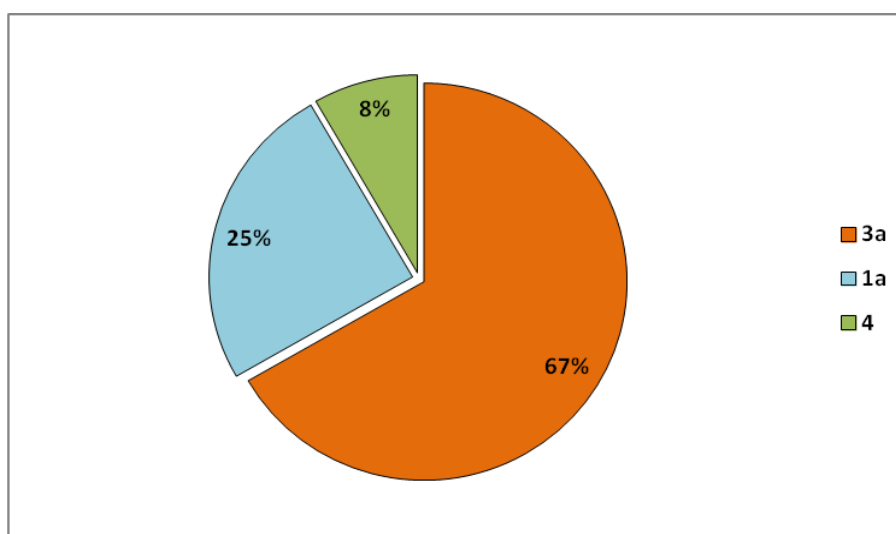


Figura 20: Distribuzione sottotipi tra gli IVDUs albanesi.

| GRUPPI A RISCHIO (n) | SOTTOTIPI n (%) | | | | |
|---------------------------|-----------------|----------|-----------|-----------|----------|
| | 1b | 1a | 2a/c | 3a | 4 |
| Popolazione generale (55) | 33 (60) | 2 (4) | 15 (27) | 2 (4) | 3 (5) |
| IVDUs (12) | - | 3 (25) | - | 8 (67) | 1 (8) |
| Totale | 33 | 5 | 15 | 10 | 4 |
| | | | | | |

Tabella 6: Distribuzione fattori di rischio e sottotipo virale.

4.2.2 Caratterizzazione dei sottotipi di HCV per sesso e per età nei soggetti albanesi

All'interno dei sottotipi, la popolazione albanese si distribuisce diversamente per quanto riguarda l'età e il sesso. Per quanto riguarda il sesso, i sottotipi 1b, 2a/c e 4 risultano maggiormente presenti nei maschi rispetto alle femmine (Figura 21).

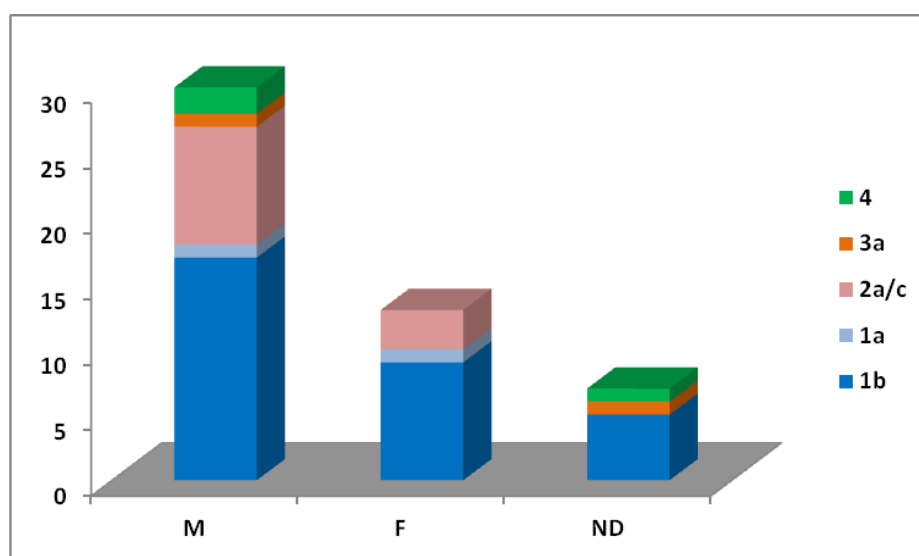


Figura 21: Distribuzione tra i sessi.

Riguardo l'età dei pazienti, la popolazione è stata suddivisa in tre gruppi: 18-28 anni, 29-39 anni, 40-55 anni. Il sottotipo 1b e 2a/c risultano maggiormente presenti nella fascia d'età più avanzata (44.4% e 40%; $p=0.02$ per il sottotipo 1b) con un *trend* crescente dalle fasce più giovani a quelle più anziane. I valori sono riportati nella tabella 7 e nella figura 22.

| | SOTTOTIPI n (%) | | | | |
|---------------------|-----------------|-----------|--------|--------|--------|
| | 1a | 1b | 3a | 4 | 2a/c |
| FASCE D'ETA' | | | | | |
| 18-28 anni | 2 (40) | 3 (9.1) | 1 (10) | 1 (25) | 3 (20) |
| 29-39 anni | | 9 (27.3) | | 1 (25) | 3 (20) |
| 40-55 anni | | 14 (42.4) | | | 6(40) |
| ND | 3 (60) | 7 (21.2) | 9 (90) | 2 (50) | 3 (20) |
| Totale | 5 | 33 | 10 | 4 | 15 |

Tabella 7: Distribuzione sottotipi e fasce d'età.

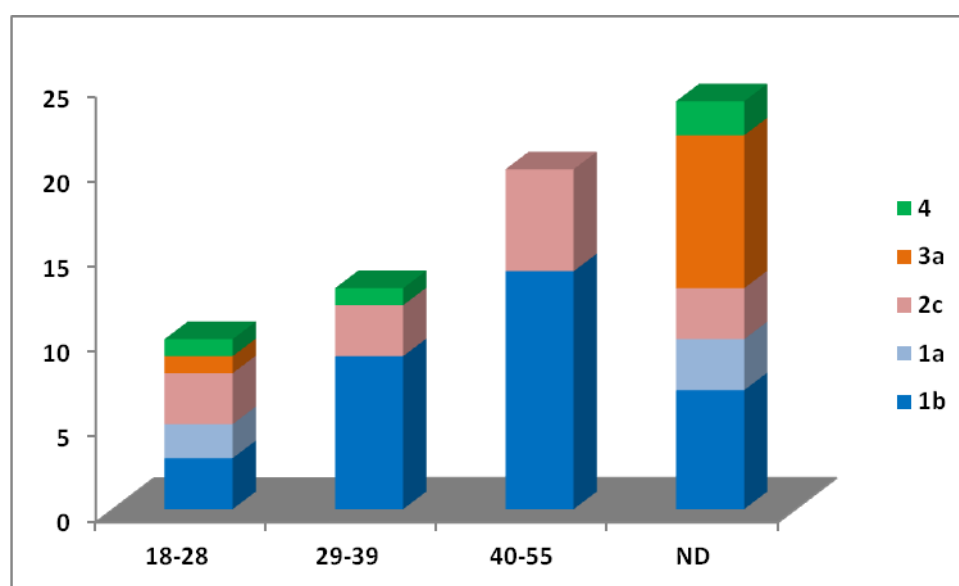


Figura 22: Distribuzione sottotipi e fasce d'età.

4.3 STORIA EPIDEMIOLOGICA DEL SOTTOTIPO 1b

4.3.1 Analisi filogeografica del sottotipo 1b

Come detto nella sezione Materiali e Metodi, dato che delle molte sequenze 1b ottenute dalle banche dati mancava la data di campionamento, per i 227 isolati del *data set* 1b è stato impostato un *evolutionary rate* (ER) pari a 8.4×10^{-4} ($0.6-1.5 \times 10^{-3}$).

L'albero filogeografico del sottotipo 1b (Figura 23) ha mostrato la presenza di 2 clade distinti che chiamiamo A e B. Il clade A include la maggior parte dei ceppi del Montenegro (26/37, 72,2%) che mostrano dei piccoli subclade significativi (*posterior probability*, pp=0.99, 0,92 e 0.93); gran parte delle sequenze italiane che formano un sottoclade altrettanto significativo (pp=0.98); un gruppo di sequenze tedesche e infine la totalità delle sequenze albanesi che si raggruppano in un unico subclade, non significativo, ma molto vicino agli isolati montenegrini.

Al contrario, nel clade B sono presenti la maggior parte delle sequenze turche (21/24, 87.5%) insieme al restante numero di sequenze montenegrine (11/37, 29.7%); in più è incluso un sottoclade significativo di 3 sequenze spagnole (pp=0.99).

Sulla base della ricostruzione filogeografica del *data set* 1b, la località più probabile della radice dell'albero è risultata Cipro (*state posterior probability* sp=0.78), seguita dal Montenegro (ps=0.18). Complessivamente, quindi, la probabilità posteriore che l'origine di HCV di sottotipo 1b sia nell'area dell'Egeo orientale è quasi dell'80%.

L'origine più probabile dei due clade A e B è Cipro (sp=0.86 e 0.89).

La stima del tMRCA (*time of most recent common ancestor*) della radice dell'albero mostra un valore di 102 anni fa (95% HPD 57-160) corrispondenti quindi a una presumibile origine dei ceppi analizzati risalente al 1912. I tMRCA dei clade A e B sono di 93 e 89 anni rispettivamente, risalenti al 1921 e al 1925. Due principali subclade montenegrini mostrano un'origine risalente a 77 anni fa (1937) e 66 anni fa (1948); il subclade italiano, incluso nel clade A, ha mostrato un tMRCA di 59 anni, risalente quindi al 1955.

Infine il subclade dei ceppi albanesi di sottotipo 1b hanno un origine di 73.3 anni fa risalente al 1941. I valori dei tMRCA e gli intervalli di confidenza (95% HPD) sono riassunti nella tabella 8.

