

UNIVERSITÀ DEGLI STUDI DI MILANO

SCUOLA DI DOTTORATO

in

"MALATTIE INFETTIVE"

DIPARTIMENTO

di Scienze Biomediche e Cliniche "L. Sacco"

XXVII CICLO



ORIGINE, EPIDEMIOLOGIA E FILOGEOGRAFIA DEI GENOTIPI D E A
DEL VIRUS DELL'EPATITE B

Coordinatore

Prof.ssa Antonella D'ARMINIO MONFORTE

Tutor

Dr. Gianguglielmo ZEHENDER

Tesi di dottorato di:

Elena GABANELLI

Matricola n.R09715

Anno Accademico 2013/2014

UNIVERSITÀ DEGLI STUDI DI MILANO

SCUOLA DI DOTTORATO
in
"MALATTIE INFETTIVE"

DIPARTIMENTO
di Scienze Biomediche e Cliniche "L. Sacco"

XXVII CICLO



ORIGINE, EPIDEMIOLOGIA E FILOGEOGRAFIA DEI GENOTIPI D E A
DEL VIRUS DELL'EPATITE B

Coordinatore

Prof.ssa Antonella D'ARMINIO MONFORTE

Tutor

Dr. Gianguglielmo ZEHENDER

Tesi di dottorato di:

Elena GABANELLI

Matricola n.R09715

Anno Accademico 2013/2014

INDICE

1. INTRODUZIONE	1
1.1 IL VIRUS DELL'EPATITE B	2
1.1.1 Classificazione e struttura del virus	2
1.1.2 Il genoma virale	4
1.1.2.1 Regione pre-S/S	5
1.1.2.2 Regione pre-C/C	8
1.1.2.3 Gene Pol	9
1.1.2.4 Gene X	10
1.1.3 Il ciclo vitale	11
1.1.4 Patogenesi	14
1.1.5 Storia naturale dell'infezione	16
1.2 EPIDEMIOLOGIA DI HBV	19
1.2.1 Prevalenza, distribuzione e modalità di trasmissione	19
1.2.2 Numero riproduttivo di base (R_0)	22
1.3 ASPETTI TERAPEUTICI	22
1.4 VARIABILITÀ GENETICA DI HBV	28
1.4.1 Sottotipi	28
1.4.2 Genotipi e sottogenotipi	29
1.4.3 Mutanti	34
1.4.4 Genotipi ricombinanti	40
1.5 IMPLICAZIONI CLINICHE DEI GENOTIPI DI HBV	42
1.6 COINFEZIONE HBV-HIV	45
1.7 INFEZIONE DA HBV IN ITALIA	47
1.8 FILODINAMICA E FILOGEOGRAFIA	49
1.8.1 Filodinamica del genotipo D di HBV	51
1.8.2 Filodinamica del genotipo A di HBV	52
2. SCOPO DELLO STUDIO	55

3. MATERIALI E METODI	57
3.1 CASISTICA	58
3.2 METODI	58
3.2.1 Modalità di raccolta, preparazione e conservazione dei campioni	58
3.2.2 Estrazione del DNA	58
3.2.3 Amplificazione del DNA virale in regione PRE-S/S E P mediante nested-PCR	60
3.2.4 Rilevamento dei prodotti di PCR	64
3.2.5 Purificazione dei prodotti di PCR	65
3.2.6 Sequenziamento delle regioni genomiche amplificate	66
3.2.6.1 Reazione di sequenza	66
3.2.6.2 Purificazione dei prodotti di sequenza	67
3.2.6.3 Sequenziamento	68
3.2.6.4 Analisi dei dati di sequenza	68
3.2.7 Creazione dei dataset	69
3.2.8 Analisi filogenetica	69
3.2.8.1 Discriminazione tra i genotipi mediante strumenti disponibili in Rete	70
3.2.8.2 Genotipizzazione delle sequenze in studio	70
3.2.8.3 Analisi dei ricombinant con Simplot	72
3.2.8.4 Likelihood mapping	72
3.2.8.5 Analisi filogenetica mediante i metodi Bayesiani	73
3.2.9 Analisi filodinamica	75
3.2.10 Analisi filogeografica	76
3.2.11 Analisi statistica	77
4. RISULTATI	78
4.1 CARATTERISTICHE DEMOGRAFICHE DELLA POPOLAZIONE IN STUDIO	79
4.2 CARATTERIZZAZIONE E DISTRIBUZIONE DEL GENOTIPO VIRALE	80
4.2.1 Distribuzione dei sottogenotipi di HBV nella popolazione in studio	83
4.2.2 Associazione tra fattore di rischio e genotipo/sottogenotipo virale	83
4.3 ANALISI DEL GENOTIPO D	85
4.3.1 Costruzione del dataset	85
4.3.2 Likelihood mapping	85

4.3.3 Analisi filodinamica	87
4.3.4 Analisi filogeografica	90
4.4 ANALISI DEL GENOTIPO A	94
4.4.1 Costruzione del dataset	94
4.4.2 Likelihood mapping	95
4.4.3 Stima dell'evolutionary rate dei sottogenotipi di HBV-A	97
4.4.4 Analisi filogenetica e filodinamica del sottogenotipo A2	98
4.4.5 Analisi filogenetica e filogeografica dei sottogenotipi A1, A3 e A6	101
5. DISCUSSIONE	106
Appendice A	117
Appendice B	120
6. BIBLIOGRAFIA	123

1. Introduzione

1.1 IL VIRUS DELL'EPATITE B

1.1.1 Classificazione e struttura del virus

Il virus dell'epatite B (HBV) è il prototipo della famiglia *Hepadnaviridae* (genere *Orthohepadnavirus*), che comprende virus epatotropi a DNA, responsabili dell'insorgenza di malattie epatiche acute e croniche ¹.

La particella virale completa (o particella di Dane) è una struttura sferica, di 42 nm di diametro, costituita da un involucro esterno di natura lipoproteica, *envelope*, e da una particella interna, *core* o *nucleocapside*, di circa 30 nm.

Il *core* è formato da un rivestimento proteico (*capside*), composto dall'*antigene del core* (HBcAg), che racchiude il DNA virale, una DNA polimerasi, responsabile della sintesi di DNA virale nelle cellule infettate, ed una proteina kinasi.

L'*envelope* è un doppio strato fosfolipidico che contiene tre proteine sintetizzate nel citoplasma degli epatociti infetti ^{2,3}, le proteine *large*, *middle* e *small* (*major*), le quali costituiscono i tre antigeni virali di superficie. La proteina *small*, conosciuta un tempo come antigene Australia, è l'*HBsAg* (Figura 1).

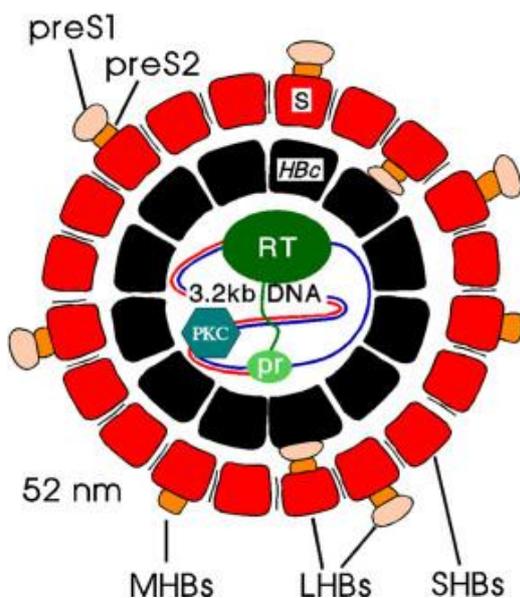


Figura 1. Struttura del virus dell'epatite B.

Nella circolazione sanguigna di pazienti infettati col virus dell'epatite B si osservano, oltre alle particelle di Dane, le particelle subvirali, che presentano forma sferica o filamentosa, con un diametro di 20-22 nm e una lunghezza variabile tra i 40-400 nm per i filamenti (Figura 2).

Esse sono costituite dal solo envelope, che contiene gli antigeni virali di superficie. Non sono infettanti perchè sono prive dall'acido nucleico e sono conseguenza di un'eccessiva produzione di proteine di superficie durante la replicazione ⁴.

La proteina S rappresenta il costituente predominante di tali particelle e il tipo di particella formata sembra essere determinato dal rapporto proteina S/proteina L, coassemblate durante la morfogenesi. In particolare particelle sferiche contengono solo alcune tracce della proteina L, mentre il coassemblamento di un'ampia porzione di proteina L con la proteina S determina la formazione di filamenti ⁴.

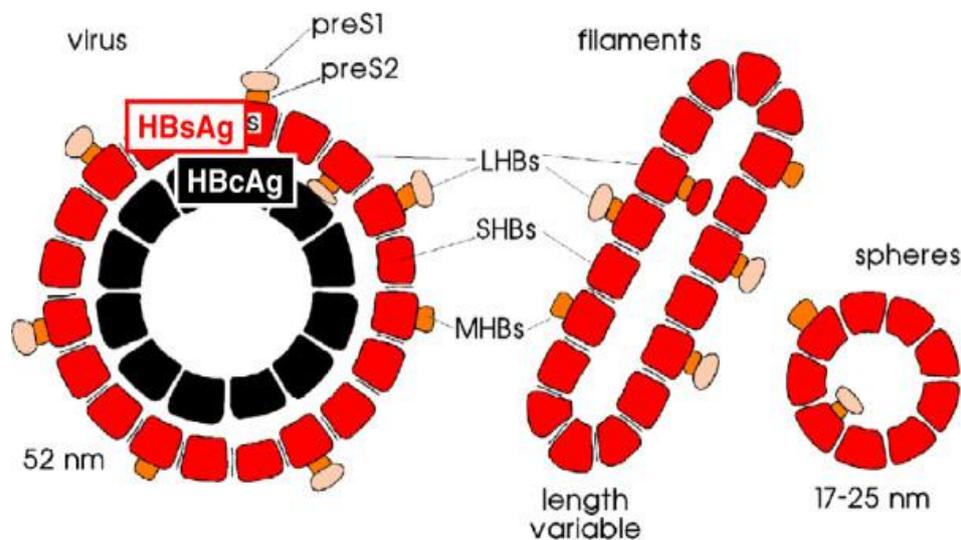


Figura 2. Virus dell'epatite B (particella di Dane) e filamenti o particelle sferiche costituite solo dall'envelope.

L'autoassemblamento delle particelle subvirali si verifica in un compartimento citoplasmatico intermedio post-ER e pre-Golgi (*ERGIC, ER-Golgi intermediate compartment*), creando delle vescicole che vengono rilasciate dalla cellula infetta ⁵.

I filamenti, assemblati nel lume del reticolo endoplasmatico (ER) e trasportati all'ERGIC, sono lunghi e ramificati e forma e dimensione limitano la loro progressione attraverso il pathway secretorio. Tale progressione richiede la conversione in particelle sferiche ⁴.

Il rapporto tra particelle subvirali e virus completo nel siero dei pazienti affetti dall'epatite B varia da 1000:1 a 10,000:1 ⁶.

Il ruolo delle particelle subvirali nel ciclo vitale di virus dell'epatite B non è chiaro. Queste particelle possono interagire con il sistema immune dell'ospite oppure favorire il processo infettivo ⁷.

1.1.2 Il genoma virale

Il genoma di HBV è costituito da una molecola circolare di DNA di circa 3200 paia di basi (da 3182 a 3248 nucleotidi a secondo del genotipo) ⁸, parzialmente a doppia elica. L'acido nucleico virale presenta una caratteristica peculiare, possedendo un filamento più lungo dell'altro. Il filamento breve S (+) ha una lunghezza variabile di 1700-2800 nucleotidi ed è solo il 50-80% del filamento lungo complementare, denominato L o filamento (-), la cui lunghezza varia da 3020 a 3320 nucleotidi. I due filamenti mantengono una configurazione circolare grazie all'appaiamento di 240 nucleotidi in posizione 5' ² (*cohesive end region*).

La sintesi del filamento a polarità positiva, per mezzo della polimerasi endogena, permette, in corso di replicazione, il completamento del tratto a catena singola e la conseguente produzione di un DNA a doppia elica completo, superavvolto e chiuso covalentemente ^{9,10}.

Il genoma di HBV presenta un'organizzazione molto compatta: la catena L infatti contiene quattro sequenze geniche ampiamente sovrapposte, denominate S, C, P e X, che coprono l'intero genoma. Esse rappresentano le quattro *open reading frames* [ORF], codificate per sette proteine distinte del virus. Al contrario, la catena S sembra non contenere ORF (Figura 3).

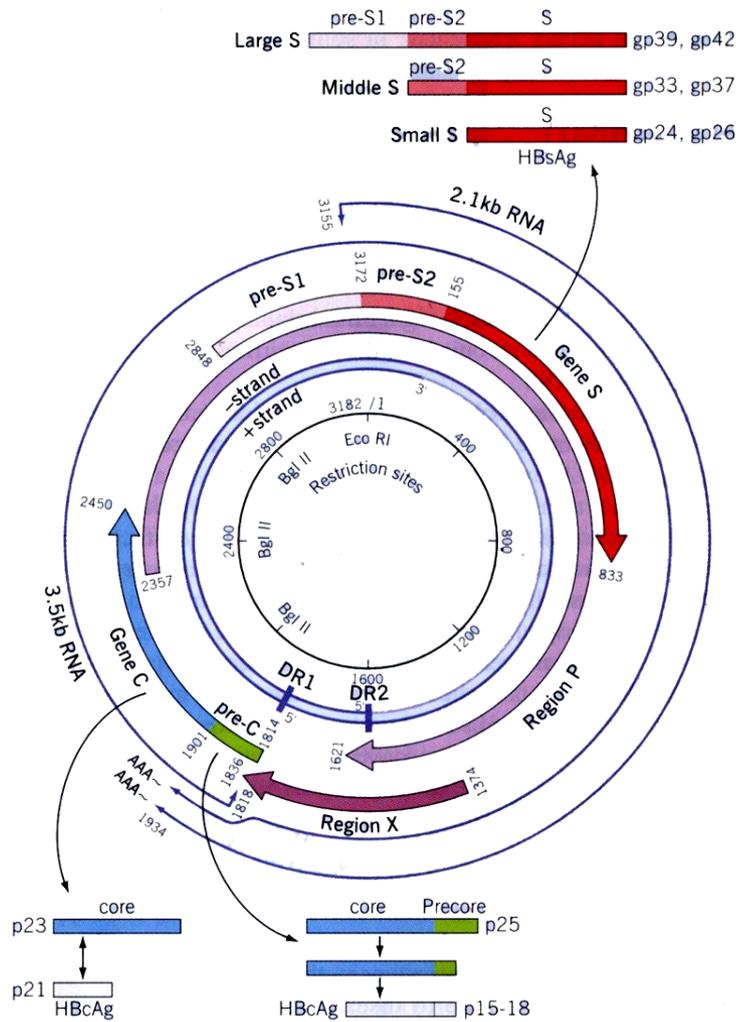


Figura 3. Struttura genomica del virus dell'epatite B.

1.1.2.1 Regione pre-S/S

Il gene *pre-S/S* è diviso in tre regioni: pre-S1 (108-119 aa in relazione al genotipo)¹¹, pre-S2 (55 aa) e il gene S (226 aa)¹². Queste codificano per i tre antigeni virali di superficie (proteine *large*, *middle* e *small*) mediante l'uso di codoni alternativi di inizio per la trascrizione. La presenza di tre diversi *start codon* e di un unico *stop codon* sono infatti alla base della sintesi delle tre differenti proteine dell'envlope, che condividono la porzione carbossiterminale.

La proteina *large* [L] (p39 e gp42) è codificata dalle regioni pre-S1, pre-S2 e dal gene S ed ha una lunghezza variabile in relazione al sottotipo/genotipo di HBV.

Essa costituisce solo il 20% delle proteine dell'envelope ed è la meno rappresentata a livello della superficie virale.