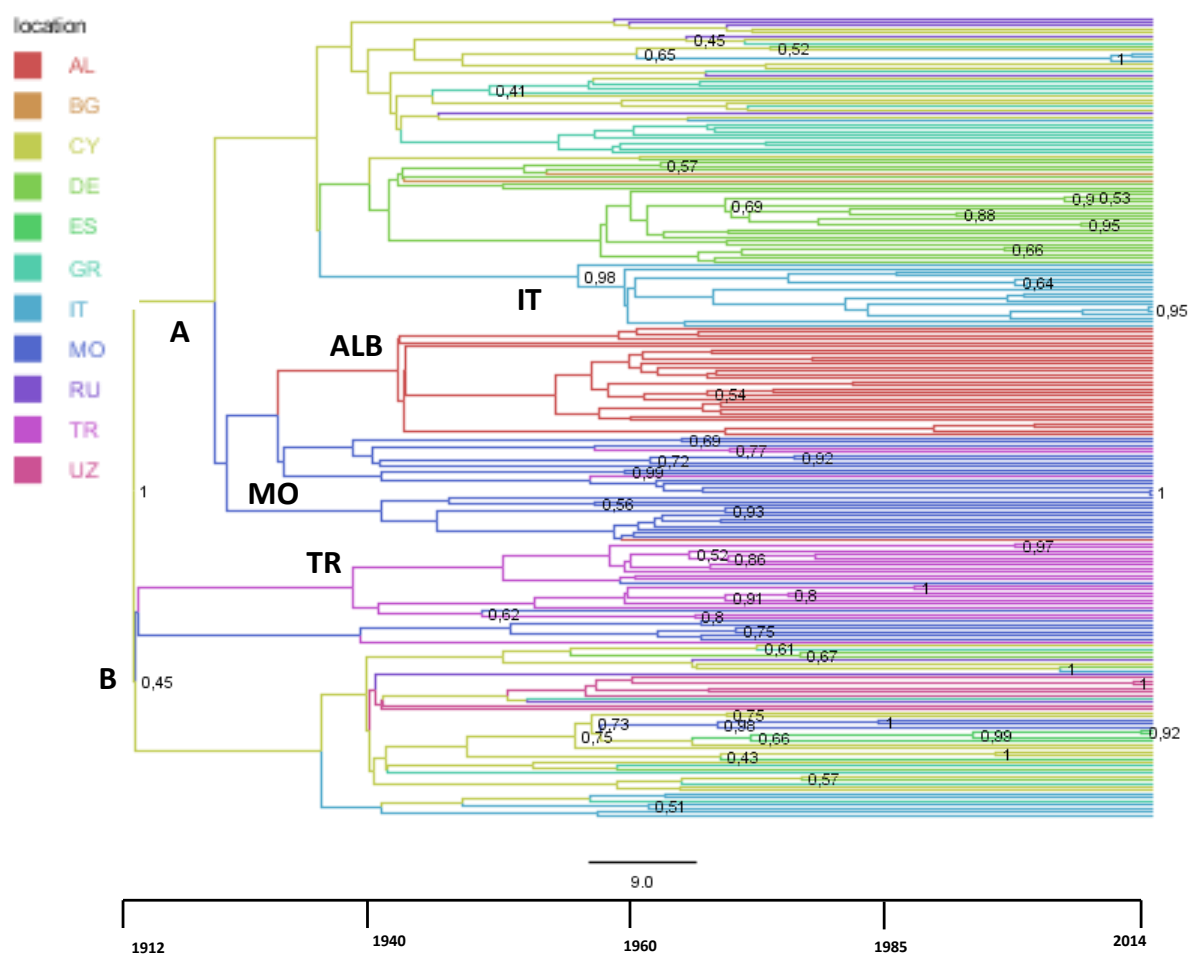


Figura 23: Albero filogeografico delle 227 sequenze del gene NS5B di HCV di sottotipo 1b, di differente provenienza geografica. La scala ai piedi dell'albero rappresenta il tempo (in anni) prima dell'ultimo anno di campionamento (2014). Nella legenda sono indicati i Paesi e i rispettivi colori.

| CLADE | Subclade | tMRCA | | Years | | Location | sp |
|-----------|----------|-------|------------|-------|-----------|----------|------|
| | | Mean | CI | Mean | CI | | |
| TREE ROOT | | 102.8 | 57.7-137.8 | 1912 | 1957-1877 | CY | 0.78 |
| A | | 93.3 | 53.3-118.7 | 1921 | 1960-1872 | CY | 0.86 |
| | MO | 77 | 48.7-131 | 1937 | 1965-1896 | MO | 0.98 |
| | IT | 58.9 | 32-94 | 1955 | 1932-1920 | IT | 0.99 |
| | ALB | 73.9 | 43-115 | 1940 | 1971-1899 | AL | 1 |
| B | | 89.2 | 51.7-137 | 1925 | 1962-1877 | CY | 0.89 |
| | TR | 70.8 | 40.2-104.7 | 1943 | 1974-1909 | TR | 0.82 |

Tabella 8: tMRCA e intervalli di confidenza (CI) per i principali clade e subclade.

Il metodo utilizzato consente, anche, di calcolare su base statistica (*Bayes Factor*, *BF*) la significatività delle migrazioni tra una località e l'altra. A questo scopo è stato utilizzato un cut-off di $BF=13$. Inoltre le sequenze sono state sottoposte ad un'analisi con il programma *Spread* 1.0.6 che permette di creare una possibile riproduzione visiva della diffusione del virus nelle regioni in studio. L'analisi ha evidenziato relazioni significative tra Cipro e la Grecia ($BF=33560$), il Montenegro e Albania ($BF=15$), la Turchia e il Montenegro ($BF=6940$) e Cipro e Montenegro ($BF=12$) (Figura 24).

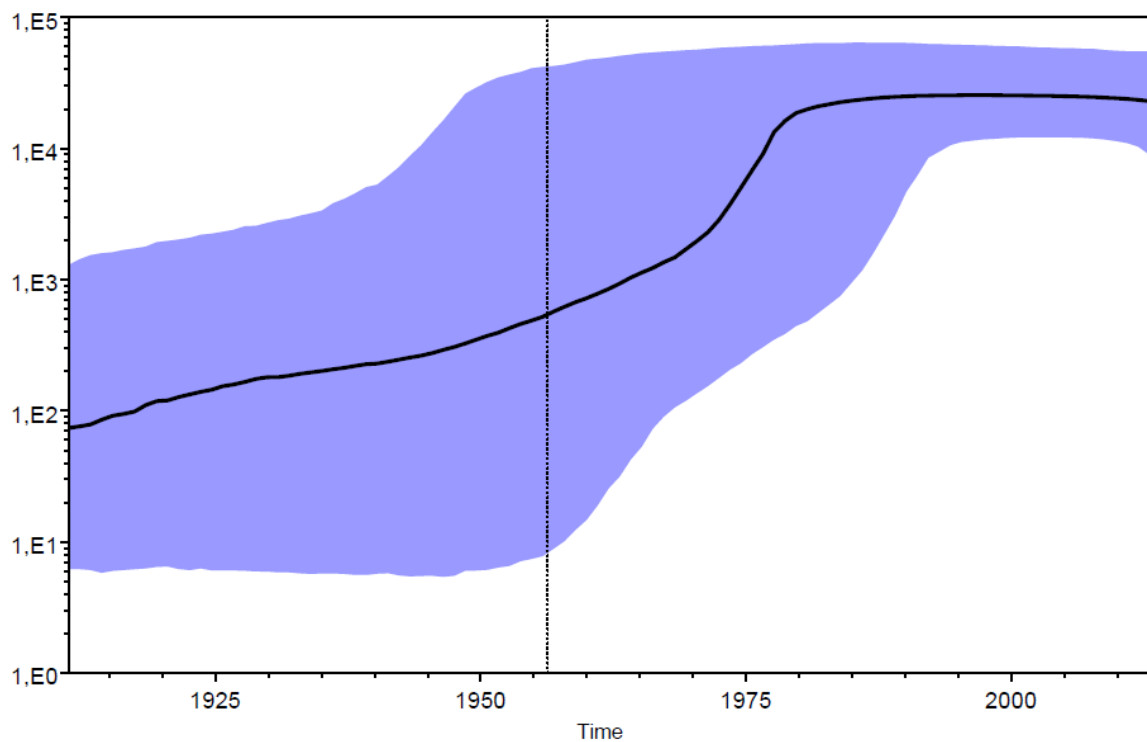


Figura 24: Analisi con *Spread* della diffusione di HCV-1b nelle regioni in studio.

4.3.2 Analisi filodinamica del sottotipo 1b

L'analisi dello *skyline plot* condotta sul *data set* completo del sottotipo1b (Figura 25) suggerisce che tale sottotipo sia presente nella popolazione già dagli anni '20 e mostra una lieve crescita fino agli anni '50-'60. In seguito si è assistito ad una crescita esponenziale del numero effettivo di infezioni che si sarebbe verificata intorno agli anni '70 fino agli anni '80, quando la curva raggiunge il *plateau* che perdura ancora oggi.

Figura 25: *Skyline* realizzato sul *data set* di isolati 1b.



4.4 STORIA EPIDEMIOLOGICA DEL SOTTOTIPO 3a

4.4.1 Analisi filogeografica del sottotipo 3a

L'*Evolutionary rate* (ER) delle sequenze del sottotipo 3a è stato stimato sulle 271 sequenze datate, mediante i metodi Bayesiani (MCMC), comparando due modelli di orologi molecolari, lo “stretto” e il “rilassato”, (BSP).

Il confronto tramite *Bayes Factor* (BF) dei valori di *marginal likelihoods* degli alberi ottenuti evidenzia come l'albero prodotto con l'orologio molecolare rilassato risulti significativamente migliore rispetto a quello ottenuto con l'orologio molecolare stretto ($2\ln\text{BF}$ orologio rilassato vs stretto = 85.9 a favore del primo). Applicando tale modello abbiamo ottenuto una stima dell'ER pari a 3.86×10^{-3} *substitution/site/year*, con un intervallo di credibilità del 95% tra 2.57×10^{-3} e 5.71×10^{-3} .

In base all'analisi filogeografica del *data set* 3a completo (Figura 26), la località più probabile della radice dell'albero è risultata il Pakistan (*state probability* sp=0.57).

Dal momento che tre sequenze pakistane si mettevano all'*outgroup* dell'albero risultava difficoltoso lo studio filogeografico. Per questa ragione l'analisi è stata ripetuta omettendo le tre sequenze. Come si può osservare in figura 27, la topologia dell'albero era la medesima dell'albero precedente così come i tempi stimati per i nodi interni non differivano significativamente.

L'albero filogeografico del sottotipo 3a (Figura 27) ha mostrato la presenza di diversi clade su base geografica. In particolare, è presente un clade significativo, chiamato clade A, (pp=0.86) eterogeneo, comprendente isolati dall'Estonia, Cipro, Lituania e Azerbaijan; successivamente si evidenzia la presenza di un altro clade significativo, chiamato clade B, (pp=0.98) che comprende la maggior parte dei ceppi della Bulgaria (11/14) e dell'Azerbaijan (16/29). Il terzo clade significativo riscontrato (clade C, pp=0.88) comprende 21 isolati pakistani insieme a pochi isolati inglesi (2/17) che nel complesso appaiono distribuiti in modo sparso in tutto l'albero. Un altro subclade pakistano era evidente ma non era significativamente supportato

Gli isolati montenegrini si dispongono in quattro gruppi significativi: il primo (clade D) formato da 12 ceppi con pp=0.86; il secondo comprende 4 ceppi con pp=1 (clade E); il terzo comprende 20 ceppi con pp=0.98 (clade G); mentre il quarto è formato da 11 ceppi con pp=0.96 (clade G). I 10 isolati albanesi si trovano sparsi nell'albero ma in alcuni casi si raggruppano vicino a sequenze montenegrine.

Sono stati inoltre stimati i tMRCA per la radice e per i principali clade significativi. La radice dell'albero mostra un tMRCA di 144 anni fa in Pakistan, seguito da un nodo principale euro-asiatico di 37 anni fa, corrispondente al 1977.

Il clade A e il clade B hanno un'origine in Azerbaijan risalente a circa 22 anni fa (1992).

Il clade C ha un'origine probabile in Pakistan 23 anni fa (1990) (95% HPD 16.2-30.64).

I quattro clade montenegrini hanno origine più recente rispettivamente di 17, 13, 12 e 6 anni fa.

Nella tabella 9 sono riassunti i tMRCA stimati, le date e le località più probabili, con i limiti credibilità inferiori e superiori dei principali *clade* descritti.

| CLADE | tMRCA | | Years | | Location | sp |
|-------------------------|-------|-----------|-------|-----------|----------|------|
| | Mean | CI | Mean | CI | | |
| TREE ROOT | 144 | 66-237.4 | 1870 | 1948-1776 | PA | 0.57 |
| Europeo/Asiatico | 37 | 23.6-57.9 | 1977 | 1990-1956 | PA | 0.35 |
| A | 22.1 | 18.9-35.5 | 1992 | 1995-1978 | AZ | 0.99 |
| B | 22.09 | 14.9-28.2 | 1992 | 1999-1985 | AZ | 0.96 |
| C | 23.3 | 16.2-30.6 | 1990 | 1998-1983 | PA | 0.99 |
| D | 17.4 | 12.3-22.4 | 1996 | 2001-1991 | MO | 0.99 |
| E | 6.86 | 9.9-22 | 2007 | 2007-1992 | MO | 1 |
| F | 13.4 | 9.7-17.5 | 2000 | 2004-1996 | MO | 1 |
| G | 12.2 | 8.3-16.5 | 2001 | 2006-1997 | MO | 1 |

Tabella 9: tMRCA e intervalli di confidenza (CI) per i principali clade.

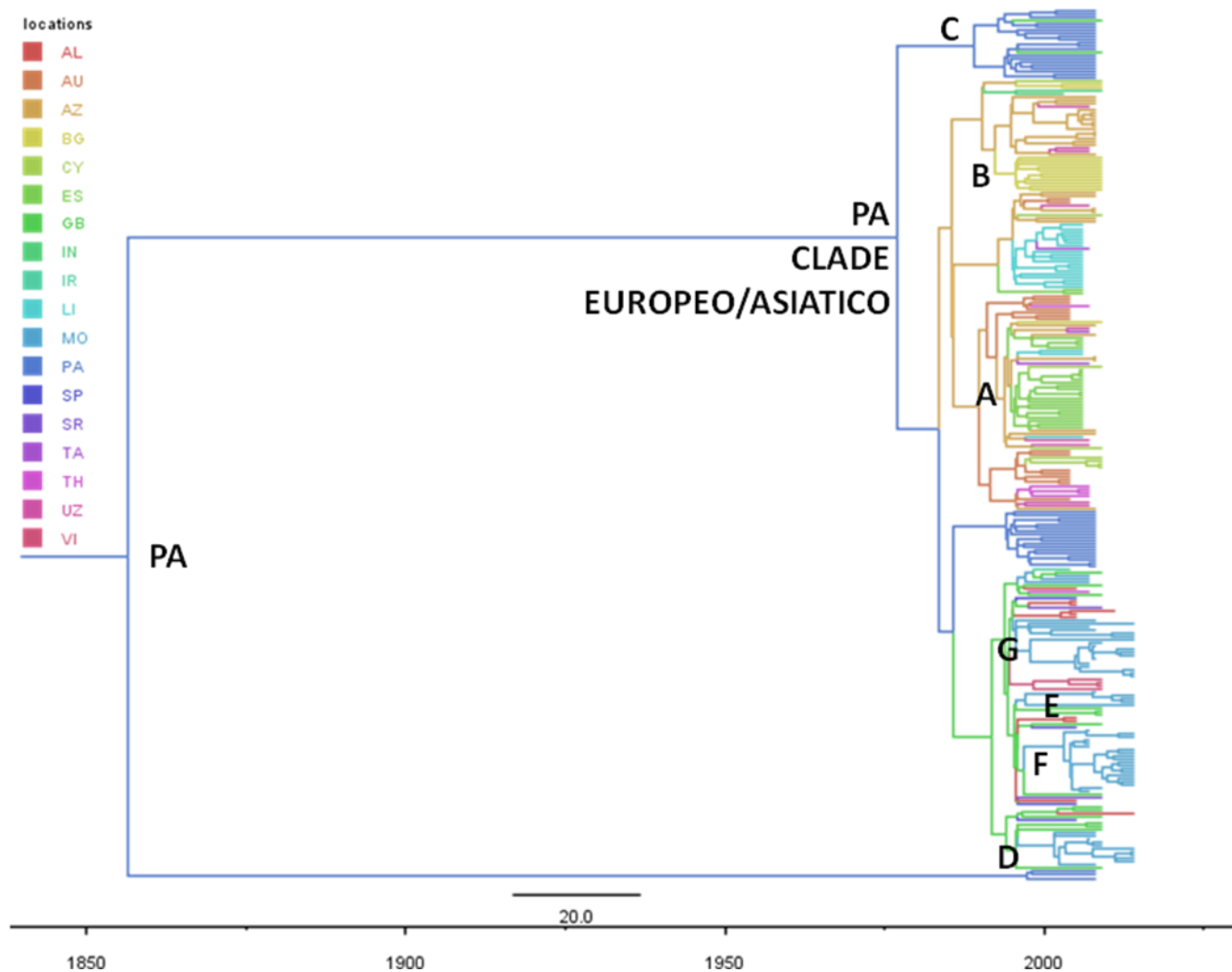


Figura26: Albero filogeografico delle 271 sequenze del gene NS5B di HCV di sottotipo 3a di differente provenienza geografica. La scala ai piedi dell'albero rappresenta il tempo (in anni) prima dell'ultimo anno di campionamento (2014). Nella leggenda sono indicati i paesi e i rispettivi colori.

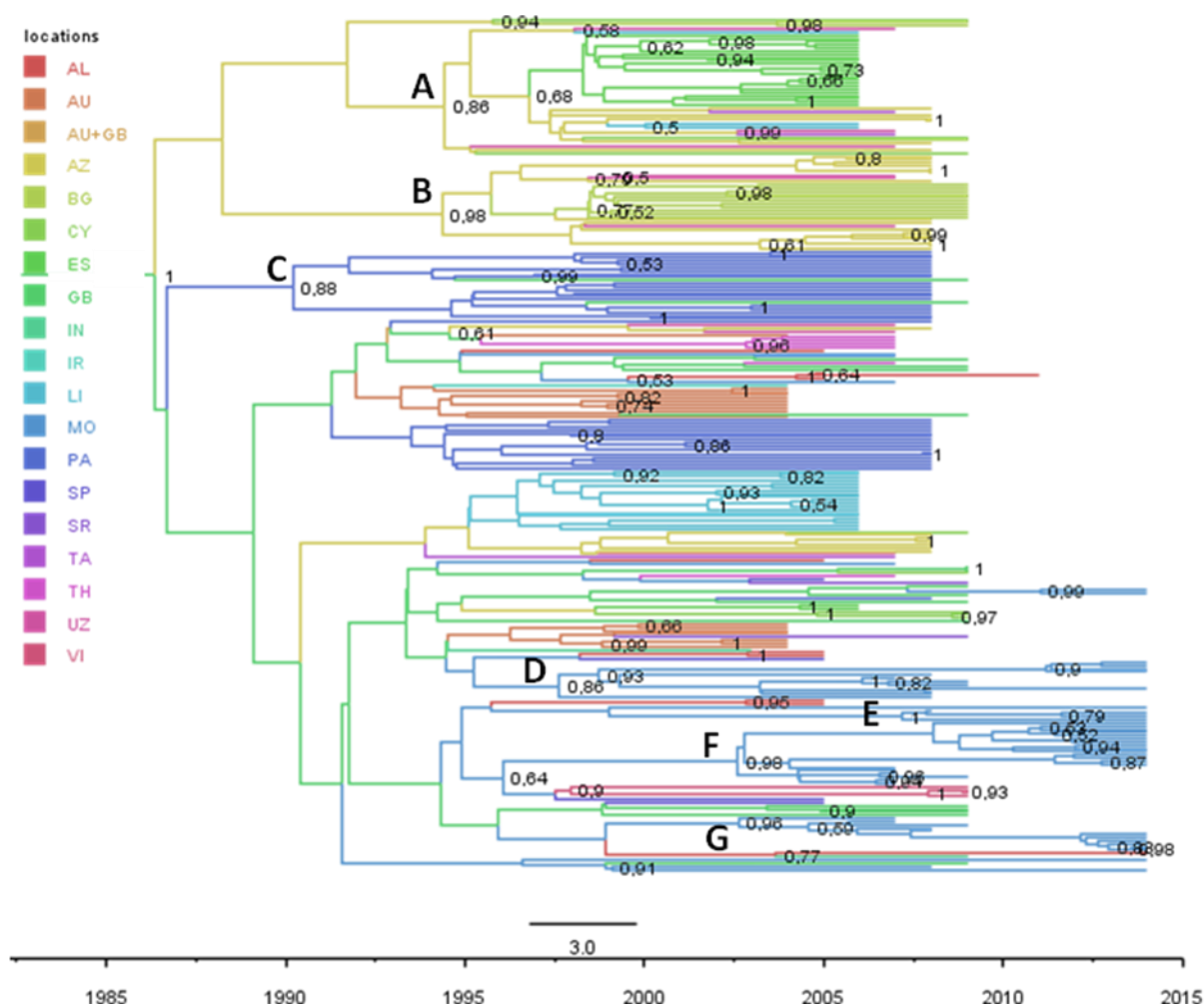


Figura 27: Albero filogeografico delle 268 sequenze del gene NS5B di HCV di sottotipo 3a di differente provenienza geografica. La scala ai piedi dell'albero rappresenta il tempo (in anni) prima dell'ultimo anno di campionamento (2014). Nella leggenda sono indicati i paesi e i rispettivi colori.

4.4.2 Analisi filodinamica degli IVDUs 3a

Dal *data set* di partenza abbiamo estrapolato 54 sequenze montenegrine e albanesi di sottotipo 3a che avevano come fattore di rischio l'uso di droghe e tramite *Beast* abbiamo effettuato l'analisi filodinamica. L'albero datato ottenuto è mostrato in Figura 28.

L'albero dei soggetti IVDUs albanesi e montenegrini mostra due clade altamente significativi relativi alle due popolazioni, chiamati A e B, mostrando una separazione netta tra gli isolati albanesi e quelli montenegrini. Il clade A contiene tutti gli isolati albanesi con due isolati montenegrini ($pp=0.9$), mentre il clade B ($pp=0.93$) contiene gli isolati montenegrini che si suddividono in 4 subclade significativi che sono presenti, anche, nella ricostruzione filogeografia globale e sono stati, allo stesso modo, nominati D, E, F e G. Tali clade dimostrano la presenza di recenti catene di trasmissione separate tra gli IVDUs in Montenegro. La radice dell'albero ha un tMRCA di 40 anni fa risalente al 1976. Sia il clade montenegrino che albanese hanno un'origine di circa 32 anni fa quindi intorno al 1982, mentre le recenti catene di trasmissione risalgono a 10-11 anni fa intorno agli anni 2000. La tabella 10 riassume la stima dei tMRCA dei clade evidenziati.

| CLADE | Subclade | tMRCA | | Years | |
|------------------|----------|-------|-----------|-------|-----------|
| | | Mean | CI | Mean | CI |
| TREE ROOT | | 39.4 | 23.1-57.6 | 1974 | 1990-1956 |
| A | | 32.5 | 19.3-45.4 | 1981 | 1994-1969 |
| B | | 32.8 | 20.5-46.2 | 1981 | 1993-1968 |
| | D | 11.6 | 7.5-18.4 | 2002 | 2006-1995 |
| | E | 8.7 | 6-12.5 | 2005 | 2008-2001 |
| | F | 12.3 | 7.8-18.7 | 2002 | 2006-1995 |
| | G | 11.9 | 7.2-18.4 | 2002 | 2007-1995 |

Tabella 10: tMRCA e intervalli di confidenza (CI) per i principali clade e subclade.