Tale proteina risulta essenziale per l'infezione e per la morfogenesi virale essendo implicata sia nel legame del virus ai recettori cellulari, nell'assemblaggio del virione e nel suo rilascio dalla cellula, che nella sollecitazione della risposta cellulare B e T¹³.

La sua presenza nel sangue della persona infetta è strettamente correlata con l'antigene *HBeAg* e con la presenza di HBV-DNA, due markers sierologici dell'infezione acuta¹⁴.

La proteina middle [*M*] (p30, gp33 e gp36)^{11,15,16} è codificata dalla regione pre-S2 e dal gene S. Essa non è essenziale per l'assemblaggio ed il rilascio del virus, ma gioca un ruolo prioritario per la penetrazione del virus nell'epatocita, poiché contiene un recettore per l'albumina umana (aa 3 e aa 16).

La proteina small [*S*], l'antigene S o *HBsAg* (p24 e gp26) è codificata dal gene S e rappresenta la più piccola tra le proteine di superficie. La sua denominazione come proteina maggiore (*major*) è da imputare alla sua presenza dal punto di vista quantitativo, sia a livello delle particelle difettive, rilasciate nella circolazione sanguigna, sia a livello del virione completo.

L'*HBsAg* appare molto precocemente nel siero delle persone infettate nel corso dell'infezione, costituendo il primo marker sierologico dell'infezione, e prevale nel siero dei portatori cronici.

L'antigene S presenta due loop: il *cytosolic loop*, tra gli aminoacidi 29 e 80, il quale interagisce con le proteine del core, e l'*immunodominant loop*, tra gli aminoacidi 101 e 163, localizzato all'interno del lume del reticolo endoplasmico e successivamente ripresentato sulla superficie di virioni secreti. All'interno della struttura dell'*immunodominant loop*, tra gli aminoacidi 124 e 147, si trova l'epitopo antigenico maggiore, denominato *determinante "a"*, che è il determinante antigenico comune a tutti i sottotipi. Tale regione è caratterizzata da una struttura a 2 loop (uno maggiore e uno minore) e contiene 8 residui cisteinici responsabili della formazione di legami disulfidici intramolecolari o intermolecolari¹⁷ (Figura 4).

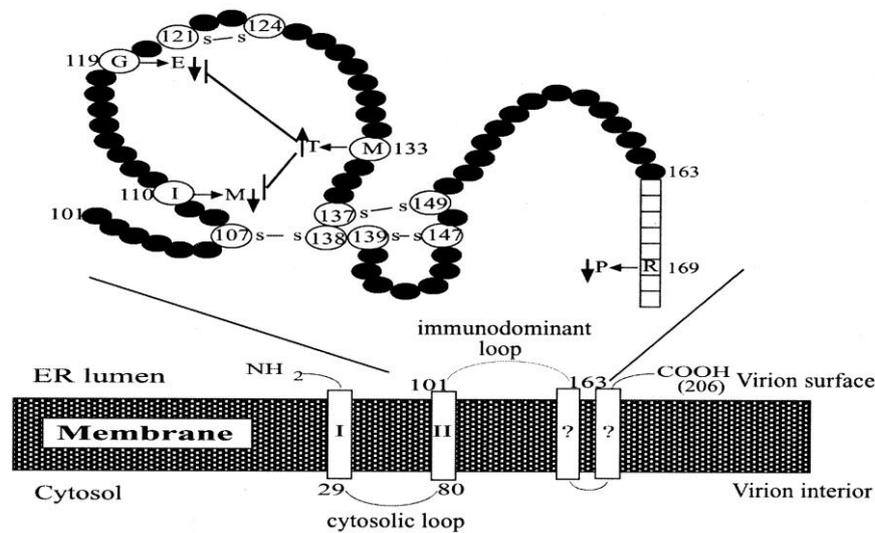


Figura 4. Struttura dell'HBsAg. Il pannello in basso rappresenta la proteina S, di 206 aminoacidi, la quale attraversa la membrana del reticolo endoplasmico, probabilmente per quattro volte ¹⁸.

Esistono poi due coppie di determinanti tra loro rispettivamente esclusivi, *d/y* e *w/r*, i quali si distinguono dalla presenza dell'aminoacido lisina o arginina in posizione 122 (*d* oppure *y*) e 160 (*w* oppure *r*) della proteina S. Se l'aminoacido nella posizione 122 è Arg (122R), il determinante sarà *y*, se è Lys (122K) il determinante sarà *d* ¹⁹. Allo stesso modo, 160R definisce il determinante *r* e 160K il *w* ¹⁹.

Successivamente, sono stati descritti quattro sottodeterminanti di *w* (*w1-w4*) di cui sono responsabili le mutazioni in posizione 127: gli aminoacidi Pro, Thr e Leu codificano rispettivamente per le varianti *w1/w2*, *w3* e *w4* ²⁰.

Un'altro determinante *q* è stato espresso dalla gran parte dei ceppi HBV appartenenti ai sottotipi *adw4* e *adr* ^{21, 22}. Si suppone che la responsabilità per quest'ultimo sia attribuibile ai residui 177 e 178 ²³.

La combinazione fra determinanti e sottodeterminanti definisce i vari sottotipi di HBV.

Il ruolo essenziale del determinante "*a*" è di scatenare la risposta immune, sia naturale che indotta dalla vaccinazione, dell'intera proteina ²⁴. Per tale motivo gli

attuali vaccini anti-epatite B sono costituiti dall'antigene S. La presenza di soli anticorpi *anti-HBs* è sufficiente per la protezione di un individuo contro l'infezione. Le mutazioni a livello della sequenza nucleotidica del *determinante "a"* danno luogo a varianti di HBV che "sfuggono" alla vaccinazione (*escape mutant*), ossia producono infezione in soggetti che hanno titoli protettivi di *anti-HBs*³.

1.1.2.2 Regione pre-C/C

La regione *pre-C/C* codifica l'antigene del core (*HBcAg*) e la proteina di secrezione (*HBeAg*), distinte a seconda del sito trascrizionale d'inizio. Tale regione è costituita dal gene C (183 aa), preceduto da una breve regione di 30 aminoacidi denominata *pre-C*.

Il trascritto della sola regione C (183 aa) codifica l'antigene del core (*HBcAg*), la proteina strutturale del capsido virale di 21 kd, all'interno della cui struttura terziaria si riconoscono due domini, il primo coinvolto nell'assemblaggio del nucleocapside, il secondo nell'encapsidazione e replicazione virale¹².

La trascrizione, a partire dalla sequenza *pre-C* (30 aa), comporta la sintesi della proteina pre-core di 24 Kd, costituita dalla proteina del core più una sequenza addizionale di 29 aa. I primi 19 aa rappresentano una sequenza segnale che indirizzano la proteina pre-core verso la via secretoria; durante il passaggio, attraverso il reticolo endoplasmico, tale sequenza viene lisata da una peptidasi cellulare provocando la trasformazione della proteina Pre-core P25 nella proteina P22. Quest'ultima viene poi secreta, attraverso l'apparato di Golgi, dove viene ulteriormente clivata in corrispondenza dell'estremità C terminale generando l'antigene solubile *HBeAg* di 16 Kd^{1,12}.

L'*HBcAg* è il principale componente strutturale del nucleocapside e possiede in posizione carbossiterminale un dominio di 35 residui aminoacidici, ricco in arginina, ad elevata affinità per l'RNA e quindi favorente la formazione del capsido²⁵.

L'antigene del core *HBcAg* non è svelabile in circolo perchè è ricoperto dall'antigene di superficie *HBsAg*. Quando l'*HBcAg* viene espresso sulla superficie

degli epatociti infettati, è il principale bersaglio della risposta immunitaria cellulo-mediata finalizzata alla distruzione degli epatociti infettati ³.

Gli anticorpi *HBcAb* diretti contro *HBcAg* si riscontrano in tutti i soggetti esposti al virus; essi non sono protettivi e non permettono la distinzione tra evento acuto e cronico, rinvenendosi sia nel soggetto guarito che nel portatore/malato cronico. La distinzione tra evento acuto e cronico si basa invece sul riscontro, nel primo caso, degli anticorpi anti-core di tipo IgM, che non permettono però la discriminazione tra epatite acuta e riacutizzazione in corso di epatopatia cronica ²⁶.

L'*HBeAg* non è un componente della particella virale, ma è presente nel siero del soggetto infetto ed è un marcatore di attiva replicazione e di infettività; viene inoltre espresso sulla superficie degli epatociti insieme all'*HBsAg* ed è un importante bersaglio della risposta immune dell'ospite, che porta alla distruzione della cellula infetta. L'eliminazione di *HBeAg* e la concomitante comparsa di anticorpi specifici si associa, salvo poche eccezioni rappresentate dalla selezione di mutanti *HBeAg* minus, al controllo immunitario dell'infezione virale.

Poiché l'*HBeAg* viene sintetizzato e secreto in circolo in grande quantità, sembra abbia la funzione di "distrarre" la risposta immunitaria (anticorpi e linfociti T citotossici) affinché non raggiunga il vero bersaglio, costituito dalla cellula infettata.

1.1.2.3 Gene Pol

Il *gene Pol* (832 aa) costituisce l'80% del genoma virale e comprende quattro domini distinti, che codificano per enzimi essenziali al ciclo di replicazione virale.

Il gene *Pol* codifica infatti la *polimerasi virale*, una proteina multifunzionale costituita dal dominio di *priming*, dal dominio di *spacer*, dal dominio di trascrittasi inversa, che catalizza la sintesi del DNA/RNA-dipendente, e dal dominio RNAsi-H che degrada l'RNA dell'ibrido DNA-RNA.

A differenza della polimerasi retrovirale, la polimerasi di HBV è priva di un dominio di integrazione.

Tale poliproteina presenta in porzione amino-terminale la DNA polimerasi-RNA dipendente, la quale si lega covalentemente all'estremità 5' del filamento di DNA(-) portando alla formazione di un ibrido DNA-RNA.

In posizione carbossi-terminale è invece localizzata la RNA-asi, proteina che porta alla proteolisi dell'ibrido, permettendo la sintesi del filamento complementare (+)¹⁰.

1.1.2.4 Gene X

Il gene X ha come prodotto una proteina che sembra avere una funzione regolatrice, esprimendo una funzione transattivante nei confronti delle altre proteine virali. E' una proteina non-strutturale (*HBx*), di 154 aminoacidi, che sembra possa attuare la sua proprietà transattivante anche nei confronti di altri geni cellulari giocando così un ruolo importante nel processo di epatocarcinogenesi ²⁷.

La proteina X è essenziale per l'infezione di HBV e desta una particolare attenzione per la sua funzione pleiotropica. *HBx* non lega direttamente il DNA ma agisce tramite un'interazione proteina-proteina ed attiva importanti fattori trascrizionali, indispensabili per la replicazione del DNA virale ²⁸.

L'infezione da epatite cronica di tipo B è associata ad un aumento della produzione di ROS (*reactive oxygen species*), a sua volta collegato a processi infiammatori, dentro agli epatociti, che sono responsabili dell'ossidazione delle macromolecole intracellulari.

I ROS possono danneggiare il DNA inducendo mutazioni (stress ossidativo), indurre l'attivazione di importanti fattori di trascrizione, quali il promotore della crescita, della proliferazione e della replicazione del DNA, e coinvolgere altre funzioni che inducono a morte cellulare o allo sviluppo di tumore.

HBx sembra attivare fattori di trascrizione cellulare, attraverso lo stress ossidativo nelle cellule, tramite la sua diretta associazione con i mitocondri portando all'innalzamento dei livelli di ROS. Infatti *HBx* interagisce fisicamente e direttamente con un canale mitocondriale esterno voltaggio-dipendente, causando una diminuzione del potenziale mitocondriale di membrana e l'incremento di ROS.

Un livello aumentato di ROS può aumentare la replicazione virale, nonostante il meccanismo attraverso il quale la proteina X attivi questi fattori di trascrizione non sia ancora completamente chiarito ²⁸.

1.1.3 Il ciclo vitale

HBV, insieme a tutti gli *Hepadnavirus*, impiega un metodo indiretto per la sua replicazione, la quale richiede, come intermedio, un filamento a RNA.

Il ciclo replicativo del virus inizia con l'interazione tra proteine virali dell'envelope e specifici recettori cellulari presenti sulla membrana plasmatica degli epatociti. La regione N-terminale, tra gli aminoacidi 21 e 47, della proteina grande dell'envelope (regione *pre-S1*) sembra essere essenziale nel mediare tale processo. A questo segue la fusione delle membrane cellulari, l'internalizzazione del virione mediante endocitosi, il quale viene privato dell'envelope, e la traslocazione del genoma virale nel nucleo cellulare (Figura 5).

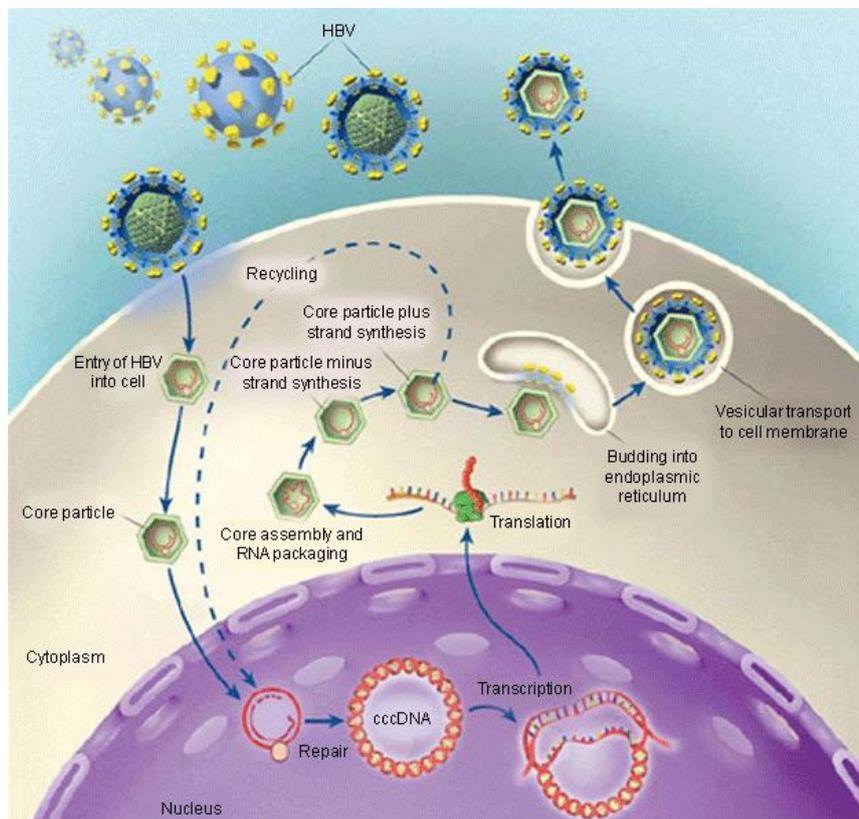


Figura 5. Ciclo vitale di HBV ¹.

A livello nucleare il genoma virale, costituito da DNA circolare parzialmente a doppia elica (*rc-DNA, relaxed circular DNA*), viene convertito in DNA circolare a doppio filamento (*ccc-DNA, covalently closed circular*) grazie all'attività di enzimi cellulari, indipendentemente dalla polimerasi virale ²⁹.

Il ccc-DNA funge da stampo per la trascrizione di trascritti genomici e subgenomici, ad opera dell'enzima RNA polimerasi II della cellula ospite; questi vengono poi trasportati a livello citoplasmatico dove ha luogo la sintesi delle proteine funzionali e strutturali del virione ³⁰.

I trascritti subgenomici di 2,4 Kb, 2,1 Kb e 0,7 Kb codificano le proteine dell'envelope (*HBsAg*) e la proteina X. I due trascritti di 3,5 Kb costituiscono l'mRNA precore, che codifica la proteina *pre-C*, processata proteoliticamente a dare l'antigene e (*HBeAg*), e l'mRNA pregenomico, che svolge la duplice funzione di codificare la proteina del core e la polimerasi e fungere da stampo per la trascrittasi inversa (Figura 6).