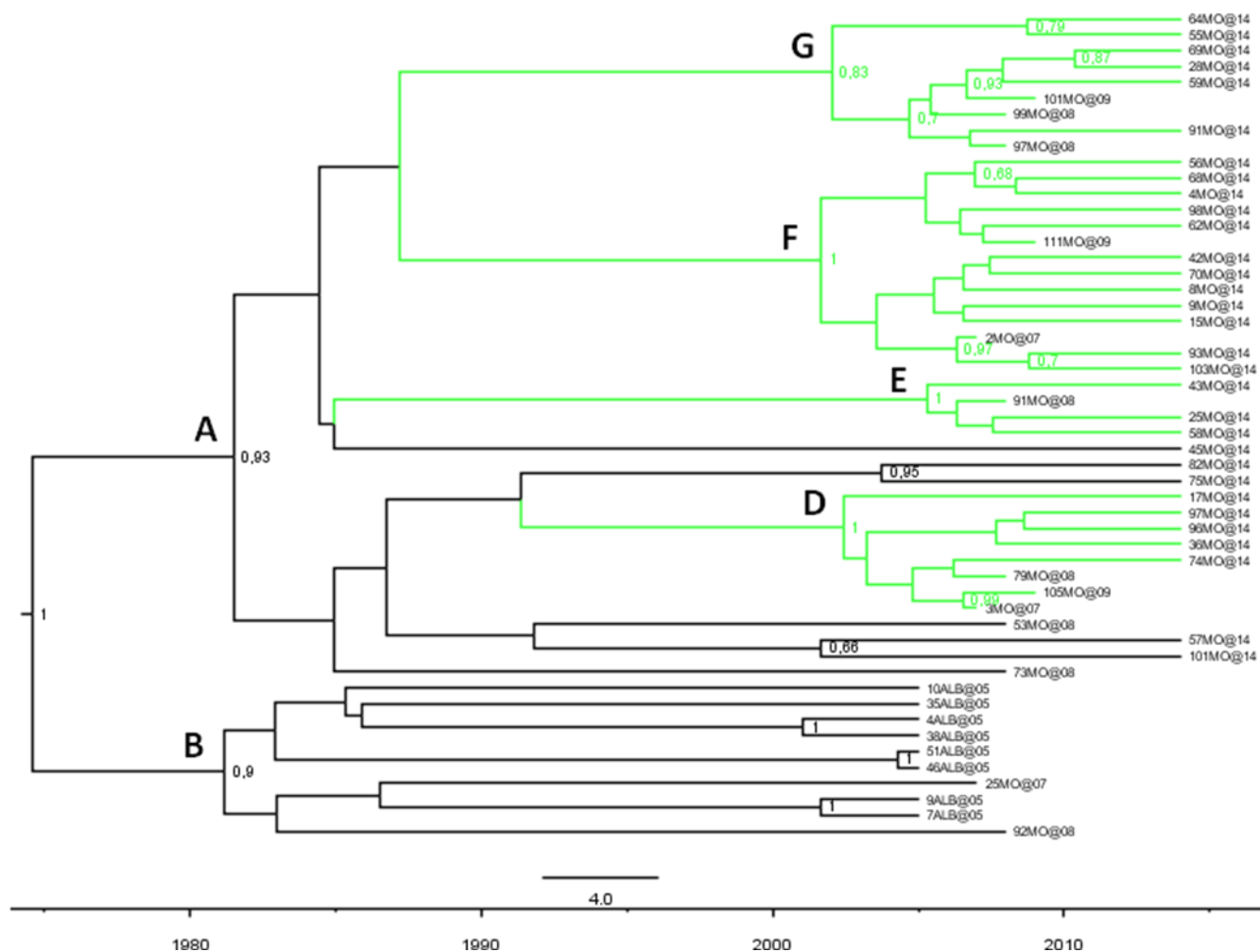


Figura 28: Albero datato delle 54 sequenze montenegrine e albanesi del gene NS5B di HCV di sottotipo 3a. La scala ai piedi dell'albero rappresenta il tempo (in anni) prima dell'ultimo anno di campionamento (2014). I clade significativi sono evidenziati in verde.

L'analisi dello *skyline plot* condotta sul *data set* (Figura 29) suggerisce che il sottotipo sia presente in tali paesi dagli anni '70; successivamente si è assistito ad una crescita esponenziale del numero effettivo di infezioni che si sarebbe verificata intorno agli anni '80-'90. A partire dagli anni '90 la curva raggiunge un *plateau*, ma la crescita riprende intorno agli anni 2000 e arriva a un *plateau* che perdura ancora oggi.

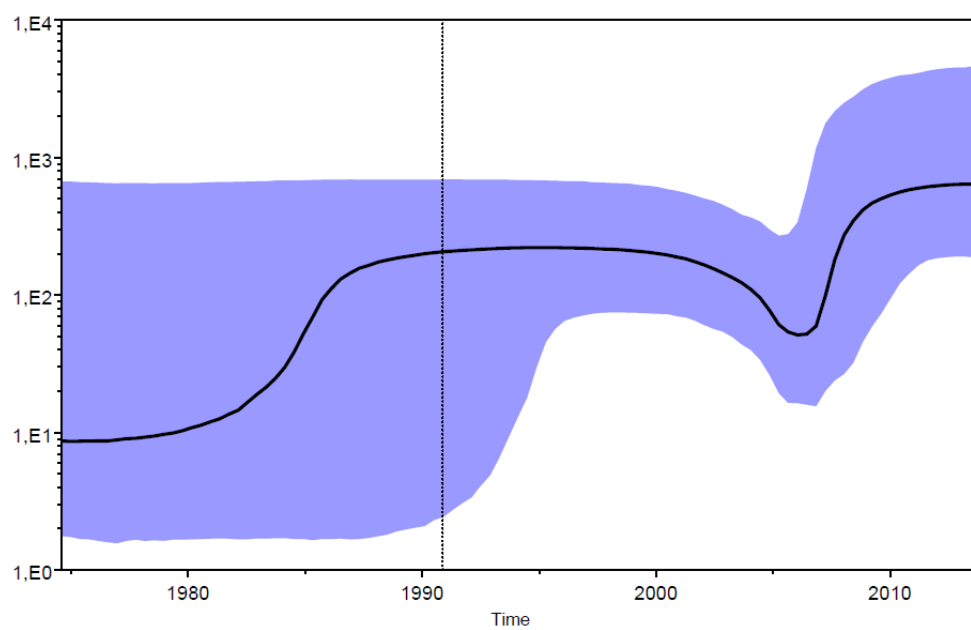


Figura 29: *Skyline* realizzato sul *data set* degli isolati albanesi e montenegrini 3a.

4.5 STORIA EPIDEMIOLOGICA DEL SOTTOTIPO 4d

4.5.1 Analisi filogeografica del sottotipo 4d

L'ER delle sequenze del sottotipo 4d è stato stimato su 106 isolati comparando due modelli di orologi molecolari, lo "stretto" e il "rilassato", (BSP).

Anche in questo caso il test *Bayes Factor* (BF) evidenzia che l'albero prodotto con l'orologio molecolare rilassato è significativamente migliore rispetto a quello ottenuto con l'orologio molecolare stretto ($2\ln\text{BF}$ orologio rilassato vs stretto = 34.7 a favore del primo). Applicando un modello di orologio molecolare rilassato abbiamo ottenuto una stima dell'ER pari a 7.53×10^{-4} *substitution/site/year*, con un intervallo di credibilità del 95% tra 2.74×10^{-4} e 12.1×10^{-4} .

L'albero mostra due clade principali: clade I (pp=1) che include tutte le sequenze turche mentre l'altro, clade II, include la maggior parte delle sequenze europee (Olanda, Italia, Francia, Montenegro e Albania) (Figura 30). In particolare, tutti i ceppi del Montenegro (29/31) erano inclusi in un unico clade altamente significativo (clade III, pp=0.98). Un gruppo di isolati italiani formano un clade, non significativo, che fa da *outgroup* al clade europeo.

L'origine più probabile dell'albero risulta essere l'Italia (ps=0.72). Per quanto riguarda la stima dei tempi di origine degli isolati, la radice dell'albero mostra un tMRCA di circa 125 anni fa che corrisponde a una presumibile origine intorno al 1890. Il clade turco è datato 48 anni fa mentre quello montenegrino mostra un'origine risalente a 50 anni fa corrispondente alla metà degli anni '60. Il clade europeo che comprende i ceppi italiani mostra un tMRCA di circa 94 anni fa, intorno quindi al 1920.

Nella tabella 11 sono evidenziati i clade principali con i rispettivi tMRCA e intervalli di confidenza.

| CLADE | tMRCA | | Years | | Location | sp |
|-------------|-------|------------|-------|-----------|----------|------|
| | Mean | CI | Mean | CI | | |
| TREE | | | | | | |
| ROOT | 125.7 | 44.3-255.8 | 1888 | 1969-1758 | IT | 0.72 |
| I | 48.2 | 19.5-96.2 | 1966 | 1994-1918 | TR | 1 |
| II | 65.8 | 26.1-131.3 | 1948 | 1987-1882 | NL | 0.62 |
| III | 49.9 | 21.1-99.3 | 1964 | 1992-1914 | MO | 0.99 |

Tabella 11: tMRCA e intervalli di confidenza (CI) per i principali clade.

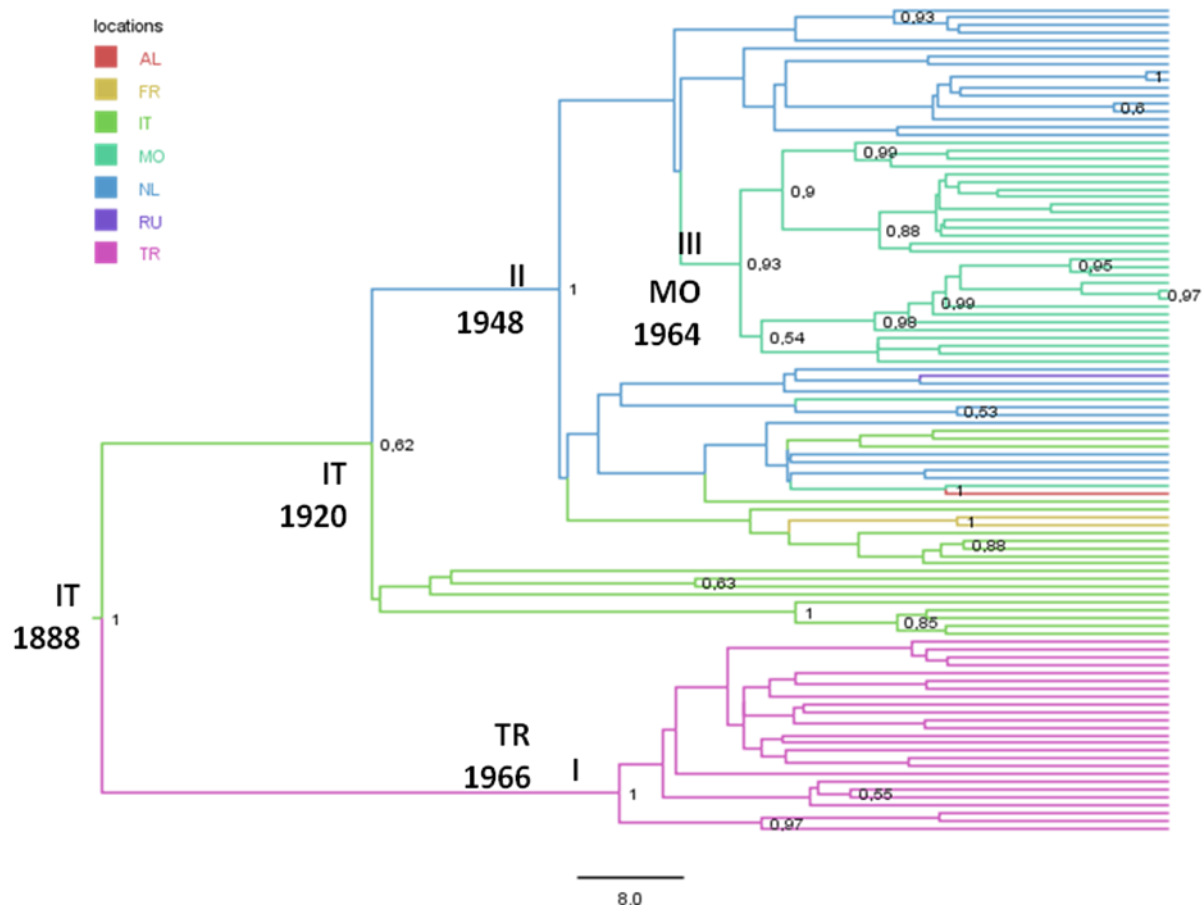


Figura 30: Albero filogeografico delle 106 sequenze del gene NS5B di HCV di sottotipo 4d di differente provenienza geografica. Gli anni e i clade principali sono riportati sui nodi. Nella leggenda sono indicati i paesi e i rispettivi colori.