Le proteine dell'envelope e quella del core vengono sintetizzate ed assemblate in modo autonomo le une dalle altre: le prime nel reticolo endoplasmico ruvido, la seconda a livello citoplasmatico.

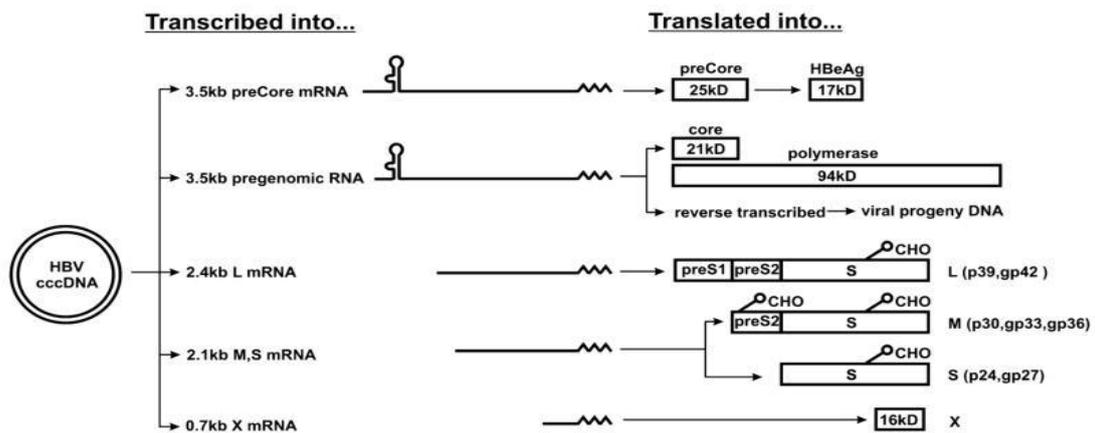


Figura 6. mRNA di HBV prodotti dalla trascrizione del ccc-DNA e proteine per cui essi specificano ³¹.

L'RNA pregenomico (*pgRNA*) e la proteina Pol vengono incorporati all'interno di proteine del core, assemblatesi a formare il nucleocapside. Quest'incorporazione è indotta direttamente dall'*HBV-RNA encapsidation signal epsilon*, che consiste di una regione terminale dell'RNA pregenomico di circa 130 nucleotidi in posizione 5', con una struttura stem-loop. Le mutazioni in questa regione portano ad una replicazione meno efficiente del virus o alla creazione di particelle virali non vitali³².

All'interno del nucleocapside, un singolo filamento dell'RNA pregenomico viene retrotrascritto dalla polimerasi virale RNA-dipendente, che in questo caso assume il ruolo di trascrittasi inversa, in un filamento di DNA genomico a polarità negativa, generando un ibrido DNA-RNA. Non appena quest'ultimo si è formato, vi è una degradazione dell'RNA pregenomico ad opera di una RNasi H, la quale lascia intatta solo una piccola porzione di *pgRNA* all'estremità 5'³³. Questo oligoribonucleotide non digerito, che comprende la sequenza DR1 (direct repeat 1) e 6 o 7 nucleotidi all'estremità 5' di DR1, viene traslocato e legato alla regione DR2 (direct repeat 2) di DNA, in posizione terminale 5'³¹. Successivamente, l'HBV polimerasi genera il filamento complementare (+) del DNA, utilizzando questo oligoribonucleotide come primer e usando il filamento (-) come stampo³⁰. La catena positiva rimane incompleta poiché la stessa polimerasi è, in seguito, privata dei precursori trifosfati.

Il nucleocapside, così costituito, viene internalizzato nel reticolo endoplasmico dove acquisisce le proteine dell'envelope, costituite dall'antigene di superficie virale e infine il virione maturo fuoriesce dalla cellula epatica. Al momento della fuoriuscita del virus, il filamento positivo cessa immediatamente di allungarsi. Questo meccanismo spiega la variabilità nella sua lunghezza.

Un aspetto particolare del ciclo replicativo di HBV, che sta alla base dell'infezione persistente, è che una piccola parte di nucleocapsidi non viene indirizzata verso la via secretoria, torna invece nel nucleo dove il DNA genomico viene convertito in *ccc-DNA*. Ciò permette di creare una riserva stabile di molecole di *ccc-DNA*, che fungano da stampo trascrizionale, senza la necessità di multiple reinfezioni³⁴.

1.1.4 Patogenesi

Ancora oggi non è stato completamente chiarito l'esatto meccanismo patogenetico delle epatiti virali acute e croniche. Poiché per nessuno dei virus epatitici maggiori (HAV, HBV, HCV, HDV, HEV) è stato inequivocabilmente dimostrato un effetto citopatico diretto, si ritiene che un ruolo fondamentale sia svolto dalla risposta immunitaria dell'organismo umano nei confronti dei virus stessi.

Le numerose ricerche condotte negli ultimi anni hanno permesso di far luce sulla patogenesi dell'epatite B e soprattutto sui meccanismi responsabili della cronicizzazione dell'infezione.

Il danno epatico è provocato dalla risposta immunitaria cellulo-mediata diretta verso gli antigeni del nucleocapside virale (*HBcAg*, *HBeAg*), espressi sulla superficie degli epatociti infettati insieme agli antigeni maggiori di istocompatibilità di classe I (*HLA-I*)³.

Dopo l'infezione, il virus B raggiunge gli epatociti e replica; le particelle complete, assieme all'eccesso di *HBsAg*, vengono liberate dall'epatocita in assenza di ogni effetto citopatico e passano nella circolazione sanguigna. Qui vengono riconosciute da macrofagi e linfociti natural killer (*NK*), componenti dell'immunità naturale, che rappresenta il primo meccanismo di protezione antivirale in corso di infezione acuta. Gli epatociti infetti producono $\text{IFN-}\alpha$ e $\text{IFN-}\gamma$, tramite i quali ha luogo il richiamo intraepatico di cellule dell'immunità innata, responsabili a loro volta della secrezione di grandi quantità di citochine. Tali citochine determinano una up-regulation dell'espressione di molecole *HLA* di classe I sulla superficie epatocitaria³⁵.

Alla fase sintomatica dell'infezione acuta corrisponde l'attivazione della risposta cellulare specifica (Figura 7)¹, sostenuta da linfociti T citotossici CD8^+ e T helper CD4^+ . I primi riconoscono gli antigeni *HBcAg* e *HBsAg* presentati sulla superficie cellulare dell'epatocita infetto, associati a molecole *MHC-I*³⁶. Questa può comportare la lisi diretta dell'epatocita infetto e il rilascio di citochine quali $\text{IFN-}\gamma$ e $\text{TNF-}\alpha$ in grado di down-regolare la replicazione virale negli epatociti circostanti³⁷.

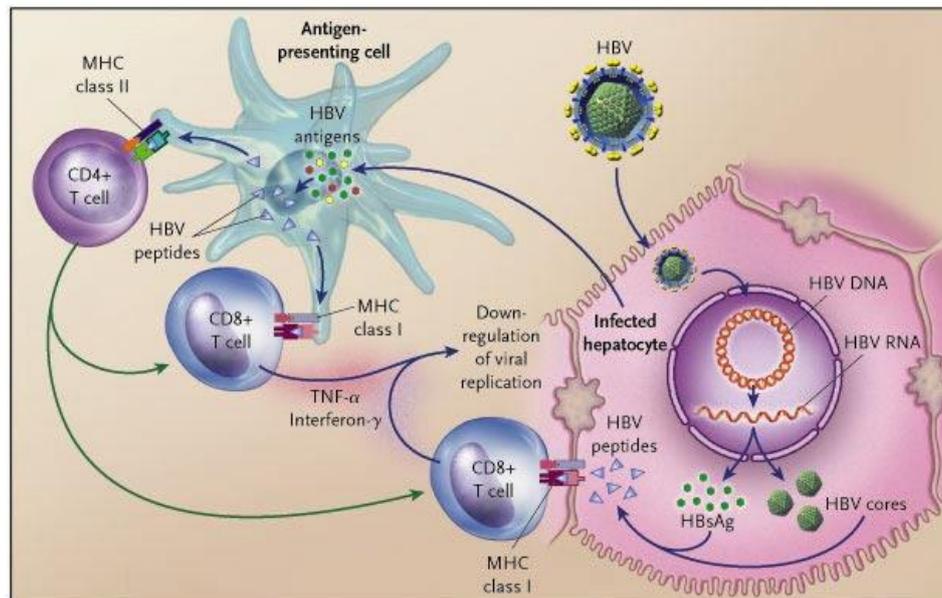


Figura 7. Risposta immunitaria cellulo-mediata ad HBV¹.

All'attivazione dei linfociti T citotossici consegue l'innescò della risposta infiammatoria aspecifica, sostenuta da TNF, radicali liberi e proteasi responsabili del danno cellulare.

Numerosi dati sono a favore del controllo della replicazione virale esercitata dai linfociti T citotossici in assenza di una completa eradicazione dell'infezione. Questo supporterebbe l'ipotesi secondo cui, dopo la risoluzione favorevole dell'evento acuto persisterebbero minime quantità di virus replicante evidenziabili solo tramite metodiche di amplificazione genica (PCR)³⁸.

A supporto dell'azione antivirale dei linfociti T citotossici si sviluppa la risposta cellulare mediata dai linfociti T helper, limitata ad antigeni proteici, in particolare *HBcAg* e *HBeAg*, presentati sulla superficie di cellule presentanti l'antigene (APC), quali i macrofagi, in associazione a molecole MHC-II³⁵.

All'attivazione dei linfociti CD4+ consegue il rilascio di citochine, responsabili della proliferazione cellulare T, dell'aumentata espressione epatocitaria di molecole MHC-I e dell'inibizione della replicazione virale.

I linfociti T helper sono, inoltre, responsabili della differenziazione delle plasmacellule in linfociti B attivati, e quindi della risposta anticorpale.

La produzione di anticorpi è un elemento critico per la neutralizzazione di particelle virali in circolo e per impedire l'ingresso dei virioni negli epatociti.

Tuttavia, si è concordi nel ritenere che l'evoluzione, favorevole o meno, dell'infezione acuta dipenda dall'efficienza della risposta cellulo-mediata specifica, in quanto la risposta umorale è vigorosa, sia nel corso dell'infezione acuta autolimitantesi che in quella ad evoluzione cronica.

1.1.5 Storia naturale dell'infezione

La storia dell'infezione naturale da HBV è correlata a diversi fattori legati alla biologia del virus, alla presenza di fattori concomitanti nell'ospite (coinfezioni con altri virus e abuso di alcol), e a caratteristiche dell'ospite stesso. Infatti assumono importanza per l'esito della malattia l'età di acquisizione dell'infezione, il sesso e lo stato di immunocompetenza del soggetto infettato ³⁹.

Nel 90% dei casi l'infezione ha un decorso del tutto asintomatico (infezione subclinica) e solo il 10 % dei soggetti infetti sviluppa l'infezione acuta caratterizzata da una manifestazione clinica variabile. Dei soggetti che mostrano sintomatologia, il 90% acquisisce l'immunità, grazie alla sierconversione ad *anti-HBs*, e va incontro di conseguenza a guarigione.

Il decorso dell'infezione acuta dipende da fattori quali, ad esempio, l'età a cui si acquisisce: la sintomatologia si manifesta più frequentemente in adulti e in bambini maggiori di 5 anni (33-50% dei casi) rispetto a bambini di età inferiore, che hanno una percentuale di sviluppo della malattia clinica del 5-15%.

Raramente la malattia acuta ha un decorso fulminante (1%) con un tasso di mortalità dell'80% ⁴⁰.

Nel caso di epatite acuta, la sintomatologia insorge con nausea, forti dolori addominali e muscolari, febbre leggera ed un aumento delle transaminasi con incrementi da 3 a 5 volte i valori normali fino a 100 volte il *range* di normalità (il limite massimo di ALT a 25°C è di 22 UI/L negli uomini e di 18 U/L nelle donne). Dopo qualche giorno a causa della presenza di bilirubina nel siero, può comparire ittero (25-35% dei casi), con presenza di urine scure, feci chiare e colorazione

giallastra di cute, mucose e congiuntive. La manifestazione dell'ittero è correlata con la presenza di un'elevata carica virale e con un breve periodo di incubazione ^{9, 32}.

Mediamente nel 10% dei casi, in seguito ad infezione acuta, si ha cronicizzazione, con persistente antigenemia. Si definisce portatore cronico di *HBsAg* un soggetto in cui l'antigenemia perduri per più di 6 mesi.

La probabilità di cronicizzazione dell'infezione da HBV è strettamente correlata con l'età di acquisizione: l'infezione acquisita in età neonatale o nei primi anni di vita mostra un'alta percentuale di cronicizzazione (90% nel primo caso e 50% nel secondo) per poi decrescere con l'aumentare dell'età fino a raggiungere una percentuale del 6-10% nei soggetti adulti ⁴¹.

Il meccanismo alla base della mancata o scarsa clearance virale in età neonatale sarebbe rappresentato dalla tolleranza immunitaria all'antigene *e*, dovuta all'incompleto sviluppo del sistema immunitario ⁴².

I meccanismi immunitari patogenetici, implicati nell'evoluzione cronica in età adulta nei pazienti immunocompetenti, non sono stati del tutto chiariti. Secondo diverse evidenze sperimentali, accumulate nel corso degli anni, i pazienti con epatite cronica B possono presentare diverse alterazioni quali: una scarsa risposta anticorpale diretta contro gli antigeni codificati dalla regione *pre-S*, che sarebbe responsabile di un fallimento della neutralizzazione virale, una ridotta o assente produzione di interferone alfa che favorirebbe la replica virale, una ridotta produzione di interleuchine e un difetto della risposta dei linfociti T suppressor verso gli antigeni del nucleocapside virale. Di tutte le alterazioni citate la più importante appare indubbiamente il difetto funzionale dei linfociti T citotossici che può essere provocato da diversi fattori: presenza di elevata e persistente viremia (le alte cariche virali determinano un "esaurimento funzionale" dei linfociti T), stato di immunotolleranza nei confronti del virus indotto dall'eccesso di *HbeAg*, espressione di alcuni particolari antigeni di istocompatibilità sulla superficie degli epatociti, deficit funzionale dei linfociti T helper (il cui compito è coadiuvare i linfociti T citotossici) ed infine infezione degli stessi linfociti T citotossici da parte del virus ³.

Merita di essere sottolineato che la selezione di mutanti virali è da considerarsi l'effetto dell'infezione persistente piuttosto che la causa.

A livello epatico, lo stato infiammatorio cronico risulta solitamente più moderato rispetto a quello correlato all'epatite acuta, anche se la persistenza per anni o decenni può favorire la progressione a cirrosi, con un maggior rischio di insorgenza del carcinoma epatocellulare (HCC).

Anche la severità del quadro clinico, è inversamente collegata all'età in cui si stabilisce la malattia cronica.

Anch'essa presenta quadri evolutivi differenti, come lo stato di "portatore sano" (80% dei casi) in cui il soggetto mostra un quadro epatico normale o caratterizzato da lesioni minime ed una funzionalità epatica inalterata. Lo stato di "portatore sano" è caratterizzato dall'integrazione del DNA virale negli epatociti, condizione che spesso risulta clinicamente stabile nel tempo.

Nel 20% dei casi invece si stabilisce una malattia cronica attiva o persistente, caratterizzata da una continua replica virale con possibile sviluppo di cirrosi e/o carcinoma epatocellulare.

Non è ancora del tutto chiaro come il virus possa indurre epatocarcinoma. Il virus potrebbe innescare direttamente lo sviluppo del tumore o quest'ultimo potrebbe insorgere indirettamente come conseguenza dell'infiammazione cronica e della cirrosi. Molti virus che inducono tumori portano oncogeni. Sembra, tuttavia, che il genoma dell'HBV non presenti oncogeni; inoltre il lungo periodo di latenza fra l'infezione da HBV e la comparsa di un tumore non è compatibile con l'ipotesi che un oncogene virale provochi l'epatocarcinoma. E' anche da considerare la carica virale di HBV-DNA presente nel paziente con epatite cronica, come ulteriore fattore di rischio di insorgenza del carcinoma epatocellulare ⁴³.

Non è da escludere tuttavia che la proteina attivatrice di geni prodotta dal gene X di HBV sia coinvolta in uno stadio precoce del processo di cancerogenesi.

Infatti a questa proteina è riconosciuta la funzione di transattivante e viene trovata frequentemente in soggetti con epatocarcinoma, anche se il suo ruolo esatto non è stato ancora chiarito.

Un punto chiave, anche se non l'unico, nell'induzione del carcinoma è rappresentato dall'integrazione del DNA del virus nel genoma dell'ospite. Molti retrovirus utilizzano l'inserzione di sequenze virali in posizione adiacente a oncogeni cellulari come strategia per indurre carcinomi e leucemie in mammiferi ed uccelli. Sebbene HBV non abbia di solito necessità di integrare il proprio DNA nel genoma dell'ospite per compiere il proprio ciclo vitale, questa integrazione può avvenire casualmente. E' un meccanismo *random*, che non avviene quindi in siti preferenziali del genoma cellulare e potrebbe attivare dei promotori cellulari ⁴⁴.

I fattori virali e quelli dell'ospite, oltre alla coinfezione con il virus dell'epatite C (HCV), virus dell'epatite D (HDV) o il virus dell'immunodeficienza umana (HIV), insieme alla coesistenza di danno al fegato dall'abuso alcolico ed al sovrappeso, possono influenzare il decorso naturale dell'infezione cronica da HBV oltre che l'efficacia delle strategie terapeutiche antivirali ⁴⁵.

1.2 EPIDEMIOLOGIA DI HBV

1.2.1 Prevalenza, distribuzione e modalità di trasmissione

Tra i principali virus epatotropi maggiori, l'HBV mantiene una rilevanza a livello mondiale. L'Organizzazione Mondiale della Sanità ha stimato che 2 miliardi di persone, circa un terzo della popolazione mondiale, ha evidenza sierologica di un'infezione passata o presente da HBV ⁴⁶ e oltre 350 milioni sono gli infetti cronici ⁴⁷; di questi il 78% vive in Asia, il 16% in Africa, il 3% in Sud America e il 3% in Europa, Nord America e Oceania ⁴⁸.

La frequenza di infezione di HBV e le modalità di trasmissione variano in modo marcato nelle differenti parti del mondo.

Le regioni ad alta endemia, dove la prevalenza di infezione cronica è superiore all'8% e il 70-98% della popolazione ha evidenza sierologica di passata o intercorrente infezione, includono la regione Asia-Pacifico, l'Africa Sub-Sahariana e

il bacino Amazzonico ⁴⁶. Circa il 45% della popolazione mondiale vive in paesi ad alta endemia per HBV ⁴⁴.

Le regioni ad endemia intermedia, dove la prevalenza di infezione cronica è compresa tra l'1 e l'8% e in cui il dato sierologico di avvenuto contatto si riscontra nel 10-60%, sono: Giappone, il subcontinente indiano, la parte meridionale dell'Oriente, Europa Centrale e Medio Oriente ⁵⁰.

Le restanti regioni si considerano a bassa endemia, e sono caratterizzate da una prevalenza di portatori *HBsAg* positivi inferiore all'1% e un tasso di esposizione del 0.5-2% ⁴⁶ (Figura 8).

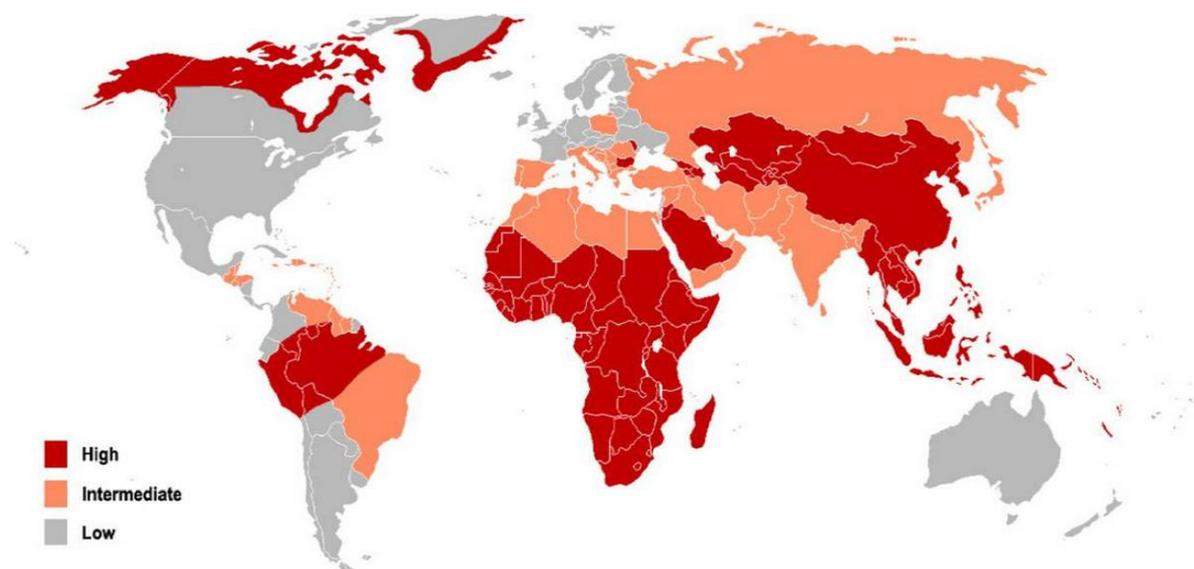


Figura 8. Distribuzione geografica dell'infezione cronica da HBV ⁴⁹.

La trasmissione verticale è un fattore determinante per lo sviluppo dell'infezione cronica di HBV ed è la via predominante dell'acquisizione dell'infezione nei paesi ad alta endemia, in particolare della regione Asia-Pacifico. Se la madre è *HBsAg* e *HBeAg* positiva, il rischio di trasmissione perinatale stimato è del 70-90% e, in questo caso, il 90% dei neonati cronicizzerà; nel caso in cui invece la madre sia *HBsAg* positiva ma *HBeAg* negativa, il rischio di trasmissione è ridotto al 10-40% e quello di cronicizzazione al 40-70% ⁵¹.

La trasmissione orizzontale del virus nella prima infanzia è determinante nei paesi endemici ed è la fonte principale dell'infezione nell'Africa Sub-Sahariana ⁵². Il tasso di cronicizzazione è del 50% all'età di 5 anni.

L'acquisizione della malattia in età adulta si contrae in paesi a bassa endemia. Si verifica a seguito di rapporti sessuali a rischio, per via parenterale (tossicodipendenti, tatuaggi, piercing, etc.) e in soggetti esposti a rischio professionale ⁵³.

Attualmente, i movimenti della popolazione e la migrazione stanno cambiando l'incidenza e la prevalenza della malattia ovunque, ma soprattutto in alcuni Paesi europei a bassa endemia.

Da epatite cronica progressiva, la malattia può evolvere in cirrosi, con un tasso annuale del 2-6% in pazienti *HBeAg* positivi e dell'8-10% in pazienti *HBeAg* negativi. L'incidenza annuale di HCC (carcinoma epatocellulare), è inferiore all'1% nei pazienti senza cirrosi e del 2-3% nei pazienti con cirrosi ⁵⁴.

L'HCC rappresenta il sesto tipo di cancro più comune nel mondo, con circa 630,000 nuovi casi ogni anno. Il 55% dei casi di tutto il mondo e l'89% dei casi che si verificano nelle regioni endemiche o iperendemiche è causato da infezione cronica da HBV. Degli oltre 350 milioni di infetti cronici da HBV circa un quarto svilupperà HCC ⁴⁶.

La sopravvivenza nella cirrosi compensata è dell'80-86% in 5 anni, mentre nella cirrosi scompensata è del 14-35% in 5 anni ⁵⁵.

L'insufficienza epatica scompensata HBV-correlata insieme all'HCC sono responsabili della morte di circa 1 milione di persone ogni anno ^{1, 56}, ed attualmente, rappresentano il 5-10% dei casi di trapianto del fegato ^{1, 52}.

La vaccinazione contro l'HBV, incorporata per lungo tempo in un programma di immunizzazione nei paesi ad alta prevalenza, quali l'Africa Sub-Sahariana e la regione Asia-Pacifico, ha consentito un'elevata riduzione del tasso di infezione cronica di HBV. La prevalenza dei portatori di *HBsAg* è scesa dall'8% a 1% ⁵³.

L'inclusione universale del vaccino nell'*Expanded Programme of Immunization* in tutti i paesi porterà, dopo un periodo opportuno, alla scomparsa dell'infezione da

HBV e successivamente, alla scomparsa delle complicanze quali cirrosi e carcinoma epatocellulare HBV-correlata ⁵⁷.

1.2.2 Numero riproduttivo di base (R_0)

Sulla base del *tasso di crescita* (r) è possibile stimare il *numero riproduttivo di base* (R_0) di un'infezione, che rappresenta il numero medio di casi secondari infettati da un singolo caso primario, in una popolazione senza difese immunitarie. L' R_0 viene calcolato in base alla durata dell'infezione nei portatori. Oltre alla durata dell'infettività, una serie di altri fattori possono influenzare l' R_0 , come la modalità e l'efficienza della trasmissione dell'infezione nella popolazione. Varie stime dell' R_0 di HBV sono state pubblicate. Medley ⁵⁸ ha suggerito che l' R_0 è compresa tra i valori di 1 e 2 nei paesi a bassa endemia. Il valore dell' R_0 cambia anche in base ai sottogenotipi. La stima dell' R_0 del sottogenotipo D3 è significativamente più alta rispetto al sottogenotipo A2 ⁵⁹. Ciò può essere dovuto alle diverse modalità di trasmissione dei sottogenotipi di HBV nella popolazione: il sottogenotipo A2 viene trasmesso principalmente attraverso la via sessuale, mentre il D3 con la via parenterale (trasfusioni di sangue e i suoi derivati, iniezioni non sicure, etc.).

1.3 ASPETTI TERAPEUTICI

Dal momento che oltre il 90% degli adulti superano spontaneamente l'infezione da HBV, l'epatite acuta non richiede particolari terapie, a parte specifici trattamenti per i sintomi quali nausea, vomito e anoressia. L'epatite cronica al contrario richiede l'utilizzo di farmaci antivirali. Sono stati utilizzati finora diversi farmaci: alcuni si sono mostrati inefficaci mentre altri presentano elevata tossicità alle dosi terapeutiche ⁶⁰. Le terapie utilizzate per l'epatite cronica si prefiggono di ridurre o di impedire la replicazione virale, di migliorare lo stato infiammatorio a livello epatico e di far regredire o stabilizzare il processo fibrotico (se presente) e prevenire lo sviluppo di cirrosi e di HCC. I farmaci utilizzati nel trattamento

terapeutico dell'epatite cronica B sono l'interferone alfa e gli analoghi nucleosidici (Figura 9), farmaci sintetizzati chimicamente, il cui ruolo, una volta fosforilati, è quello di competere con i normali nucleotidi durante l'incorporazione nelle sequenze di DNA virale e inibendo di conseguenza la replicazione virale.

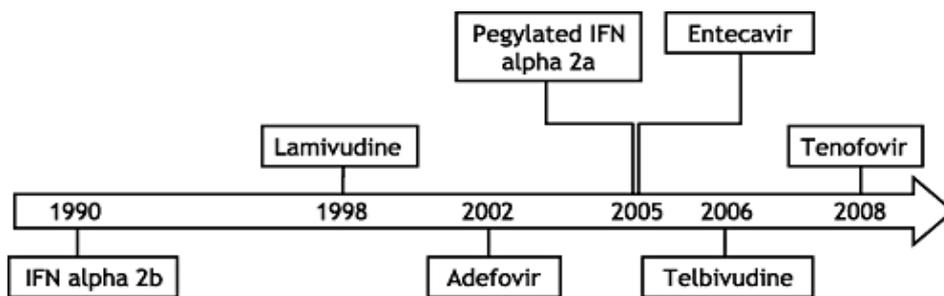


Figura 9. Terapie fino ad oggi approvate per il trattamento dell'epatite cronica B ⁶¹.

L'interferone alfa convenzionale (INF- α 2b), approvato nel 1991, è stato il primo agente utilizzato per il trattamento dell'epatite cronica B. Ha una doppia modalità di azione, comportando un'azione antiproliferativa ed antivirale, ma anche immunomodulante, stimolando l'attività dei linfociti T-*helper*, la maturazione dei linfociti B ed inibendo i T-*suppressor*. Viene utilizzato per pazienti con infiammazione attiva, innalzamento delle ALT per un periodo superiore a 6 mesi e livelli di viremia moderati ($<10^8$ IU/ml) ⁴⁹.

Si è potuto valutare che in soggetti con epatite cronica B, positivi per *HBeAg*, si ottiene una risposta virologica nel 20-50% dei pazienti soggetti, a terapia da almeno 6 mesi, con l'eliminazione dell'*HBeAg* e del HBV-DNA circolante.

Una limitazione di questo farmaco è legata alla modalità della sua somministrazione (per via parenterale) e agli effetti collaterali; inoltre il farmaco non può essere somministrato a pazienti con cirrosi scompensata e mostra un effetto scarso o nullo se la carica virale di HBV è elevata ⁶².

Nel 2005 una nuova formulazione di INF- α , l'interferone alfa-2a peghilato, è stata approvata nel trattamento dell'epatite B cronica. Da allora l'INF- α convenzionale è stato gradualmente sostituito dall'INF- α 2a peghilato ⁶¹. Nonostante l'estensione

del trattamento a 48 settimane, rispetto alle 24 di trattamento con l'INF- α convenzionale, è stato mostrato che il tasso di sieroconversione (33%) è praticamente identico a quello dell'interferone alfa convenzionale (32%). Inoltre, dopo tre mesi di *follow-up* per pazienti *HBeAg* negativi con un basso livello di HBV-DNA, il tasso di HBV-DNA non rilevabile mediante PCR, è solo del 18% ⁶³.

Il primo analogo nucleosidico ad essere stato approvato, nel 1998, nella terapia dell'epatite cronica B, è la lamivudina.

La lamivudina viene metabolizzata direttamente negli epatociti e mostra una maggiore tollerabilità rispetto all'IFN. Inoltre può essere somministrata per via orale, dando una maggiore sicurezza di impiego anche in soggetti con cirrosi scompensata. Porta anche ad un ripristino della risposta immunitaria (mediata dalle cellule T) del paziente, a causa della rapida diminuzione della viremia indotta dal farmaco.

L'efficacia del farmaco sembra aumentare in relazione alla durata del trattamento, passando dal 17% di sieroconversione ad *anti-HBe* ad 1 anno di terapia al 27%, 36% e 47% rispettivamente a 2, 3 e 4 anni dall'inizio della terapia con l'analogo nucleosidico.

Uno svantaggio è però caratterizzato dalla possibilità di comparsa di mutanti farmacoresistenti. Le mutazioni più comuni cadono nel dominio YMDD del gene della polimerasi. Si è notato come la percentuale di insorgenza di mutazioni aumenti con l'aumentare del tempo al quale un soggetto è sottoposto a trattamento con la lamivudina (fino al 76% dei pazienti dopo 5 anni di terapia) ⁶⁴.

A dispetto di questi problemi è stato osservato che il trattamento a lungo termine è in grado di ridurre la progressione della malattia in termini di sviluppo di cirrosi e di epatocarcinoma ^{64,65}. Questi benefici sono stati osservati sia in pazienti cirrotici che non cirrotici ⁶⁵.

Attualmente l'uso della lamivudina è limitato dall'alto tasso di insorgenza della resistenza.