4.5.2 Analisi filodinamica del sottotipo 4d

L'analisi filodinamica dell'intero *data set* 4d (Figura 31) ha mostrato una presenza costante del sottotipo dall'origine fino agli anni '80, quando è iniziata una crescita esponenziale dell'epidemia globale 4d dagli anni '90 agli inizi del 2000 raggiungendo un *plateau* che perdura ancora oggi.

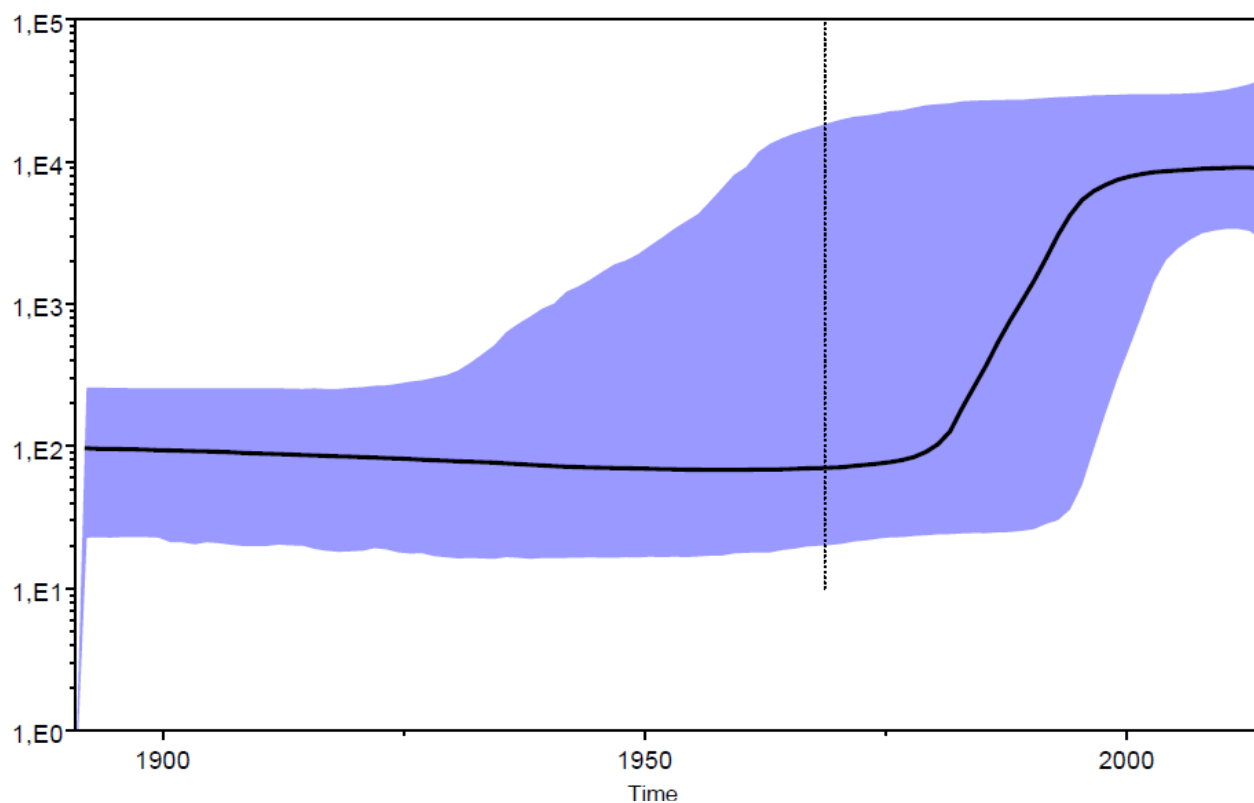


Figura 31: Skyline del *data set* completo 4d.

5. DISCUSSIONE

L'epidemia del virus HCV in Europa è in continua evoluzione e i parametri epidemiologici quali la prevalenza, l'incidenza, la trasmissione e la distribuzione genotipica sono cambiati in maniera sostanziale negli ultimi 15 anni⁸⁹.

Il trattamento attuale dell'epatite C ha le potenzialità di cambiare radicalmente la situazione nei prossimi anni. Tuttavia aggiornare i dati di prevalenza e di distribuzione genotipica è fondamentale per lo sviluppo di strategie di gestione ed eliminazione dell'infezione. Diversi studi recenti hanno riportato stime globali e regionali di prevalenza di HCV ma con il limite di considerare solo gruppi selezionati e non popolazioni ad alto rischio che sono i maggiori *driver* recenti dell'infezione (movimenti migratori, profughi, traffici illegali)⁹⁰.

Attualmente le informazioni sulla circolazione dei genotipi/sottotipi di HCV in Europa dell'Est e nell'area del Mediterraneo sono scarse. Non vi sono dati riguardanti la prevalenza di HCV nella popolazione generale in Montenegro. Altri studi epidemiologici sui Balcani occidentali hanno riscontrato una bassa frequenza di infezioni di HCV in Albania e tra i rifugiati del Kosovo; tuttavia nella stessa area si registra un'elevata prevalenza di HBV⁹¹.

In questo lavoro di tesi si è voluta pertanto studiare la prevalenza e la distribuzione dei diversi sottotipi in due Paesi nell'area dei Balcani, quali il Montenegro e l'Albania; si è cercato inoltre di ricostruire la storia epidemiologica dell'infezione negli ultimi decenni, studiandone la dinamica di popolazione mediante un metodo di analisi basato sulla teoria della coalescenza.

I nostri dati dimostrano che nella totalità dei soggetti albanesi e montenegrini, il virus dell'epatite C mostra un'elevata eterogeneità geno/sottotipica. In particolare, il sottotipo più diffuso nella zona dei Balcani occidentali risulta essere l'1b, presente circa nel 30.5% della casistica globale (Albania e Montenegro). Il sottotipo 3a risulta il secondo più diffuso nella popolazione in studio con una prevalenza del 29.2%, simile all'1b, seguito dall'1a e il 4d presenti rispettivamente nel 19.2% e 14.5%. Tra i pazienti albanesi il sottotipo 2a/c è il secondo sottotipo più frequente, presente nel 27% della

popolazione. Curiosa è la presenza, anche se di sole due sequenze, del sottotipo 4o in Albania; questo dato necessita, però, di ulteriori approfondimenti.

Questi dati sono in accordo con la distribuzione riscontrata in uno studio sulla prevalenza genotipica del virus HCV in Serbia e Montenegro⁶⁵.

Analizzando la distribuzione dei genotipi in relazione ai comportamenti a rischio nella casistica globale, abbiamo osservato una prevalenza significativamente più elevata del sottotipo 3a nei soggetti che fanno uso di droghe per via iniettiva (61%), mentre il sottotipo 2c è risultato completamente assente in questo gruppo.

Il sottotipo 1b è risultato significativamente più frequente (53%) nella infezioni di tipo iatrogeno, suggerendo un'associazione di questo sottotipo con tale modalità di trasmissione.

Nella maggioranza dei casi il fattore di esposizione all'infezione non era noto, trattandosi per lo più di infezioni acquisite in comunità-sporadiche (*community-acquired*). Tuttavia è ipotizzabile che vari fattori, legati ad una trasmissione di tipo parenterale tra cui anche l'esposizione iatrogena, abbiano un ruolo maggiore nella sua diffusione. Sorprendentemente il sottotipo 4d è risultato molto diffuso nella popolazione generale e anche tra i tossicodipendenti.

Tutto questo rivela una certa complessità nella situazione epidemiologica di HCV, nell'area del Mediterraneo dell'Est. In particolare il sottotipo 1b è presente nelle forme sporadiche, in pazienti anziani con epatite C cronica; mentre il sottotipo 3a è prevalente nei tossicodipendenti.

Uno degli scopi del presente studio è di delineare la storia epidemiologica recente dell'infezione da HCV nei Balcani Occidentali attraverso lo studio della filogeografia e della filodinamica dei sottotipi di HCV più diffusi.

In particolare, abbiamo focalizzato la nostra attenzione sui sottotipi epidemici prevalenti nella nostra popolazione cioè: 1b, 3a e 4d.

Le analisi filodinamiche e filogeografiche sono state eseguite su *data set* separati contenenti i soli isolati 1b, 3a e 4d per studiare la storia epidemiologica di HCV in anni recenti.

La nostra ricostruzione spazio-temporale della filogenesi di HCV del sottotipo 1b, suggerisce un'origine di tale sottotipo risalente ai primi decenni del secolo scorso

(1900), come già precedentemente stimato da altri autori. In particolare è stato suggerito dagli stessi autori che HCV-1b abbia avuto una larga diffusione in Europa tra gli anni '40 e '80, presumibilmente a causa delle trasfusioni di sangue e della tossicodipendenza^{92,93}.

In accordo con queste osservazioni, l'analisi filodinamica dell'epidemia in Montenegro suggerisce che anche in quest'area geografica, cruciale per l'Est Europa, il numero di trasmissioni di HCV-1b sia cresciuto in modo esponenziale tra gli anni '50 e gli anni '80, anche in questo caso probabilmente a causa dell'uso di sangue (o sue frazioni) contaminate, intensificatosi nel periodo delle due guerre mondiali. D'altra parte va anche ricordato l'uso promiscuo di siringhe non sterilizzate, che è continuato in Europa almeno fino agli anni '50-'60 e oltre, fino a quando sono state introdotte le siringhe monouso⁹⁴. In particolare, in Europa dell'Est e Asia centrale il programma di riduzione del danno che prevedeva, tra l'altro, un uso proprio di siringhe da iniezione, è iniziato nelle repubbliche della ex-Yugoslavia nella metà degli anni '80⁹⁵. Proprio l'implementazione di regole più stringenti nell'uso di siringhe sterili insieme all'introduzione dei test di screening per HCV alla fine degli anni '80, sembra essere in associazione al *plateau* osservato nel numero di infezioni da HCV-1b negli ultimi anni.

Lo studio filogeografico ha suggerito un'origine dell'HCV-1b circolante in Europa, nell'area dell'Egeo, in particolare a Cipro. Tale risultato è rafforzato dai risultati di uno studio da noi condotto nel 2011, in cui la radice dell'albero filogeografico, comprendente una serie di isolati europei di HCV-1b, vedeva come località più probabile proprio la Grecia⁸⁸ e Cipro come seconda area più probabile.