L'adefovir dipovixil è il secondo analogo nucleosidico ad essere stato approvato dalla FDA (*Food and Drug Administration*) per il trattamento di pazienti cronici infetti con HBV. Sebbene questo farmaco sia meno efficace nella soppressione dell'HBV-DNA, se comparato alla lamivudina ⁶⁶, la probabilità rispetto a quest'ultima di comparsa di resistenza è minore (29% dopo 5 anni di trattamento con adevofir) ⁶⁷. Questo basso tasso di sviluppo di resistenza potrebbe essere associato alla sua minima struttura aciclica flessibile che ostacola l'insorgenza di resistenza, dovuta all'ingombro sterico.

L'adefovir è principalmente usato nei pazienti lamivudina-resistenti in quanto il suo impiego, come farmaco di prima linea per pazienti con epatite B cronica non precedentemente trattati, è limitato dalla modesta attività come soppressivo antivirale e dalla potenziale tossicità renale ⁶¹.

L'entecavir è stato approvato per il trattamento dell'epatite cronica B nel 2005. È un nucleoside, appartenente a un nuovo sottogruppo, i ciclopentani, ed è caratterizzato da un'attività soppressiva anti-HBV particolarmente efficace. Diversi studi, effettuati su pazienti trattati con entecavir, hanno evidenziato un significativo miglioramento nell'istologia epatica, dei livelli di HBV-DNA e di ALT, rispetto a quelli trattati con lamivudina. Il 67% dei pazienti *HBeAg*-positivi e il 90% dei pazienti *HBeAg*-negativi hanno raggiunto livelli sierici di HBV-DNA non rilevabili mediante PCR a un anno di trattamento ^{68,69}. Inoltre un recente studio ha mostrato che oltre il 91% dei pazienti presenta una riduzione della carica virale a valori non rilevabili (< 12 IU/ml) dopo tre anni di terapia ⁷⁰, e che questa elevata percentuale persiste dopo cinque anni di trattamento continuo con entecavir ⁷¹.

Un ulteriore vantaggio della terapia con entecavir è rappresentato dalla probabilità molto bassa di comparsa di resistenza (solo l'1,2% dopo 5 anni nei pazienti naïve al trattamento) ⁷². Questo basso tasso di resistenza è correlato sia alla profonda soppressione virale, che alla richiesta di almeno tre siti di mutazione genetica per conferire la farmacoresistenza. Per questi motivi l'entecavir è attualmente l'agente di prima linea nel trattamento dei pazienti naïve affetti da epatite B cronica.

Poiché due dei tre siti richiesti per l'insorgenza della resistenza all'entecavir si sovrappongono a quelli che conferiscono resistenza alla lamivudina, essendo entrambe i farmaci analoghi nucleosidici, l'entecavir non rappresenta, invece, il farmaco di prima scelta per il trattamento di pazienti lamivudina-resistenti ⁷³. La probabilità di comparsa di resistenza all'entecavir è del 51% nei pazienti con mutazioni pre-esistenti responsabili di lamivudina-resistenza, dopo cinque anni di trattamento con entecavir ⁷².

Inoltre, a dispetto dei numerosi miglioramenti ottenuti con entecavir, il tasso di sieroconversione *HBeAg*, resta relativamente basso (solo il 23%) ⁷⁴. La sieroconversione *HBsAg* si verifica nel 5,1% dei pazienti dopo 96 settimane di trattamento con entecavir⁷⁵, con un ulteriore 1,4% entro il quinto anno per i pazienti che continuano a riceverlo ⁷¹.

Nel 2006, è stata approvata per il trattamento di pazienti, di età superiore ai 16 anni, con epatite cronica B, la telmivudina. Questo antivirale, appartenente alla classe degli analoghi nucleosidici, interferisce con l'azione di un enzima virale, la DNA polimerasi, che è implicato nella formazione del DNA del virus, impedendo così al virus di moltiplicarsi e diffondersi. In diversi trial l'efficacia della telbivudina nell'epatite cronica è stata comparata con quella della lamivudina. L'efficacia è stata misurata basandosi sul tasso di risposta al trattamento dopo un anno, definito dal più basso livello di DNA virale in circolazione nel sangue, associato sia ad un ritorno al normale dei livelli di un enzima epatico (ALT) o alla scomparsa dal sangue di un marcatore del virus dell'epatite B. La telbivudina è risultata essere più potente della lamivudina nell'abbassamento della carica virale a valori non rilevabili, dopo un anno di trattamento, sia in pazienti *HBeAg*-positivi che in pazienti *HBeAg*-negativi: il tasso di HBV-DNA non rilevabile è del 60% contro il 40% per gli *HBeAg*-positivi e del 88% contro il 71% per gli *HBeAg*-negativi, rispettivamente ⁷⁶. Tassi simili di perdita e sieroconversione *HBeAg* e di normalizzazione dei livelli di ALT sono stati riscontrati nei due gruppi ⁷⁷.

Una sola mutazione, rtM204I, è stata osservata nei pazienti con resistenza alla telbivudina. Sebbene il tasso di farmacoresistenza risulti essere più basso se

comparato a quello della lamivudina, la comparsa di resistenza virale alla telbivudina è ancora elevata (25% nei pazienti *HBeAg*-positivi e 11% negli *HBeAg*-negativi dopo due anni di trattamento) ⁷⁸, suggerendo cautela nell'uso prolungato di questo farmaco.

Inoltre, anche se riportate in percentuali molto piccole di pazienti dopo uno o due anni di terapia, esiste un rischio di miopatia e neuropatia in seguito al trattamento con telbivudina ⁷⁹.

L'ultimo agente ad essere stato approvato, nel 2008, per il trattamento dell'epatite cronica B è il tenofovir disoproxil fumarato, un analogo nucleotidico appartenente alla stessa categoria dell'adefovir, gli aciclici fosfonati. Questo farmaco ha tre caratteristiche principali: come l'entecavir e la telbivudina causa una profonda soppressione del HBV-DNA, è efficace nel trattamento dell'HBV lamivudino-resistente (molto più dell'adefovir) e, di grande importanza, è caratterizzato da una probabilità di sviluppare resistenza molto bassa, simile a quella dell'entecavir (nessun caso è stato osservato dopo quattro anni di terapia) ⁸⁰. Alcuni trial clinici hanno mostrato che la maggior parte dei pazienti hanno raggiunto livelli non rilevabili di carica virale (<400 copie/mL), la normalizzazione dei livelli di ALT e la perdita di *HBeAg* ⁷⁷. Dopo 96 settimane di tenofovir, la perdita di *HBsAg* e la sieroconversione *HBsAg* si è verificata nel 7.0% e 5.6% dei pazienti *HBeAg*-negativi e nel 3.8% e 1.9% dei pazienti *HBeAg*-positivi. La tossicità renale è stata riportata in piccole proporzioni in pazienti infetti da HIV trattati con questo farmaco, ma nessuna tossicità renale è stata riportata in pazienti con epatite B cronica immunocompetenti ⁶¹.

Il tenofovir rappresenta, per questi motivi, l'agente ideale per pazienti resistenti alla lamivudina o alla telbivudina, oltre che per pazienti con epatite cronica, naïve al trattamento.

Attualmente, è in fase di studio un altro agente, appartenente allo stesso gruppo di adefovir e tenofovir, il besifovir. Trail di fase I e II hanno mostrato che questo farmaco ha una profonda attività soppressiva antivirale, sia nei pazienti naïve al

trattamento che in quelli lamuvudino-resistenti ⁸¹. Comparato con entecavir, il trattamento con besifovir a 48 settimane raggiunge un'egual potenza di soppressione del virus, in assenza di tossicità renale ⁸². Inoltre, essendo un fosfonato aciclico, è atteso un basso tasso di resistenza.

1.4 VARIABILITÀ GENETICA DI HBV

La variabilità genetica di HBV è osservata sia come evoluzione dei genotipi e sottotipi, in seguito alla divergenza del genoma virale, nella popolazione portatrice del virus, sia come emergenza di mutazioni in ogni soggetto infetto.

La replicazione genomica di HBV è un processo caratterizzato da un'elevata frequenza di mutazione che fa sì che questo virus evolva molto più rapidamente rispetto ad altri virus a DNA, in risposta a fattori ambientali. Tali fattori hanno contribuito alla creazione di un'ampia diversificazione di ceppi HBV. Genotipi, sottogenotipi e sottotipi *HBsAg*, rappresentano popolazioni virali geneticamente stabili che condividono una comune storia evolutiva ⁸³.

1.4.1 Sottotipi

Il raggruppamento di isolati HBV in sottotipi *HBsAg* risale a più di 30 anni fa e costituisce la prima dimostrazione della diversità di tale virus.

Secondo i risultati di diverse analisi sierologiche su *HBsAg*, i ceppi del virus dell'epatite B sono classificati in 10 sottotipi, che rappresentano la manifestazione fenotipica di mutazioni a livello genomico, e sono: *ayw1*, *ayw2*, *ayw3*, *ayw4*, *ayr*, *adw2*, *adw3*, *adw4q-*, *adrq+*, *adrq-* ^{19, 20, 29}.

Epitopi *HBsAg*, implicati nell'espressione dei sottotipi specifici, sono localizzati in una regione che include i due loop esterni della molecola (aa 110-180) e rendono i ceppi HBV antigenicamente differenti ⁸³.

La classificazione sierologica di HBV è stata usata per motivi epidemiologici, e in qualche caso, per rintracciare anche le eventuali infezioni nosocomiali da HBV, per

definire la correlazione tra malattia e sottotipo specifico responsabile. Tuttavia, nell'ultimo decennio, la determinazione sierologica è stata sostituita gradualmente dalla genotipizzazione ³². I genotipi hanno un ruolo nella determinazione della gravità clinica e nello sviluppo dell'infezione cronica ⁸⁴. Inoltre, lo stato di epatite B cronica *HBeAg* negativa è genotipo-specifica ⁸⁴.

1.4.2 Genotipi e sottogenotipi

I ceppi del virus dell'epatite B sono classificati in 8 genotipi, denominati con le lettere maiuscole A, B, C, D, E, F, G, H. I primi quattro genotipi (A-D) sono stati descritti da Okamoto nel 1988 ⁸⁵, mentre gli altri (E-H) sono stati caratterizzati nel corso degli ultimi due decenni ²⁹. Recentemente, è stato proposto il nono genotipo "I" evidenziato nella Cina Nord-Occidentale ⁸⁶, India dell'Est ⁸⁷, Laos ⁸⁸ e Vietnam ⁸⁹, nonostante sia ancora oggetto di dibattito ⁹⁰, poiché potrebbe essere un ceppo ricombinante HBV/ACG ⁸⁹. Recentemente, è stato presentato il decimo genotipo, nominato provvisoriamente "J", isolato in un paziente giapponese con HCC ⁹¹.

La definizione "genotipo" si basa su una divergenza superiore al 7.5% sulla sequenza dell'intero genoma o superiore al 4% sulla sequenza dell'antigene s ⁸⁴ per i ceppi appartenenti a genotipi diversi, e inferiore al 7.5% per i ceppi dello stesso genotipo ⁹².

La lunghezza del genoma completa di HBV può variare da 3182 a 3248 nucleotidi, a seconda del genotipo, a causa di delezioni e inserzioni (Tabella 1) ⁸.

Dato che la capacità codificante del genoma di HBV è ottimizzata con la sovrapposizione degli *open reading frames* ed ogni nucleotide partecipa alla codifica di almeno una proteina virale, queste differenze nella lunghezza incidono su una o più proteine. Per esempio, il gene P, che codifica la polimerasi virale (*Pol*), si sovrappone parzialmente al gene C, il quale codifica l'*HBcAg*. I ceppi del genotipo A hanno un'inserzione di due codoni, in prossimità della fine del gene C, e di conseguenza sia l'*HBcAg* che *Pol* hanno due aminoacidi in più. Il genotipo G, invece, è caratterizzato da un'inserzione di 36 nucleotidi, immediatamente dopo il codone iniziale del gene C, che va ad aumentare la lunghezza dell'*HBcAg* di 12 aminoacidi

ma non va ad interessare il gene *Pol*, e da una delezione di un codone nella regione *PreS1*, che riduce di un aminonacido sia *PreS1* che *Pol* ²⁹.

Tabella 1. *Caratteristiche fondamentali dei genomi e le differenze tra i genotipi di HBV.*

Genotipo	Lunghezza genotipo in bp	Differenze in ORF
A	3221	Inserzione dell'aa 153 e 154 in HBc
B	3215	
C	3215	
D	3182	Delezione dell'aa 1-11 in preS1
E	3212	Delezione dell'aa 11 in preS1
F	3215	
G	3248	Inserzione di 12 aa in HBc Delezione dell'aa 11 in preS1
H	3215	
J	3182	Delezione di 11 aa in preS1 ⁹¹

Per quanto concerne la distribuzione geografica (Figura 10), i genotipi A e il D sono presenti a livello mondiale, anche se il genotipo D è relativamente raro nell'Europa del Nord e in America, ma è molto frequente nel bacino del Mediterraneo e in Medio Oriente. I genotipi B e C sono stati trovati essenzialmente in Estremo Oriente ⁹².

Il genotipo E è limitato all'Africa Sub-Sahariana, anche se sono riportati in letteratura dei casi rari in Francia, Inghilterra ed Arabia Saudita, probabilmente a causa dell'immigrazione ⁹³. Tale genotipo sembra essersi diffuso nella popolazione umana recentemente poiché i discendenti dei neri americani importati tra i 200 e i 400 anni fa negli Stati Uniti d'America non presentano questo genotipo ma il sottogenotipo A2, come i discendenti degli immigrati del Nord Europa ⁴⁹.

Il genotipo F è stato distribuito principalmente in America Orientale e Meridionale e la sua origine risiede nella popolazione Amerindiana.

Il caso del genotipo G è meno chiaro: inizialmente è stato rilevato negli Stati Uniti, Francia e Germania e, recentemente, in Messico con un'alta prevalenza ^{29, 94, 95}. L'origine geografica di questo genotipo rimane ancora sconosciuta ⁹⁶.

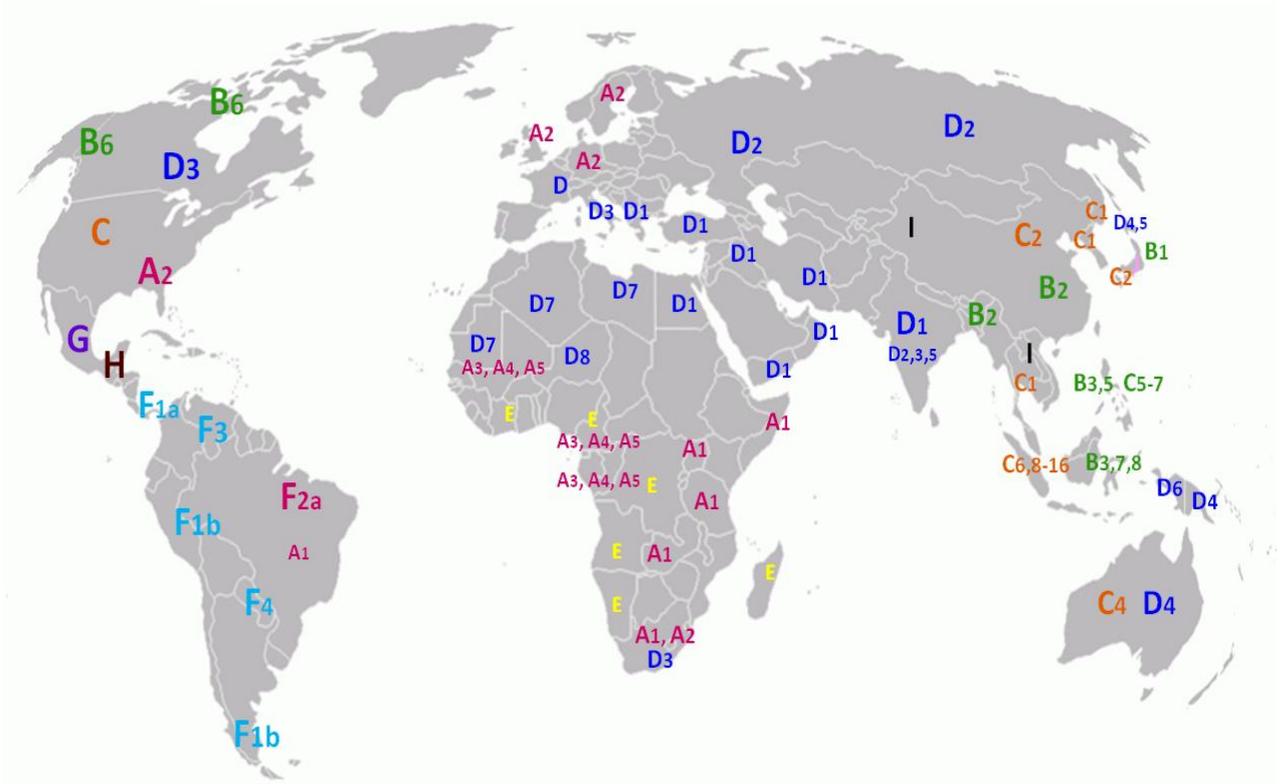


Figura 10. La distribuzione geografica di genotipi e sottogenotipi di HBV .

Il genotipo H è noto come il genotipo amerindio, in quanto isolato in gran parte in America Centrale e California, in quest'ultima probabilmente a seguito dell'immigrazione, e in Giappone ^{29,97}.

Quindi, i genotipi F e H possono essere definiti come genotipi del Pacifico/America Latina ²⁹.

I genotipi I e J sono stati descritti precedentemente.

Dall'analisi della distribuzione geografica si può notare come i genotipi A, D, E, F e G siano diffusi in aeree nelle quali la trasmissione orizzontale costituisce la principale via di infezione, a differenza dei genotipi B e C che sono prevalenti nelle aeree ad endemia maggiore, dove la modalità di trasmissione, verticale o perinatale, assume un ruolo essenziale nella diffusione del virus.

L'evoluzione genotipica di HBV ha portato allo sviluppo di varianti differenti chiamati "sottogenotipi". Un sottogenotipo diviene sospetto quando la divergenza intragenotipica basata sulla sequenza dell'intero genoma è superiore al 4% ⁴⁸.

I sottogenotipi conosciuti finora sono 48 e la loro classificazione secondo i diversi genotipi è la seguente: A (A1-A7), B (B1-B9), C (C1-C16), D (D1-D9), F (F1a/b-F4), I (I1-I2).

Solo durante il periodo 2008-2014 sono identificati quindici nuovi sottogenotipi nel Sud Est Asiatico, che sono il C7 nelle Filippine e il B7, B8, B9, C6, C8, C9, C10, C11, C12, C13, C14, C15 C16 e D6 in Indonesia ⁹⁸⁻¹⁰⁴.

Questi dati sono riportati in Figura 10 e in Tabella 2.

Tabella 2. *Sottogenotipi di HBV e origine geografica.*

	Sottogenotipo	Origine geografica	Referenza
A	A1	Africa, Asia, Brasile	105-107
	A2	Nord Europa, Nord America, Sud Africa	105, 108
	A3	Camerun, Gabon, Rwanda, Gambia	109, 110
	A4	Mali, Gambia	111, 112
	A5	Nigeria, Rwanda, Camerun, Haiti (pop. africana)	110, 112-114
	A6	Congo, Rwanda	114, 115
	A7	Camerun, Rwanda	110, 114
B	B1	Giappone	116
	B2	Asia (Giappone escluso)	8
	B3	Indonesia, Filippine	117
	B4	Vietnam, Cambogia	117
	B5	Filippine	118
	B6	Alaska, Canada Settentrionale, Groenlandia	119, 120
	B7-B9	Indonesia	99-101
C	C1	Sud-est asiatico (Vietnam, Birmania, Thailandia, Cina Meridionale)	121, 122
	C2	Estremo Oriente (Korea, Giappone, Cina Settentrionale)	8

	C3	Isole del Pacifico (Micronesia, Melanesia, Polinesia)	117, 119
	C4	Australia	123
	C5	Filippine, Vietnam	118
	C6	Indonesia, Filippine	98
	C7	Indonesia, Filippine	100, 124
	C8-C16	Indonesia	102-104
D	D1	Europa, Medio Oriente, Asia, Tunisia, Egitto	93, 125-128
	D2	Europa, Marocco, India	129-131
	D3	Sud Africa, Asia, Europa, USA, Canada Settentrionale	105,119, 120,129, 131
	D4	Australia, Giappone, Papua Nuova Guinea	29, 117
	D5	India Orientale, Giappone	131
	D6	Indonesia	98
	D7	Tunisia, Marocco	128
	D8	Nigeria	132
	D9	India Orientale	133
E		Africa Subsahariana, Regno Unito, Francia, Arabia Saudita	93
F	F1a	America Centrale e Sud America	134, 135
	F1b	Argentina, Giappone, Venezuela, USA	29
	F2	Sud America (Brasile, Venezuela, Nicaragua)	136
	F3	Venezuela, Panama, Colombia, Bolivia	29, 137
	F4	Bolivia, Francia, Argentina	29, 137
G		USA, Germania, Giappone, Francia, Messico	29, 94, 95, 138
H		Usa, Giappone, Nicaragua, Messico	29, 97, 138
I	I1	Laos, Vietnam, Cina Settentrionale	86, 88
	I2	Laos, Vietnam	88
J		Giappone	91

I genotipi e sottogenotipi hanno una distribuzione geografica specifica che riflette una evoluzione antica ²⁹. Tuttavia, dal momento che i pazienti possono essere infetti cronici, il quadro evolve a causa del fenomeno della migrazione umana ⁴⁸. Ad esempio, in alcuni studi sono stati trovati genotipi dominanti nell'area Africana (genotipo E e A) in regioni caratterizzate dalla dominanza del genotipo F/H. Essi derivano da genotipi di discendenti Africani, migrati in Sud America, durante la tratta degli schiavi, poche centinaia di anni fa. Ad esempio il sottogenotipo A1, prevalente in Brasile, è stato sicuramente importato con gli schiavi provenienti dall'Africa Orientale e si è inoltre diffuso attraverso la Somalia alla costa asiatica dell'Oceano Indiano ⁴⁹.

1.4.3 Mutanti

Il primo studio che si prefissò di stimare il tasso di mutazione di HBV fu riportato nel 1987 da Okamoto e collaboratori ³² ed era basato sulla comparazione di differenti cloni di HBV isolati dal siero di una donna di 54 anni infetta dalla nascita. Il tasso di mutazione fu stimato tra 1,4 e 3,2 x 10⁻⁵ sostituzioni per sito per anno, circa dieci volte più elevato rispetto a quello degli altri virus a DNA, ma più basso di quello dei retrovirus e di altri virus a RNA. Uno studio successivo, condotto nel 1989 da Orito, portò alla scoperta di un tasso di mutazioni silenti pari a 5x10⁻⁵ sostituzioni per sito per anno, ponendo la percentuale di mutazione di HBV tra i virus a DNA e i retrovirus ¹³⁹.

Il tasso di sostituzione in vivo dipende da un numero di fattori tra loro separati ma interdipendenti. Gioca senza dubbio un ruolo essenziale il peculiare meccanismo di replicazione virale, basato su uno stadio di trascrizione inversa da parte della DNA polimerasi virale, particolarmente suscettibile di errori per la mancanza dell'attività di proof-reading, attività esonucleasica di correzione di errori di incorporazione nucleotidica.

Altri fattori che, viceversa, limitano il numero di mutazioni possibili, spiegando la minor frequenza di mutazione rispetto a quella di un genoma retrovirale, sono la presenza di un genoma compatto con ORF sovrapposte, la necessità di conservare sequenze DR (direct repeats), promotori e altri elementi cis-attivanti implicati

nella replicazione del genoma e il coinvolgimento di una sola copia di RNA pregenomico, che a differenza dei retrovirus, limita le chance di ricombinazione omologa.

Considerando che il tasso di sostituzioni sinonime è considerevolmente maggiore di quelle non-sinonime (Hannoun stimò le sostituzioni non-silenti come l'88% di tutte le sostituzioni osservate in pazienti infettati verticalmente ¹³⁹) è stato suggerito che la variazione genomica di HBV è vincolata dai cambiamenti aminoacidici. Questo porta all'ipotesi che ogni proteina e ogni dominio di un ORF, avendo dei limiti strutturali e funzionali, tenda a conservare la sequenza aminoacidica nel tempo. Assumendo che la forza selettiva operi in modo indipendente su ogni ORF, le sostituzioni nucleotidiche nelle regioni sovrapposte sono il risultato di una doppia selezione su due ORF sovrapposte. L'insorgenza di una mutazione a livello di geni sovrapposti, ma non correlati tra loro, che porta alla modificazione di più di un prodotto genico, dovrebbe perciò produrre una variante virale con alterata antigenicità e/o replicazione e storia naturale che porta un vantaggio selettivo sul ceppo selvatico (*wild type*).

Nonostante la maggior parte delle mutazioni produce virus incapaci di sopravvivere o senza alcun effetto fenotipico, una mutazione può essere "fissata" e trasmessa alla progenie virale, grazie a una selezione positiva che favorisce la sopravvivenza del virus mutato con il genotipo che meglio riesce ad adattarsi alle condizioni in cui si trova.

Le forze selettive che operano sui mutanti sono rappresentate dall'immunità indotta, dalla vaccinazione e dal trattamento con farmaci antivirali. Tale selezione avviene sia a livello cellulare, a favore del virus che replica in modo più efficace, che a livello extracellulare, a seconda delle caratteristiche del sistema immunitario dell'ospite. La risposta dell'ospite appare come un fattore molto importante, essendoci evidenze che il tasso di mutazione nei decenni risulti irrilevante in caso di infezione silente, quando cioè la risposta immune dell'ospite è minima ¹³⁹.

Nel complesso, se si considerano l'elevata efficienza di replicazione del virus (con rilascio fino a 10^{11} particelle virali al giorno), la sua permanenza in soggetti cronici, la pressione immunitaria esercitata dall'ospite e dalle terapie antivirali, si può

comprendere come possano accumularsi singole mutazioni e far emergere ceppi di rilevanza clinica ed epidemiologica.

Il virus dell'epatite B offre un chiaro esempio di come un virus possa eludere il sistema immunitario e sopravvivere nell'ospite. La variazione di un determinante antigenico rende il virus mutato non più riconoscibile dagli anticorpi prodotti contro il ceppo selvatico permettendone la replicazione (*escape mutant*). Questo è valido sia per un riconoscimento anticorpale sia di tipo cellulo-mediato ²⁵.

Mutazioni virali sono state scoperte sia in soggetti affetti da epatite acuta o fulminante, sia in quelli con infezione di tipo cronico e possono riguardare diverse parti del genoma.

Sono state identificate delle mutazioni nella regione pre-core, che bloccano la secrezione dell'antigene. Infatti sono stati segnalati casi di pazienti che, nonostante mostrassero una sierconversione ad *anti-HBe*, mostravano una malattia progressiva caratterizzata da un'elevata concentrazione di HBV-DNA nel siero, mantenendo così la replicazione virale ed il danno epatico ⁹⁴. La mutazione puntiforme più recentemente identificata in questi soggetti consiste nella sostituzione di una G con una A, in posizione 1896 (G1896A), localizzata nella struttura ϵ del gene pre-core ¹⁴⁰. Questa sostituzione comporta l'introduzione di uno stop codon (TAG) al posto del codone che codifica per il triptofano, comportando la prematura terminazione della trascrizione e mancata sintesi dell'*HBeAg*. Tale mutazione non influisce sulla produzione dell'antigene *HBcAg*, in quanto la sua traduzione inizia a valle del codone di stop.

Un secondo gruppo di mutazioni, che si verificano in posizione 1762 e 1764, colpisce il promotore basale del core, determinando una riduzione del legame dei fattori di trascrizione fegato-specifici e la conseguente riduzione dei livelli di trascrizione dell'mRNA precore e core. La comparsa della doppia mutazione A1762T più G1764A comporta un decremento, ma non la scomparsa, della produzione di *HBeAg*, e un incremento della carica virale. Questo tipo di mutazioni si trovano più frequentemente negli individui infetti con il genotipo A ¹⁴⁰.

L'epatite portata da queste mutazioni virali (*HBe minus*), oltre a mostrare un decorso in genere più grave rispetto al *wild type*, presenta anche una minore risposta alla terapia con gli interferoni ²⁵.

Inoltre sono state osservate delle mutazioni a livello del gene core in soggetti immunocompromessi ed in pazienti orientali con epatite cronica o carcinoma epatocellulare. Tali varianti virali comportano la perdita degli epitopi riconosciuti dalle cellule B e T, conferendo un vantaggio selettivo al virus contro la risposta immunitaria dell'ospite. Non sembra che queste varianti siano associate ad un'induzione dell'infezione cronica ¹⁴¹.

Mutazioni del gene X possono coinvolgere elementi regolatori che controllano la replicazione, quali il promotore basale del core ed Enhancer II. Poiché il promotore basale del core comprende i nucleotidi 1742-1802 e si sovrappone con il gene X nel concomitante frame di lettura, le mutazioni A1762T e G1764A sopracitate causano cambiamenti anche a livello del gene X. Inoltre, delezioni o inserzioni a livello del promotore basale del core, shiftano il frame del gene X, portando alla produzione di proteine X tronche ¹⁴⁰.

Mentre i mutanti core sembrano capaci di evadere la risposta di tipo cellulomediata, i mutanti dell'envelope sfuggono al riconoscimento mediato dagli anticorpi neutralizzanti.

Il determinante "a" costituisce l'epitopo immunodominante della proteina di superficie *HBsAg* e deriva da un regione genomica molto conservata, è infatti localizzato tra i residui 124 e 147 dell'*HBsAg* ed ha una struttura a doppia ansa. È stato dimostrato che le cisteine in posizione 124 e 147 e la prolina in posizione 142 sono essenziali per la reattività antigenica ¹⁴² e di conseguenza sono incluse in sequenze molto conservate. Variazioni della struttura primaria del determinante "a" possono variare drasticamente la struttura antigenica conformazionale dell'*HBsAg*, tanto da permettere la selezione di *escape mutant* in grado di replicare anche in presenza di anticorpi neutralizzanti *anti-HBs* e di conseguenza di sfuggire alla risposta immunitaria ¹⁴³.

Il mutante G145R rappresenta un prototipo di queste varianti virali. Si tratta di una mutazione puntiforme, con la sostituzione di una guanina in una adenina in

posizione 587 del gene S, che determina anche una sostituzione aminoacidica glicina con arginina a livello del determinante "a" dell'*HBsAg*. Questa mutazione comporta una modificazione delle caratteristiche antigeniche del determinante "a", incidendo sulla idrofobicità della sequenza, tenuto conto che l'arginina è un residuo aminoacidico di dimensioni maggiori rispetto alla glicina.

L'idrofobicità della sequenza, che spesso è correlata con l'antigenicità, risulta notevolmente aumentata in pazienti infettati con varianti virali *escape*.

Anche le regioni pre-S1 e pre-S2, essendo altamente immunogeniche, sono potenzialmente sottoposte alla pressione del sistema immunitario, poiché esposte sulla superficie insieme all'*HBsAg*. La regione Pre-S rappresenta, infatti, quella con la maggiore eterogeneità di tutto il genoma¹⁴⁰.

Sono state riconosciute numerose mutazioni puntiformi, ampie delezioni nella regione pre-S1 e altre delezioni in grado di inibire l'espressione dell'intera proteina pre-S2. Alcune mutazioni comportano l'eliminazione dei siti di riconoscimento da parte delle cellule B e T, senza però alterare il sito di legame agli epatociti; in questo modo il virus sfugge alla risposta dell'ospite, pur mantenendo la capacità di infettare nuove cellule e di indurre uno stato di infezione cronica¹⁴⁴.

E' stato inoltre osservato come mutazioni nella regione preS/S inducano l'acquisizione di funzioni transattivanti, non presenti nel gene *wild type*, con la produzione di peptidi coinvolti nello sviluppo del carcinoma epatocellulare¹⁴⁵.