Cipro è una piccola isola ma è posizionata al crocevia tra Europa, Africa e Asia, dando origine ad un alto flusso di turisti e immigrati politici dall'Europa dell'Est e dal Medio Oriente. E' noto come a Cipro coesistano tutti i sottotipi menzionati, con prevalenze variabili⁹⁶. Considerata la grande eterogeneità di sottotipi presenti, è assai probabile che Cipro rappresenti un bivio tra tre continenti in cui tale sottotipo si può essere diversificato.

Nella presente analisi gli isolati 1b analizzati segregavano in 2 clade principali, uno comprendente principalmente i ceppi dell'Europa Occidentale, l'altro più limitato all'area Sud Orientale, contenente tra l'altro sequenze di provenienza turca. Ambedue i clade sono originati intorno agli anni '20. In un recente studio sull'epidemiologia

molecolare di HCV in Turchia⁸⁸ abbiamo osservato una netta prevalenza del sottotipo 1b, che si sarebbe diffuso, anche in questo caso, a cominciare dagli anni '20 per cause iatrogene; mentre la tossicodipendenza, epidemiologicamente legata ad altri sottotipi di HCV (1a, 3a), non sembra aver rivestito un ruolo primario. In particolare le sequenze del Montenegro tendono a raggrupparsi (non significativamente) in 2 gruppi principali: il primo, più grande, incluso nel clade Occidentale, datato negli anni '30, il secondo incluso nel clade turco, originato più tardi alla fine degli anni '60. Va ricordato che in Montenegro vive da tempo una minoranza etnica turca/musulmana (che rappresenta circa lo 3.31% della popolazione montenegrina), residuo di una popolazione molto più diffusa presente durante l'occupazione Ottomana, per lo più emigrata dalla Turchia alla fine degli anni '20⁹⁷.

È importante ricordare come l'Asia minore e il Mediterraneo Orientale sono stati, nell'epoca tra gli anni '20 e '30, teatro delle guerre balcaniche, un'immane tragedia seguita alla caduta dell'Impero Ottomano dopo la prima guerra mondiale, che ha visto una serie di conflitti militari, come quello Greco-Turco, a cui è conseguito lo spostamento forzato di più di 2 milioni di persone nell'area. Tali eventi possono aver creato le condizioni favorevoli a una rapida diffusione dell'infezione da HCV e un rimescolamento dei sottotipi prevalenti in diversi paesi dell'area geografica⁸⁸.

Le sequenze 1b albanesi tendono a raggrupparsi in un unico clade poco significativo proprio accanto al clade montenegrino. A causa della bassa significatività, non abbiamo potuto trarre conclusioni in relazione all'origine di tale sottotipo in Albania, anche se i clade, montenegrini e albanese, mostrano una stretta relazione evolutiva, suggerendo un possibile legame epidemiologico tra queste due aree. Per questa ragione HCV-1b necessita di essere studiato più approfonditamente su una più ampia casistica.

Profondamente diversa è, invece, la storia del sottotipo 3a in Europa, come appare anche dalla nostra analisi filogeografica. In particolare, l'analisi del *data set* completo, mostra che il sottotipo esiste da almeno 144 anni e la sua origine sembra essere localizzata nell'area del Pakistan. Recenti studi hanno suggerito che il sottotipo 3a sia endemico in Medio Oriente e sia presente, almeno dagli anni '40, nel subcontinente indiano. Negli anni '50-60 poi, ci sarebbe stata una rapida crescita dell'epidemia e una sua diffusione più ampia in tutto il mondo, a causa principalmente del traffico di droga

che partendo da questa area, devastata alla fine degli anni '80 dalla guerra in Afghanistan, si è diretta verso i paesi industrializzati⁹⁸.

Si stima che dal 60% a più dell'80% dei tossicodipendenti per via iniettiva siano attualmente HCV positivi in 37 paesi del mondo (circa 10 milioni di persone)⁹⁹.

La più vasta popolazione di tossicodipendenti HCV positiva vive in Europa dell'Est e in Sud Est Asiatico. E' noto che il sottotipo 3a è significativamente associato alla tossicodipendenza^{100, 101}, così come il sottotipo 1a^{92,93,102}. L'analisi filodinamica mostra una crescita esponenziale delle infezioni sostenuta da questo sottotipo tra gli anni '80 e '90 e il raggiungimento, nell'ultimo decennio, di un *plateau* nel numero di infezioni, osservato anche da altri autori in Europa¹⁰³, probabilmente in relazione ad un cambiamento nei comportamenti a rischio dei tossicodipendenti.

L'analisi filogeografica ha mostrato una netta segregazione degli isolati virali su base geografica. In particolare alcuni isolati bulgari, oggetto di una nostra recente pubblicazione, formano un gruppo altamente significativo, incluso nel clade A insieme a ceppi dell'Azerbaijan suggerendo la presenza del sottotipo nell'area compresa tra l'Asia centrale e l'Europa dell'Est¹⁰⁴. Allo stesso modo isolati pakistani ed estoni segregavano in clade significativi. Gli isolati Montenegrini, invece, formano 4 clade significativi anche in relazione con alcuni ceppi albanesi. Questi 4 clade mostrano un'origine recente intorno alla fine degli anni '90 e l'inizio del 2000 suggerendo la presenza di catene di trasmissione del sottotipo 3a dovute probabilmente a scambi tra tossicodipendenti in tale area. Tali clade, piccoli e significativi, sono tipici della trasmissione localmente ristretta che si verifica nelle popolazioni di tossicodipendenti e che ha generato una diversificazione genetica di HCV-3a nelle comunità locali.

Dobbiamo, quindi, ipotizzare che HCV-3a sia stato introdotto in Montenegro e Albania negli anni '80 e che sia stato soggetto ad una crescita esponenziale che sembra persistere anche recentemente, come suggerito dall'analisi filodinamica.

La presenza del sottotipo 4d di HCV in Montenegro è particolarmente interessante. Il genotipo 4 è la variante più comune in Egitto, Medio Oriente e Africa centrale ed è presente a livello mondiale nel 20% dei soggetti infetti da HCV. Il genotipo 4 è stato anche isolato in una quota significativa di pazienti viventi nei paesi occidentali dove rappresenta fino al 4.7% degli infetti. Il genotipo 4 è fortemente eterogeneo,

presentando numerosissimi sottotipi (circa 20) con una divergenza genetica significativa. Il sottotipo 4a è quello più diffuso in Egitto e, sulla base delle ricostruzioni filodinamiche, la sua diffusione sarebbe avvenuta tra gli anni '30 e '80 a causa di campagne di massa per il trattamento della schistosomiasi¹⁰⁵. Fuori dall'Africa altri sono i sottotipi più frequenti. Per esempio in Italia, Spagna ma anche Arabia Saudita il sottotipo 4c/d è nettamente più frequente. In Europa, il genotipo 4 è presente tra tossicodipendenti ma anche in casi sporadici non-tossicodipendenti (*community acquired*)¹⁰⁶. In particolare, in alcune regioni del sud Italia, tra cui la Calabria, è stata osservata una elevata prevalenza di genotipo 4, particolarmente in soggetti di età avanzata, suggerendo che il sottotipo sia endemico in quest'area geografica¹⁰⁷.

La nostra analisi filogeografica del sottotipo 4d mette in luce che il sottotipo si distribuisce in gruppi omogenei per regione geografica. Dall'albero filogeografico risulta che gli isolati montenegrini segregano in un clade altamente significativo, a eccezione di un isolato. Un gruppo di 9 isolati italiani formano un clade significativo con un nodo interno che appare più antico degli altri (tMRCAs= 83 anni, 1931). Il clade montenegrino, invece, ha una data di origine alla fine degli anni '60 (1964).

La filogeografia conferma la netta segregazione degli isolati nell'albero, con l'unica eccezione rappresentata dagli isolati italiani. La più probabile località di origine della radice dell'albero è risultata proprio il nostro paese. Quest'osservazione suggerisce che l'Italia potrebbe rappresentare il luogo di ingresso del sottotipo 4d in Europa in un'epoca risalente tra gli anni '30 e i primi anni '40. Tra l'altro risulta interessante come i ceppi italiani, inclusi nel clade più antico, sono per lo più individui calabresi non tossicodipendenti della zona di Crotona. I ripetuti contatti degli italiani con alcune aree africane negli anni '20-'30 potrebbe rappresentare la via di ingresso di HCV-4d nel nostro paese. Il clade montenegrino, invece, è datato intorno agli anni '60 e può essere ricondotto agli isolati italiani del secondo gruppo. In questo caso altre vie di trasmissione, come la tossicodipendenza, potrebbero essere state responsabili dell'epidemia più recente in Italia e in Montenegro: circa il 20% dei soggetti con genotipo 4d del Montenegro erano dichiaratamente tossicodipendenti. Tale epidemia è evidenziata anche dallo *skyline plot* che mostra una crescita esponenziale delle infezioni sostenuta da questo sottotipo nel periodo tra gli anni '80 e il 2000.

Naturalmente, il numero scarso di sequenze e le poche località europee rappresentate, limita un po' la significatività dell'analisi e richiede ulteriori studi futuri.

In conclusione, la vasta diversità di sottotipi riscontrata in questo studio dimostra come i Balcani Occidentali rappresentino, nell'epidemia di HCV, un crocevia tra l'Europa Occidentale e Orientale, il Medio Oriente e l'Africa con potenziali conseguenze importanti sulla epidemiologia dell'infezione da HCV.

Molti sono i fattori che contribuiscono alla diffusione di HCV, tra cui i flussi immigratori da aree endemiche, le guerre, la continua espansione dell'uso di droghe e i comportamenti ad alto rischio per infezioni sessualmente trasmissibili che hanno giocato e giocano un ruolo fondamentale in quest'area, punto di incrocio di varie epidemie presenti in continenti diversi.

Da qui deriva, a nostro avviso, la necessità di una sorveglianza continua dell'epidemia di HCV nella regione in studio, anche con l'aiuto degli strumenti avanzati di filogenesi, impiegati nella presente tesi, per ricostruire la storia recente dell'infezione.