Per eradicare l'infezione cronica da HBV, è necessaria una terapia antivirale prolungata e per tale motivo si assiste alla selezione di ceppi resistenti con mutazione nel gene della polimerasi virale, analogamente a quanto accade durante la terapia antivirale in corso di infezione da HIV¹³⁹.

L'insorgenza di mutazioni a livello del gene della polimerasi di HBV è stata riscontrata in seguito a trattamento terapeutico antivirale con analoghi nucleosidici, quali la Lamivudina, il Famciclovir e l'Adefovir.

Trattamenti a lungo termine con la lamivudina possono far sorgere resistenze, a seguito di alterazioni nel dominio altamente conservato YMDD (tiroxina, metionina, aspartato, aspartato). Mutazioni insorgono quando la metionina (M) è mutata in una valina (V) o in una isoleucina (I). Le mutazioni più frequenti sono

rappresentate dalle sostituzioni metionina/valina (M552V) o metionina/isoleucina (M552I), che portano entrambe alla sostituzione amminoacidica ai codoni 195 e 196 nel gene S, per sovrapposizione delle 2 ORF ¹⁴⁶.

Spesso associata a queste mutazioni è L528M, la cui nuova nomenclatura la caratterizza come la rtL180M, mentre la M552V e la M552I sono indicate come la rtM204V e la rtM204I ¹⁴⁷.

Altre mutazioni frequenti sono rappresentate dalla rtM202I/V, spesso in associazione con la rtL178M. Risulta abbastanza frequente anche la V555I, la cui attuale denominazione risulta essere rtV207I.

E' importante sottolineare come l'incidenza delle mutazioni nel motivo YMDD aumenti progressivamente con la durata del trattamento con lamivudina.

Diversi studi mostrano il 24% di ceppi HBV mutati ad 1 anno di terapia, il 47-56% a 2 anni fino a raggiungere una percentuale del 69-75% a 3 anni ¹⁴⁸.

L'insorgenza di mutanti YMDD si manifesta sia in individui in corso di terapia immunosoppressiva, in seguito a trapianto di fegato, sia in individui perfettamente immunocompetenti ¹⁴⁹.

L'insorgenza di mutazioni durante il trattamento con la Lamivudina è solitamente associato con un innalzamento delle ALT nel siero, valore che si abbassa con il tempo, suggerendo che il farmaco continua ad esercitare il suo effetto soppressivo ¹⁵⁰.

Il Famciclovir è, a differenza, il pro-farmaco del penciclovir-trifosfato, un analogo nucleosidico che inibisce la DNA polimerasi, sia DNA che RNA-dipendente. Questo farmaco è utilizzato per il trattamento di pazienti con epatite cronica di tipo B ed in soggetti che hanno subito trapianto di fegato ortotopico. L'insorgenza di virus mutanti famciclovir-resistenti, è stata rilevata sia in pazienti immunodepressi quanto in soggetti immunocompetenti, sebbene mostri una frequenza più bassa rispetto a quella dei mutanti lamivudina-resistenti ¹⁴⁹.

La resistenza di HBV al farmaco Adefovir, un analogo nucleosidico inibitore della trascrittasi inversa, si verifica meno frequentemente (circa il 2% dopo 2 anni) rispetto alla resistenza alla Lamivudina ed è conferita dalla sostituzione di una

Treonina con una Asparagina nel codone 236 (rtN236T), localizzato a livello del dominio D, e dalla mutazione rtA181T a livello del dominio B.

1.4.4 Genotipi ricombinanti

La presenza di lunghe delezioni ed inserzioni nel genoma virale suggerisce che altri meccanismi, oltre a quelli sopracitati, possano avere un ruolo nell'insorgenza di varianti virali, come ad esempio la ricombinazione tra diversi ceppi di HBV coinfectanti uno stesso individuo.

La coinfezione di un paziente con due ceppi di HBV, che appartengono a due genotipi differenti, può causare lo scambio di materiale genetico tra di loro, dando luogo alla nascita di un ceppo ricombinante. Tuttavia, questo meccanismo rimane enigmatico. Attualmente, non sono ancora state illustrate le modalità con cui si verifica lo scambio di materiale genetico tra due genomi hepadnavirali a livello trascrizionale ⁸.

Numerosi autori hanno descritto cambiamenti nel genoma di HBV che sembrano essere conseguenze di un evento ricombinante (Figura 11) ⁸.

Due recenti studi hanno cercato di analizzare la prevalenza globale dei ceppi ricombinanti di HBV ^{151, 152}. Circa l'87% dei presunti ricombinanti erano B/C e A/D; altri ricombinanti comprendevano ibridi A/B/C, A/C, A/E, A/G, C/D, C/F, C/G, C/U (U per genotipo sconosciuto) e B/C/U. I risultati hanno indicato che i genotipi A e C mostrano una tendenza maggiore per la ricombinazione rispetto ad altri genotipi ¹⁵¹ e che le regioni principalmente coinvolte nella ricombinazione intergenotipica sono la regione fiancheggiante DR1 (nt 1640-1900) la regione preS1/S2 (nt 3150-100), la porzione terminale 3' del gene C (nt 2330-2450) e la porzione terminale 3' del gene S (nt 650-830) ^{151, 152}.

Da questi studi è inoltre emerso che, in generale, i ceppi ricombinanti hanno la tendenza ad avere la stessa distribuzione geografica dei genotipi parentali.

Ceppi ricombinanti B/C sono stati trovati in Asia Orientale e del Sud ¹¹⁶ ceppi A/D in Italia ¹⁵¹, Sud Africa ^{151, 153} e India ¹⁵¹, A/C in Cina ¹⁵¹, C/D in Tibet ¹⁵⁴, India ¹³³ e Cina Nord-Occidentale ¹⁵¹, A/E in Camerun ^{109, 151}, A/B/C e D/E/A in Africa ^{112, 151}, C/G in Thailanda ^{151, 155}, C/F in Bolivia ¹⁵¹ A/G in America ¹⁵¹. Recentemente, un

nuovo ceppo ricombinante D/E è stato identificato nella Repubblica Centrafricana, in Irlanda (in un paziente proveniente dall’Africa Occidentale), in Tunisia e in Guinea, in pazienti dove tutte le sequenze avevano lo stesso profilo ricombinante 128, 156-158.

Sono stati anche descritti gli eventi ricombinanti tra le sequenze di HBV umano e quello di scimpanzé ¹⁶⁰ o gibbono ¹⁵².

Alcuni dei ceppi ricombinanti intergenotipici hanno dato luogo al sottogenotipo prevalente dominante in zone geografiche specifiche. La ricombinazione tra genotipi B e C ha portato allo sviluppo di due ceppi differenti con distinte distribuzioni geografiche ¹¹². I ceppi di genotipo B non ricombinanti sono stati trovati in Giappone (sottogenotipo B1), mentre ceppi derivati dalla ricombinazione tra i genotipi B e C sono diffusi nel resto dell’Asia (sottogenotipo B2) ¹¹⁶.

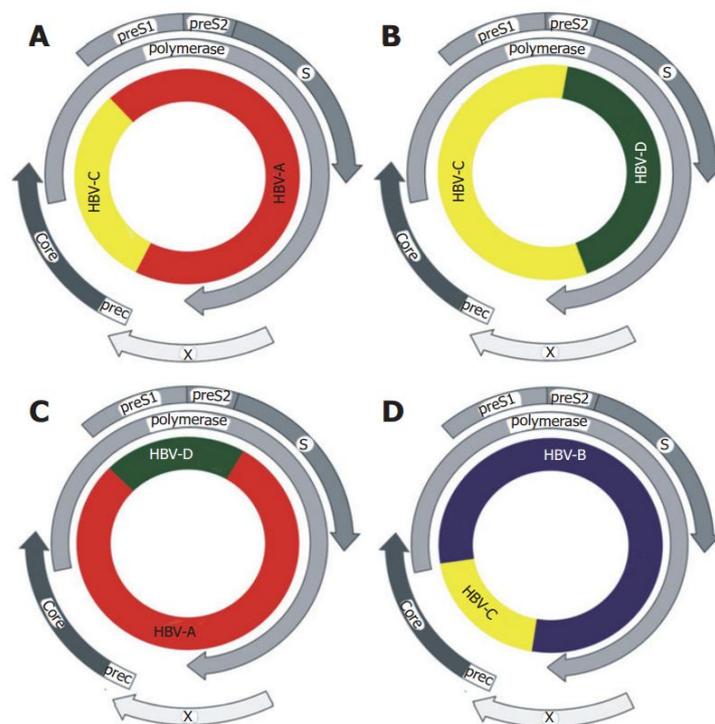


Figura 11. Presentazione schematica dell’organizzazione del genoma in ceppi ricombinanti tra vari genotipi di HBV⁸. Gli isolati provengono da **A:** Vietnam ¹⁵⁹; **B:** Tibet ¹⁵⁴; **C:** Africa ¹⁵³ e **D:** Asia ¹¹⁶. Le ORF sono rappresentate da frecce; il cerchio interno rappresenta il genoma di HBV.

Il ceppo ricombinante di HBV tra genotipi C e D costituisce il sottogenotipo dominante di HBV in Tibet ¹⁵⁴, anche se il genotipo D, che ha partecipato alla ricombinazione, si trova raramente in quell'area geografica. Il ceppo ricombinante HBV/DE, descritto nel 2010 in Niger ¹³², è stato classificato come sottogenotipo D8. Si suppone che la sua nascita è favorita dall'alta endemia di HBV in Niger (la prevalenza dell'*HBsAg* è del 17.6-29.8% ^{161, 162} e, secondo un studio molto più recente, del 19.2% ¹⁶³) e dalla coesistenza di genotipi differenti, D e E, nei paesi circostanti. In particolare, il genotipo D è prevalente in Algeria, Marocco e Tunisia ¹⁶⁴, che si trovano a nord del Niger, mentre il genotipo E è prevalente in Mali, Burkina Faso, Benin e Nigeria ^{165, 166}, che invece si trovano a ovest.

La ricombinazione intergenotipica può dirigere a caratteristiche biologiche differenti del virus dell'epatite B e ad un decorso clinico diverso nei pazienti ¹⁵¹. Ad esempio, il ceppo ricombinante B/C può influenzare la gravità clinica della malattia ¹¹⁶.

Sono stati inoltre descritti casi di ricombinazione intragenotipica. In uno studio di Simmonds & Midgley sono state descritte ricombinazioni all'interno dei genotipi A, D e F/H ¹⁵². Queste contribuiscono alla variabilità definita in ogni genotipo e, nello stesso tempo, possono ostacolare la comprensione della storia evolutiva di HBV ²⁹.

1.5 IMPLICAZIONI CLINICHE DEI GENOTIPI DI HBV

Risultati provenienti da diversi studi indicano che il genotipo di HBV può influenzare l'esito, a breve e a lungo-termine, dell'infezione da HBV ¹⁶⁷.

Relativamente agli studi di comparazione clinica dei genotipi virali noti bisogna premettere che, gran parte di essi, hanno il limite di essere focalizzati sulla comparazione del genotipo B con il C e del genotipo A con il D, a causa della relativa distribuzione geografica.

Alcuni studi condotti in Estremo Oriente hanno suggerito una più severa evoluzione clinica e una maggior propensione a sviluppare una fibrosi epatica

avanzata, cirrosi e HCC da parte di pazienti infettati da genotipo C rispetto a quelli con genotipo B ¹⁶⁷.

Tra questi un recente studio, condotto a Taiwan, mostrava che il genotipo C era associato a un maggior rischio di sviluppare HCC rispetto al genotipo B ¹⁶⁸.

In diversi studi ¹⁶⁹⁻¹⁷¹ è stato riportato che, mentre il genotipo C è maggiormente associato allo sviluppo di cirrosi ed epatocarcinoma in soggetti aventi più di 50 anni, il genotipo B sembra essere associato a uno sviluppo precoce del tumore, in soggetti più giovani e non cirrotici.

Il meccanismo patogenetico che sta alla base della più severa prognosi dell'epatopatia cronica da genotipo C non è ancora ben conosciuto.

Secondo alcuni studi, il genotipo C, rispetto al genotipo B, è caratterizzato da una maggiore latenza temporale necessaria per la sieroconversione ad *anti-e*. Ciò, come spiegato precedentemente, comporta un'infezione cronica persistente con multipli episodi di riacerbazione e remissione, che determinano una maggiore attività della patologia epatica ^{172, 173}. In particolare, risultati ottenuti da studi, condotti sempre su pazienti Taiwanesi, indicavano che infezioni da genotipo C erano associate a un minor tasso sieroconversione *HBeAg* rispetto al genotipo B ¹⁷⁴ e che il tasso di sieropositività per *HbeAg*, in bambini cronicamente infetti da genotipo C, dopo vent'anni era significativamente maggiore rispetto a quelli infetti da genotipo B.

Per quanto riguarda i genotipi A e D, diffusi in Europa e in India, i dati su larga scala sono limitati.

Secondo uno studio condotto in Spagna, pazienti affetti da epatopatia cronica di genotipo A hanno maggiore probabilità di eliminare HBsAg e HBV-DNA, una migliore evoluzione clinica ed una minore mortalità da scompenso epatico rispetto a quelli infettati da genotipo D e F ¹⁷⁵.

Questa osservazione è analoga a quella riportata in India, paese nel quale il genotipo D è risultato essere associato più frequentemente ad evoluzione cronica severa e alla degenerazione neoplastica in età giovanile rispetto al genotipo A ¹⁷⁶.

Per quanto riguarda la tendenza alla cronicizzazione delle infezioni acute da HBV, secondo recenti studi il genotipo A sarebbe correlato a un maggior rischio di progressione a infezione cronica rispetto al genotipo B e C. Inoltre in un numero

limitato di studi è stata riportata una maggiore cronicità dell'infezione acuta da genotipo A e D rispetto a quella da genotipo B e C ¹⁶⁷.

Alcuni genotipi sono stati associati ad andamento fulminante dell'epatite. Uno studio condotto in Vietnam ha evidenziato una maggiore frequenza del genotipo B; in particolare all'interno di tale genotipo, i ceppi associati ad epatite fulminante differirebbero da quelli non associativi per una mutazione nel gene X ¹⁷⁷. Dall'analisi dei casi con epatite B acuta negli Stati Uniti è stata trovata un'alta prevalenza di genotipo D ¹⁷⁸. Inoltre diverse epidemie di epatite fulminante sono state causate da coinfezione tra genotipi differenti di HBV con il virus dell'epatite D (HDV); in America ³² e in Africa ¹⁷⁹ è stata dimostrata la coinfezione di HDV con il genotipo F, mentre in Russia con il genotipo D ¹⁸⁰.

Per quanto concerne la risposta dei diversi genotipi di HBV alla terapia, pazienti con genotipo C e D sono più resistenti al trattamento con IFN α rispetto a quelli con genotipo A e B ¹⁸¹; probabilmente, la minore possibilità di risposta terapeutica è correlata al più rapido sviluppo di mutanti pre-core caratteristici dei genotipi C e D.

Uno studio multicentrico effettuato su pazienti *HBeAg*-positivi trattati con PEG-IFN α ha mostrato che il tasso di clearance di *HBeAg* a tre anni dal trattamento era maggiore nei pazienti con genotipo A e B rispetto a quelli con genotipo C e D. Tra pazienti *HBeAg* negativi trattati con PEG-IFN α , studi di follow-up a lungo termine hanno mostrato che la clearance il tasso di clearance di *HBsAg* era significativamente maggiore nei pazienti con genotipo A (20%) rispetto a quelli con genotipo non-A (6-9%)¹⁸².

Contrariamente ai genotipi A-D, pazienti infetti con il genotipo E-J sono rari e la loro risposta alla terapia con IFN α resta ancora sconosciuta ¹⁸². Secondo alcuni studi preliminari i genotipi E, F e H sarebbero più sensibili a questo tipo di trattamento rispetto al G ¹⁸³.