6.BIBLIOGRAFIA

- ¹ Choo, Q. L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W. & Houghton, M. Isolation of a cDNA clone derived from a bloodborne non-A, non-B viral hepatitis genome. *Science* 1989; 244: 359-362.
- ² Dienstag JL. NANB hepatitis I. Recognition, epidemiology and clinical features. *Gastroenterology* 1983; 85: 439-462.
- ³ Alter HJ, Seeff LB. Recovery, persistence, and sequelae in hepatitis C virus infection: a perspective on long-term outcome. *Semin Liver Dis* 2000; 20: 17-35.
- ⁴ A.J. Alter, R.H. Purcell, P.V. Holland and H. Popper. Transmissible agent in 'non-A, non-B' hepatitis. *Lancet* 1978; 1: 459-463.
- ⁵ E. Tabor, R.J. Gerety, J.A. Drucker, L.B. Seeff, J.H. Hoofnagle, D.R. Jackson et al. Transmission of non-A, non-B hepatitis from man to chimpanzee. *Lancet* 1978; 1: 463-466.
- ⁶ L.I. He, D. Alling, T. Popkin, M. Shapiro, H.J. Alter and R.H. Purcell. Determining the size of non-A, non-B hepatitis virus by filtration. *J Infect Dis* 1987; 156: 636-640
- ⁷ S.M. Feinstone, K.B. Mihalik, A.J. Alter, W.T. London and R.H. Purcell. Inactivation of hepatitis B virus and non-A, non-B hepatitis by chloroform. *Infect Immunol* 1983; 41: 816-821.
- ⁸ H.J. Alter and M. Houghton. Clinical medical research award. Hepatitis C virus and eliminating post-transfusion hepatitis. *Nat Med* 2000; 6: 1082-1086
- ⁹ G. Kuo, Q.L. Choo, H.J. Alter, G.L. Gitnick, A.G. Redeker, R.H. Purcell et al. An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* 1988; 244: 362-364.
- ¹⁰ H.J. Alter, R.H. Purcell, J.W. Shih, J.C. Melpolder, M. Houghton, Q.L. Choo et al. Detection of antibody to hepatitis C virus in prospectively followed transfusion recipients with acute and chronic non-A, non-B hepatitis. *N Engl J Med* 1989; **321**:1494-1500.
- ¹¹ Q.L. Choo, G. Kuo, A.J. Weiner, L.R. Overby, D.W. Bradley and M. Houghton. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989; 244: 359-366.
- ¹² Francki RIB, Fauquet CM, Knudson DL, Brown F. Classification and nomenclature of viruses: fifth report of the international committee on taxonomy of viruses. *Arch Virol* 1991; 2: S223
- ¹³ Brett D. Lindenbach, Heinz-Jurgen Thiel, Rice CM. Flaviviridae: The viruses and their replication. *Fields Virology*, ed 5. Philadelphia 2007; 1101-1152.
- ¹⁴ I.R. Brey and R. Bartenschlager. Membranous Replication Factories Induced by Plus-Strand RNA Viruses. *Viruses* 2014; 6, 2826-2857
- ¹⁵ Major ME, Feinstone SM. The molecular virology of hepatitis C. *Hepatology* 1997; 25: 1527-38.
- ¹⁶ Q.L. Choo, K.H. Richman, J.H. Han, K. Berger, C. Lee, C. Dong et al. Genetic organization and diversity of the hepatitis C virus. *Proc Natl Acad Sci USA* 1991; 88: 2451-2455.

- ¹⁷ Houghton M. Hepatitis C virus. *Fields Virology*, 3rd ed. Philadelphia: Lippincott-Raven 1996; 1035-1058
- ¹⁸ Strader DB, Seeff LB. Hepatitis C: A brief clinical overview. *ILAR J* 2001; 42: 107-116
- ¹⁹ Simmonds P., Holmes E.C., Cha T.A. et al. Classification of hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. *J. Gen. Virol* 1993; 74: 2391-2399
- ²⁰ Wang, C., Le, S. Y., Ali, N., and Siddiqui, A. (1995). An RNA pseudoknot is an essential structural element of the internal ribosome entry site located within the hepatitis C virus 5' noncoding region. *RNA* 1, 526-537.
- ²¹ Tanaka, T., Kato, N., Cho, M. J., Sugiyama, K., and Shimotohno, K. (1996). Structure of the 3' terminus of the hepatitis C virus genome. *J Virol* 70, 3307-3312.
- ²² Suzuki, R., Sakamoto, S., Tsutsumi, T., Rikimaru, A., Tanaka, K., Shimoike, T., Moriishi, K., Iwasaki, T., Mizumoto, K., Matsuura, Y., et al. (2005). Molecular determinants for subcellular localization of hepatitis C virus core protein. *J Virol* 79, 1271-1281.
- ²³ Ray R B, Steele R, Meyer K and Ray R Hepatitis c virus core protein represses p21WAF1/Cip1/Sid1 promoter activity. *Gene* 1998; 208: 331-336.
- ²⁴ Ray RB, Lagging LM, Meyer K, Steele R, Ray R. Transcriptional regulation of cellular and viral promoters by the hepatitis C virus core protein. *Virus Res* 1995; 37: 209-220
- ²⁵ Cocquerel, L., Wychowski, C., Minner, F., Penin, F., and Dubuisson, J. (2000). Charged residues in the transmembrane domains of hepatitis C virus glycoproteins play a major role in the processing, subcellular localization, and assembly of these envelope proteins. *J Virol* 74, 3623-3633.
- ²⁶ Hijikata M, Kato N, Ootsuyama Y, et al. Hypervariable regions in the putative glycoprotein of hepatitis C virus. *Biochem Biophys Res Commun* 1991; 175: 220-228
- ²⁷ Farci, P., Shimoda, A., Wong, D., Cabezon, T., De Gioannis, D., Strazzer, A., Shimizu, Y., Shapiro, M., Alter, H. J., and Purcell, R. H. (1996). Prevention of hepatitis C virus infection in chimpanzees by hyperimmune serum against the hypervariable region 1 of the envelope 2 protein. *Proc Natl Acad Sci USA* 93, 15394-15399.
- ²⁸ Carrere-Kremer, S., Montpellier-Pala, C., Cocquerel, L., Wychowski, C., Penin, F., and Dubuisson, J. (2002). Subcellular localization and topology of the p7 polypeptide of hepatitis C virus. *J Virol* 76, 3720-3730.
- ²⁹ Sakai, A., Claire, M. S., Faulk, K., Govindarajan, S., Emerson, S. U., Purcell, R. H., and Bukh, J. (2003). The p7 polypeptide of hepatitis C virus is critical for infectivity and contains functionally important genotype-specific sequences. *Proc Natl Acad Sci USA* 100, 11646-11651.
- ³⁰ Gonzalez, M. E., and Carrasco, L. (2003). Viroporins. *FEBS Lett* 552, 28-34.
- ³¹ Gretton, S. N., Taylor, A. I., and McLauchlan, J. (2005). Mobility of the hepatitis C virus NS4B protein on the endoplasmic reticulum membrane and membrane-associated foci. *J Gen Virol* 86, 1415-1421.
- ³² Piccininni, S., Varaklioti, A., Nardelli, M., Dave, B., Raney, K. D., and McCarthy, J. E. (2002). Modulation of the hepatitis C virus RNA-dependent RNA polymerase activity by the non-structural (NS) 3 helicase and the NS4B membrane protein. *J Biol Chem* 277, 45670-45679.

- ³³ Tu, H., Gao, L., Shi, S. T., Taylor, D. R., Yang, T., Mircheff, A. K., Wen, Y., Gorbalenya, A. E., Hwang, S. B., and Lai, M. M. (1999). Hepatitis C virus RNA polymerase and NS5A complex with a SNARE-like protein. *Virology* 263, 30- 41.
- ³⁴ Behrens, S.-E., Tomei, L., and De Francesco, R. Identification and properties of the RNA-dependent RNA polymerase of hepatitis C virus. *EMBO J.* 1996; 15: 12-22.
- ³⁵ Lohmann V, Körner F, Herian U, Bartenschlager R Biochemical properties of hepatitis C virus NS5B RNA-dependent RNA polymerase and identification of amino acid sequence motifs essential for enzymatic activity. *J Virol* 1997; 71: 8416-28.
- ³⁶ Baril, M., and Brakier-Gingras, L. (2005). Translation of the F protein of hepatitis C virus is initiated at a non-AUG codon in a +1 reading frame relative to the polyprotein. *Nucleic Acids Res* 33, 1474-1486.
- ³⁷ Neumann, A. U., Lam, N. P., Dahari, H., Gretch, D. R., Wiley, T. E., Layden, T. J. & Perelson, A. S. Hepatitis C viral dynamics in vivo and the antiviral efficacy of interferon-alpha therapy. *Science* 1998; 282, 103-107.
- ³⁸ Flint, M., Thomas, J., Maidens, C., Shotton, C., Levy, S., Barclay, W. & McKeating, J. A. (1999). Functional analysis of cell surface-expressed hepatitis C virus E2 glycoprotein. *Journal of Virology* 1999 73, 6782-6790.
- ³⁹ Jean Dubuisson and Francois-Loic Cosset. Virology and cell biology of the hepatitis C virus life cycle- An update. *Journal of Hepatology* 2014;61: s3-s13.
- ⁴⁰ Bartenschlager, R, Lohman V. Replication of Hepatitis C virus. *J Gen Virol* 2000; 81:1631-48
- ⁴¹ Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet* 2009; 10: 540-50.
- ⁴² Ogata N, Alter HJ, Miller RH, Purcell RH. Nucleotide sequence and mutation rate for the H strain of hepatitis C virus. *Proc Natl Acad Sci U S A* 1991; 88: 3392-3396
- ⁴³ González-Candelas F, López-Labrador FX, Bracho MA. Recombination in hepatitis C virus. *Viruses* 2011; 3: 2006-24.
- ⁴⁴ Pybus OG, Charleston MA, Gupta S, Rambaut A, Holmes EC, et al. The epidemic behavior of the hepatitis C virus. *Science* 2001; 292: 2323–2325.
- ⁴⁵ Smith DB, Bukh J, Kuiken C, Muerhoff AS, Rice CM, Stapleton JT, Simmonds P. Expanded Classification of Hepatitis C Virus Into 7 Genotypes and 67 Subtypes: Updated Criteria and Genotype Assignment Web Resource. *Hepatology* 2014; 59: 318-327.
- ⁴⁶ Jackoviak P, Kuls K, Budzko L, Mania A, Figlerowicz M, Finlerowicz M. Phylogeny and molecular evolution of the hepatitis C virus. *Infect Gen Evol* 2014; 21: 67-82.
- ⁴⁷ Messina JP, Humphreys I, Flaxman A, Brown A, Cooke GS, Pybus OG, Barnes E. Global Distribution and Prevalence of Hepatitis C Virus Genotypes. *Hepatology* 2015; 61: 77-87.
- ⁴⁸ Ansaldi F, Orsi A, Sticchi L, Bruzzone B, Icardi G. Hepatitis C virus in the new era: Perspectives in epidemiology, prevention, diagnostics and predictors of response to therapy. *World Journal of Gastroenterology* 2014;20: 9633-9652.
- ⁴⁹ Alter MJ. Epidemiology of hepatitis C in the West. *Semin Liver Dis* 1995; 15: 5-14.
- ⁵⁰ Simmonds P. The origin of hepatitis C virus. *Curr Top Microbiol Immunol.* 2013; 369: 1-15.

- ⁵¹ World Health Organization. Media Centre: Hepatitis C. July 2015 Available from: <http://www.who.int/mediacentre/factsheets/fs164/en/>.
- ⁵² Alter MJ. Epidemiology of hepatitis C virus infection. *World J Gastroenterol* 2007; 13: 2436-41.
- ⁵³ Lavanchy D. Evolving epidemiology of hepatitis C virus. *Clin Microbiol Infect* 2011; 17: 107-115.
- ⁵⁴ Hepatitis C--global prevalence (update). *Wkly Epidemiol Rec* 1999; 74: 425-427.
- ⁵⁵ Shepard CW, Finelli L, Alter MJ. Global epidemiology of hepatitis C virus infection. *Lancet Infect Dis.* 2000; 5: 558-67.
- ⁵⁶ Williams IT, Bell BP, Kuhnert W, Alter MJ. Incidence and transmission patterns of acute hepatitis C in the United States, 1982-2006. *Arch Intern Med* 2011; 171: 242-248.
- ⁵⁷ Tanaka Y, Hanada K, Orito E, Akahane Y, Chayama K, Yoshizawa H, Sata M, Ohta N, Miyakawa Y, Gojobori T, Mizokami M. Molecular evolutionary analyses implicate injection treatment for schistosomiasis in the initial hepatitis C epidemics in Japan. *J Hepatol* 2005; 42: 47-53.
- ⁵⁸ Di Bisceglie AM. Hepatitis C. *Lancet* 1998, 351: 351-355.
- ⁵⁹ Donahue JG, Muñoz A, Ness PM, Brown DE Jr, Yawn DH, McAllister HA Jr, Reitz BA, Nelson KE. The declining risk of post-transfusion hepatitis C virus infection. *N Engl J Med.*1992; 327: 369-73.
- ⁶⁰ Alter MJ, Hadler SC, Judson FN, Mares A, Alexander WJ, Hu PY, Miller JK, Moyer LA, Fields HA, Bradley DW, et al. Risk factors for acute non-A, non-B hepatitis in the United States and association with hepatitis C virus infection. *JAMA* 1990; 264: 2231-5.
- ⁶¹ De Paschale M, Bevilacqua L, Casiraghi MA, Gualdoni GB, Re T, Viganò P, Baldacci MP, Ferrara A, Novelli C, Mirri P, Gatti C, Albrisi L, Biagiotti S, Chianese R, Rossi U. Distribuzione dei genotipi dell'HCV in relazione all'età e ai fattori di rischio. *La trasfusione di sangue.* 2001; 26: 105-112.
- ⁶² Chiaramonte M, Stroffolini T, Caporaso N, Coppola R, Craxì A, Gaeta GB, Sagnelli E, Zanetti AR. Hepatitis-C virus infection in Italy: a multicentric sero-epidemiological study (a report from the HCV study group of the Italian Association for the Study of the Liver). *Ital J Gastroenterol* 1991; 23: 555-8.
- ⁶³ Marcellusi A, Viti R, Capone A, Mennini FS. The economic burden of HCV-induced diseases in Italy. A probabilistic cost of illness model. *Eur Rev Med Pharmacol Sc* 2015; 19: 1610-1620.
- ⁶⁴ Ansaldi F, Bruzzone B, Salmaso S, Rota MC, Durando P, Gasparini R, Icardi G. Different seroprevalence and molecular, epidemiology patterns of hepatitis C virus infection in Italy. *J Med Virol* 2005; 76: 327-32.
- ⁶⁵ Svrtlih N, Delic D, Simonovic J, Jevtovic D, Dokic L, Gvozdenovic E, Boricic I, Terzic D, Pavic S, Neskovic G, Zerjav S, Urban V. Hepatitis C virus genotypes in Serbia and Montenegro: the prevalence and clinical significance. *World J Gastroenterol* 2007; 13: 355-60.
- ⁶⁶ Hoofnagle JH. Hepatitis C: the clinical spectrum of disease. *Hepatology* 1997; 26: 15S-20S.
- ⁶⁷ McHutchison JG. Understanding hepatitis C. *Am J Manag Care* 2004; 10: S21-9.
- ⁶⁸ Benvenuto L, Pontisso P, Cavalletto D, Noventa F, Chemello L, Alberti A. Lack of correlation between hepatitis C virus genotypes and clinical course of hepatitis C virus-related cirrhosis. *Hepatology* 1997; 25: 211-215.

- ⁶⁹ Moradpour D, Cerny A, Heim MH, Blum HE. Hepatitis C: an update. *Swiss Med Wkly* 2001; 131: 291-8.
- ⁷⁰ Hoofnagle JH, Mullen KD, Jones DB, et al. Treatment of chronic non-A, non-B hepatitis with recombinant Human alfa interferon. A preliminary report. *N Engl J Med* 1986; 315: 1575-1578.
- ⁷¹ Fusco DN, Chung RT. Novel therapies for hepatitis C: insights from the structure of the virus. *Annu Rev Med* 2012; 63: 373-87.
- ⁷² Poynard P, Leroy V, Cohard M, et al. Meta-analysis of interferon randomized trials in the treatment of viral hepatitis C: effects of dose and durations. *Hepatology* 1996; 24: 778-789.
- ⁷³ Rehman S, Ashfaq UA, Javed T. Antiviral drugs against hepatitis C virus. *Genet Vaccines Ther* 2011; 9: 11.
- ⁷⁴ Pawlotsky JM, New antiviral agents for hepatitis C. *F1000 Biol Rep* 2012;4:5.
- ⁷⁵ Wedemeyer H, Dore GJ, Ward JW. Estimates on HCV disease burden worldwide – filling the gaps. *Journal of Viral Hepatitis* 2015; 22, (Suppl. S1), 1–5.
- ⁷⁶ L. Lu, T. Nakano, Y. He, Y. Fu, C.H. Hagedorn, B.H. Robertson. Hepatitis C virus genotype distribution in China: predominance of closely related subtype 1b isolates and existence of new genotype 6 variants. *J Med Virol* 2005; 75: 538-49.
- ⁷⁷ BioEdit Home Page <http://www.mbio.ncsu.edu/BioEdit/page2.html>.
- ⁷⁸ PHYLIP Home Page <http://evolution.genetics.washington.edu/phylip.html>.
- ⁷⁹ Posada, D. jModelTest: phylogenetic model averaging. *Mol Biol Evol* 2008; 25:1253-1256.
- ⁸⁰ Drummond AJ, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 2007; 7: 214-222.
- ⁸¹ Pybus OG, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nature* 2009; 10: 540-550.
- ⁸² Tsan-Yuk Lam T, et al. Use of phylogenetics in the molecular epidemiology and evolutionary studies of viral infection. *Clin Lab Sci* 2010; 47: 5-49.
- ⁸³ Bromham L, Penny D. The modern molecular clock. *Nat Rev Genet* 2003; 4: 216-224.
- ⁸⁴ Rutschmann F. Molecular Dating of phylogenetic trees: a brief review of current methods that calculate divergence times. *Divers Distrib* 2006; 12: 35-48.
- ⁸⁵ Kass R.E., Raftery A.E. Bayes factor. *J Am Stat Assoc* 1995; 90: 773-795.
- ⁸⁶ Lanave C., Preparata G., Saccone C., Serio G. A new method for calculating evolutionary substitution rates. *J Mol Evol* 1984; 20(1): 86-93.
- ⁸⁷ TRACER Home Page: <http://beast.bio.ed.ac.uk/tracer>
- ⁸⁸ Ciccozzi M, Ciccaglione AR, Lo Presti A, Yalcinkaya T, Taskan ZP, Equestre M, Costantino A, Bruni R, Ebranati E, Salemi M, Gray R, Rezza G, Galli M, Zehender G. Reconstruction of the evolutionary dynamics of the hepatitis C virus 1b epidemic in Turkey. *Infect Genet Evol* 2011; 11: 863-8.
- ⁸⁹ Esteban JI, Sauleda S, Quer J. The changing epidemiology of hepatitis C virus infection in Europe. *J Hepatol* 2008; 48(1):148-62.
- ⁹⁰ Gower E, Estes C, Blach S, Razavi-Shearer K, Razavi H. Global epidemiology and genotype distribution of the hepatitis C virus infection. *J Hepatol*. 2014; 61(1 Suppl): S45-57.

- ⁹¹ Chironna M, Germinario C, Lopalco PL, Carrozzini F, Quarto M. Prevalence of hepatitis virus infections in Kosovar refugees. *Int J Infect Dis.* 2001; 5(4): 209-13.
- ⁹² Pybus OG, Charleston MA, Gupta S, Rambaut A, Holmes EC, Harvey PH. The epidemic behaviour of the hepatitis C virus. *Science* 2001; 292(5525): 2323-5.
- ⁹³ Magiorkinis G, Magiorkinis E, Paraskevis D, Ho SY, Shapiro B, Pybus OG, Allain JP, Hatzakis A. The global spread of hepatitis C virus 1a and 1b: a phylodynamic and phylogeographic analysis. *PLoS Med* 2009; 6(12):e1000198.
- ⁹⁴ Ernest Drucker. Injectable heroin substitution treatment for opioid dependency. *The Lancet* 2001; 27: 1385.
- ⁹⁵ Sarang A, Stuijckte R, Bykov R. Implementation of harm reduction in Central and Eastern Europe and Central Asia. *Int J Drug Policy* 2007; 18(2): 129-35.
- ⁹⁶ Victoria L Demetriou and Leondios G Kostrikis Molecular epidemiology of hepatitis C infection in Cyprus within the general population and high-risk cohorts. *BMC Research Notes* 2011; 4:468.
- ⁹⁷ <http://www.rastko.rs/istorija/srbi-balkan/vrudic-montenegro.html>
- ⁹⁸ Khan A et al. Epidemic spread of hepatitis C virus genotype 3a and relation to high incidence of hepatocellular carcinoma in Pakistan. *J Med Virol.* 2009; 81(7):1189-97.
- ⁹⁹ Nelson PK, Mathers BM, Cowie B, Hagan H, Des Jarlais D, Horyniak D, Degenhardt L. Global epidemiology of hepatitis B and hepatitis C in people who inject drugs: results of systematic reviews. *Lancet* 2011; 13378(9791):571-83.
- ¹⁰⁰ Bastie A, Pawlotsky JM, Roudot-Thoraval F, Dhumeaux D. Hepatitis C virus infection. *Epidemiology. Pathol Biol* 1995;43(8): 674-80.
- ¹⁰¹ Morice Y, Cantaloube JF, Beaucourt S, Barbotte L, De Gendt S, Goncales FL, Butterworth L, Cooksley G, Gish RG, Beaugrand M, Fay F, Fay O, Gonzalez JE, Martins RM, Dhumeaux D, Vanderborcht B, Stuyver L, Sablon E, de Lamballerie X, Pawlotsky JM. Molecular epidemiology of hepatitis C virus subtype 3a in injecting drug users. *J Med Virol.* 2006;78(10): 1296-303.
- ¹⁰² Pybus OG, Cochrane A, Holmes EC, Simmonds P. The hepatitis C virus epidemic among injecting drug users. *Infect Genet Evol* 2005; 5(2): 131-9.
- ¹⁰³ Matheï C, Van Dooren S, Lemey P, Van Damme P, Buntinx F, Vandamme AM. The epidemic history of hepatitis C among injecting drug users in Flanders, Belgium. *J Viral Hepat* 2008; 15(6): 399-408.
- ¹⁰⁴ Ciccozzi M, Zehender G, Cento V, Lo Presti A, Teoharov P, Pavlov I, Bogdanova V, Perno CF, Ciotti M. Molecular analysis of hepatitis C virus infection in Bulgarian injecting drug users. *J Med Virol.* 2011; 83(9): 1565-70.
- ¹⁰⁵ Kamal SM, Nasser IA. Hepatitis C genotype 4: What we know and what we don't yet know. *Hepatology* 2008; 47(4): 1371-83.
- ¹⁰⁶ Delwaide J, Reenaers C, Gerard C, Vaira D, Bastens B, Servais B, Bekhti A, Bataille C, Wain E, De Leeuw P, Daenen G, Mesureur T, Sente JM, Belaïche J; Groupe Liégeois d'Etude des Virus Hépatotropes. HCV genotype 4 in Belgium: three distinct patterns among patients from European and African origin. *Eur J Gastroenterol Hepatol* 2006; 18(7): 707-12.

¹⁰⁷ Matera G, Lamberti A, Quirino A, Focà D, Giancotti A, Barreca GS, Guadagnino V, Liberto MC. Changes in the prevalence of hepatitis C virus (HCV) genotype 4 in Calabria, Southern Italy. *Diagn Microbiol Infect Dis.* 2002; 42(3): 169-73.