L'influenza esercitata dal genotipo sulla risposta agli analoghi nucleosidici non è ancora conosciuta. Chien e colleghi furono i primi a sostenere una risposta alla lamivudina molto maggiore nei pazienti infetti da genotipo B rispetto a quelli con genotipo C ¹⁸⁴. Comunque, due studi effettuati ad Hong Kong mostravano risultati

contraddittori^{185, 186}. È stato mostrato che lo sviluppo di resistenza alla lamivudina è simile comparando il genotipo B con il C¹⁸⁷, e il genotipo A con il D¹⁸⁸. In più nessuna differenza statistica è stata trovata nella risposta all'adefovir dipivoxil¹⁸⁹, all'entecavir¹⁹⁰ e alla telbivudina¹⁹¹ tra pazienti con differenti genotipi.

Nonostante ciò, uno studio molto recente ha indicato una significativa associazione del genotipo B con lo sviluppo di resistenza alla lamivudina, nei primi 12 mesi, rispetto al genotipo C¹⁹², suggerendo la necessità di ulteriori studi.

1.6 COINFEZIONE HBV – HIV

Poiché i virus dell'epatite B (HBV) e della sindrome dell'immunodeficienza acquisita (HIV) condividono le stesse modalità di trasmissione, non stupisce il dato secondo cui più dell'80% dei pazienti sieropositivi per HIV presenti dei markers sierologici di attuale o pregressa infezione da HBV¹⁹³. Si stima inoltre che l'8-11% della popolazione sieropositiva per HIV sia affetta da epatopatia cronica da HBV.

L'impatto della coinfezione da HBV sulla storia naturale dell'infezione da HIV è ancora controverso.

HBV non è un virus puramente epatotropo ed è noto che possa infettare attivamente i linfociti¹⁹⁴. È stata segnalata un'interazione diretta con HIV, legata all'interazione della proteina X di HBV, con sequenze regolatrici di HIV, che provocherebbe l'attivazione del provirus¹⁹⁵.

Tuttavia, solo due studi hanno suggerito l'ipotesi di un effetto sfavorevole della coinfezione con HBV sul decorso dell'infezione da HIV: il primo ha rilevato una più rapida progressione dell'infezione da HIV in pazienti con marcatori di infezione presente o passata da HBV; il secondo ha osservato una più elevata mortalità in pazienti affetti da AIDS conclamato in presenza di infezione cronica da HBV¹⁹⁶.

Altri studi non confermano tale influenza peggiorativa^{193, 197}.

Al contrario vi sono molte evidenze che la storia naturale dell'infezione da virus dell'epatite B sia modificata dalla coinfezione con HIV¹⁹⁸.

Nei pazienti coinfecti è più frequente l'evoluzione cronica dell'infezione da HBV, si repertano livelli più elevati di replicazione del virus dell'epatite B, il tasso di sieroconversione da *HBeAg* ad *anti-HBe* risulta inferiore ed è più frequente l'incidenza di riattivazioni nonostante la sieroconversione ad *anti-HBs* ¹⁹⁹.

Tale tipo di evoluzione è compatibile con la perdita del controllo immunitario della replicazione virale, in seguito alla deplezione linfocitaria dovuta all'infezione da HIV ¹⁹⁸. Nonostante la persistente ed intensa replicazione virale, normalmente, nei pazienti coinfecti, i livelli degli indici di citolisi epatica sono inferiori rispetto ai pazienti con sola infezione da HBV: tale dato si spiega con il modello immunopatogenetico del danno epatico correlato all'infezione da HBV. Secondo questo modello i linfociti T citotossici, caratteristicamente depleti in corso di infezione da HIV, svolgerebbero un ruolo chiave nella clearance virale, innescando la risposta infiammatoria responsabile della malattia epatica.

Nonostante ci siano dati a favore di una diminuita progressione della malattia epatica nei soggetti coinfecti, vi sono studi che indicano, al contrario, un significativo aumento della mortalità per malattia epatica in questa popolazione: 14.2/1000 vs 0.8/1000 rispetto ai pazienti infettati con il solo HBV; a ciò si associa spesso una aumentata progressione in cirrosi ²⁰⁰.

Tale trend verso l'evoluzione sfavorevole della malattia epatica è stata segnalato a partire dal 1996, quando è stata introdotta la HAART.

La più lunga aspettativa di vita associata all'assunzione della terapia antiretrovirale può permettere all'infezione da HBV di evolvere nella sua storia naturale a cirrosi, l'instaurarsi della quale può essere indotto, d'altro canto, dalle più frequenti riattivazioni della malattia epatica conseguenti all'immunoricostruzione causata dalla HAART.

Il danno epatico può essere inoltre aggravato dalla stessa tossicità intrinseca della maggior parte dei farmaci antiretrovirali ²⁰¹.

1.7 INFEZIONE DA HBV IN ITALIA

L'epidemiologia dell'epatite B in Italia è ampiamente variata negli ultimi decenni. Indagini sieropidemiologiche, effettuate sulla popolazione generale, negli anni '80 e '90, hanno mostrato un cambiamento nel livello di endemicità, da intermedia a bassa, con un'attuale prevalenza di portatori cronici intorno all'1,5%, un'incidenza di infezioni acute sintomatiche <1/100,000 abitanti/anno e un quasi completo azzeramento delle infezioni in età infantile ²⁰². A causa del progressivo calo delle nuove infezioni il numero dei portatori cronici di HBV sembra essere in costante decremento ²⁰³.

La progressiva riduzione dell'esposizione all'infezione da HBV, osservata già a partire dalla fine degli anni '80, è stata una conseguenza delle migliorate condizioni socio-economiche, della riduzione della circolazione intra-familiare, della maggiore conoscenza sulle vie di trasmissione e di una migliore prevenzione, oltre che alla campagna educativa sull'infezione da HIV, le cui modalità di trasmissione sono comuni ad HBV ^{203, 204}.

La strategia preventiva di immunizzazione ha dato il maggiore contributo alla riduzione dell'infezione da HBV, specialmente per i soggetti con un *range* di età di 15-24 anni ^{205, 206}. In individui appartenenti a questo gruppo d'età il tasso di infezione per 100,000 abitanti è diminuito da 17, nel 1990, a meno di 0.5 nel 2012. In Italia la vaccinazione anti-epatite-B è diventata obbligatoria per tutti i nuovi nati e gli adolescenti nel dodicesimo anno di età, con la legge del 27 maggio 1991, anche se nel 1983 era già stato avviato un programma di vaccinazione allo scopo di proteggere i gruppi di individui a rischio di infezione da HBV. Risulta inoltre obbligatorio, secondo la stessa legge del 1991, lo *screening* per *HBsAg* di tutte le donne gravide nell'ultimo trimestre di gestazione allo scopo di sottoporre i neonati di madri infette ad un protocollo specifico di profilassi passiva/attiva tramite somministrazione di immunoglobuline e vaccino ²⁰⁷.

Tale strategia vaccinale è stata introdotta allo scopo di proteggere i neonati da possibili infezioni da HBV, dato che, vista l'età di acquisizione, hanno un maggior rischio di cronicizzazione. Inoltre la somministrazione del vaccino in fase

preadolescenziiale permette di prevenire la potenziale esposizione ad HBV con l'avvicinarsi dell'età in cui si possono verificare comportamenti a rischio, sia di natura sessuale che comportamentale (assunzione di droghe) o professionale ²⁰⁸.

Il tasso di copertura, che ha raggiunto una percentuale del 94% tra gli adolescenti secondo le ultime stime, presenta delle differenze di tipo geografico: in un'indagine condotta su oltre 113.000 ragazzi si stima una percentuale del 98% nelle regioni del Nord, del 97% al centro e del 70% al Sud, dove il tasso risulta più basso a causa del precoce abbandono della scuola dell'obbligo, soprattutto per i giovani appartenenti alle famiglie meno abbienti ²⁰⁹.

Secondo il sistema di sorveglianza per l'epatite acuta SEIEVA (Sistema Epidemiologico Integrato dell'Epatite Virale Acuta) l'incidenza dell'infezione acuta da HBV in Italia è calata da 10/100,000, nel 1987, a 0.9/100,000, nel 2012.

La vaccinazione contro HBV ha permesso inoltre di ridurre anche l'incidenza dell'infezione da virus dell'epatite Delta, incapace di replicarsi in assenza di HBV. È stato mostrato che l'infezione acuta da HDV è decrementata durante il periodo 1987-2004 da 3.1 a 0.5 casi per 1 milione di abitanti ²¹⁰.

Nonostante si siano raggiunti dei risultati soddisfacenti, è stato stimato che un numero di portatori cronici di *HBsAg* compreso tra 500,000 e 1 milione, molti dei quali hanno acquisito l'infezione in passato, siano ancora presenti in Italia, suggerendo che HBV resta un problema rilevante di Sanità pubblica nel nostro Paese ²¹¹. L'analisi dei fattori di rischio nei nuovi individui infetti, mette in evidenza come l'uso di droghe intravenose, altre vie di esposizione parenterale (pratiche che prevedono il contatto con aghi e siringhe non sterili come tatuaggi, piercing, manicure, pedicure), contatti familiari con portatori cronici di *HBsAg*, così come l'esposizione a partner multipli abbiano un ruolo principale nella trasmissione del virus ²⁰⁷.

Un altro aspetto recente del problema è dato dall'intensa immigrazione da aree di intermedia ed elevata endemia per HBV come l'Est Europeo, l'Africa Sub-sahariana e l'Asia. Si stimano in circa tre milioni gli immigrati in Italia, con presenza di un buon numero di clandestini che sfuggono a qualsiasi misura di profilassi. Tuttavia nonostante il flusso migratorio da paesi ad elevata endemia abbia determinato

l'arrivo di nuovi portatori cronici di HBV, l'Italia resta ad oggi un paese a bassa endemia per questo virus.

Inoltre in Italia parte della popolazione adulta non risulta, tutt'oggi, coperta dalla vaccinazione: il 40-50% degli individui che si sottopone a vaccinazione volontaria (soprattutto categorie a rischio) non conclude il ciclo di vaccinazione; a ciò si aggiunge il fatto che circa il 5% dei soggetti sani, che ricevono le tre dosi di vaccino previste, non sviluppano una immunità protettiva nei confronti dell'HBV (*"non responders"*)²¹². Questi fatti unitamente ad un progressivo calo dell'attenzione sul problema anche da parte dei sanitari, come denunciato in un recente studio di Raimondo e colleghi²¹¹, potrebbe determinare il ritorno in una incrementata circolazione di HBV, in presenza di popolazione recettiva all'infezione, e potrebbe tradursi in un potenziale incremento di casi di epatite acuta nella popolazione italiana. Risulta pertanto necessario un continuo impegno per ottenere una copertura vaccinale più completa, che copra anche gruppi a rischio non facilmente raggiungibili.

1.8 FILODINAMICA E FILOGEOGRAFIA

Il concetto di " filodinamica " è stato introdotto per la prima volta da Grenfell nel 2004²¹³, e descrive come la notevole varietà dei processi evolutivi filogenetici dei patogeni sia influenzata dalla concomitanza di fattori quali la risposta immunitaria dell'ospite, che esercita una pressione selettiva sul patogeno, fenomeni di deriva genica "a collo di bottiglia" e dinamiche epidemiche di popolazione inter-ospite.

L'analisi filodinamica comunemente utilizza un modello temporale, definito orologio molecolare (*Molecular Clock*)²¹⁴, che si basa sull'assunto che le mutazioni si accumulino nel tempo in modo pressoché costante. Tale modello consente di delineare le relazioni intercorrenti tra le distanze genetiche e il tempo, permettendo, al tempo stesso, di datare gli eventi di divergenza evolutiva²¹⁵. 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.

L'utilizzo di modelli di un orologio molecolare per stimare i tassi evolutivi e i tempi di divergenza di sequenze eterocrone (sequenze campionate in intervalli di tempo relativamente lunghi) hanno recentemente reso possibile lo sviluppo di metodi basati sulla coalescenza, per studiare le dinamiche delle popolazioni virali a livello intra- e inter-ospite su scala temporale. La teoria della coalescenza si basa sulla correlazione tra il tempo in cui due *taxa* o due *lineage* hanno condiviso uno stesso antenore e le dimensioni della popolazione in esame, rappresentando così un ottimo strumento per valutare le modificazioni di una popolazione virale nel tempo ²¹⁶.

Le dinamiche demografiche e di migrazione della popolazione umana, svolgono inoltre un ruolo importante nel determinare la forma della struttura filogenetica delle popolazioni virali, per cui la filogeografia (studio della diffusione a livello geografico dei ceppi virali) rappresenta una parte cruciale della caratterizzazione filodinamica di un'infezione ²¹⁷. L'approccio filogeografico si basa sull'esistenza di una relazione tra evoluzione molecolare di un organismo e la sua distribuzione spaziale. Questa relazione richiede normalmente milioni di anni di evoluzione perchè possa essere verificata, ma nel caso di virus ad RNA, caratterizzati da una elevatissima velocità evolutiva e dalla produzione di milioni di generazioni in un tempo brevissimo, è possibile osservare tale relazione anche nell'intervallo di tempo di pochi anni che intercorre tra due campionamenti. La combinazione di analisi genetiche e geografiche rivelano non solo il luogo di origine delle infezioni, ma anche la modalità di trasmissione e il tasso di diffusione spazio-temporale.

Originariamente metodi statistici per ricostruire la distribuzione geografica e spaziale di un virus sulla base dei luoghi di campionamento di isolati virali erano basati sulla massima parsimonia, criterio che giustifica l'intera filogenesi con il minimo numero di migrazioni ²¹⁴. Diversi approcci filogenetici per inferire la distribuzione spaziale di un organismo sulla base della filogenesi, sono attualmente disponibili. Più recentemente, è stato sviluppato un metodo di inferenza statistica (metodo bayesiano) basato sui modelli di coalescenza che

permette la simultanea ricostruzione della storia temporale e spaziale di un'epidemia sulla base degli isolati campionati in modo casuale a tempi conosciuti e in luoghi diversi ²¹⁸.

1.8.1 Filodinamica del genotipo D di HBV

Il genotipo D è il genotipo più diffuso nel nord-est Europa, nella regione orientale e centrale del Mediterraneo, in Africa Settentrionale, nel Medio Oriente ed è inoltre molto diffuso nel sub-continente indiano e in un gruppo di isole dell'Oceano Indiano (Nicobare e Andaman) caratterizzate da alti livelli di endemicità per HBV ²¹⁹.

Nove sottogenotipi di HBV-D sono stati finora descritti (Figura 12): il D1 è il sottogenotipo più diffuso in Grecia, Turchia e nord Africa ^{220, 221}; il D2 nell'Europa nord-orientale (Russia, Bielorussia, Estonia) e in Albania ^{222, 223}; il D3 in Italia e in Serbia ^{130, 224}, il D4 è il sottogenotipo dominante in Oceania ¹¹⁷; il D5 nelle tribù primitive che vivono in India, dove si trovano anche una serie di diversi sottogenotipi D¹³¹; il D6 in Papua e Indonesia ⁹⁸ e il D7 in Tunisia e Marocco ^{128, 225}. Infine i sottogenotipi recentemente descritti D8 e D9 si trovano in Nigeria e India e sono stati riconosciuti come forme ricombinanti di HBV-D con HBV-E ¹³² e HBV-C ¹³³.

In Italia, il sottogenotipo D3 è associato all'esposizione parenterale e alla tossicodipendenza ²²⁴. In particolare il virus isolato da soggetti IVDUs (*intravenous drug users*) è caratterizzato da un certo numero di mutazioni a livello dei geni S e P ²²⁴, mutazioni descritte anche in soggetti IVDUs che vivono in altri paesi industrializzati ^{117, 226}.

Uno studio preliminare effettuato in Giappone ha suggerito un'origine relativamente recente del genotipo D che risale ai primi anni del XX secolo ²²⁶. In particolare tale studio ha mostrato una crescita esponenziale dell'epidemia a causa di questo genotipo a partire dal 1940, con un rapido incremento del numero di infezioni attorno agli anni settanta ed il raggiungimento del plateau a partire dal 1980, dato in linea con la diminuzione dell'incidenza di infezioni da HBV acuta nei paesi sviluppati.

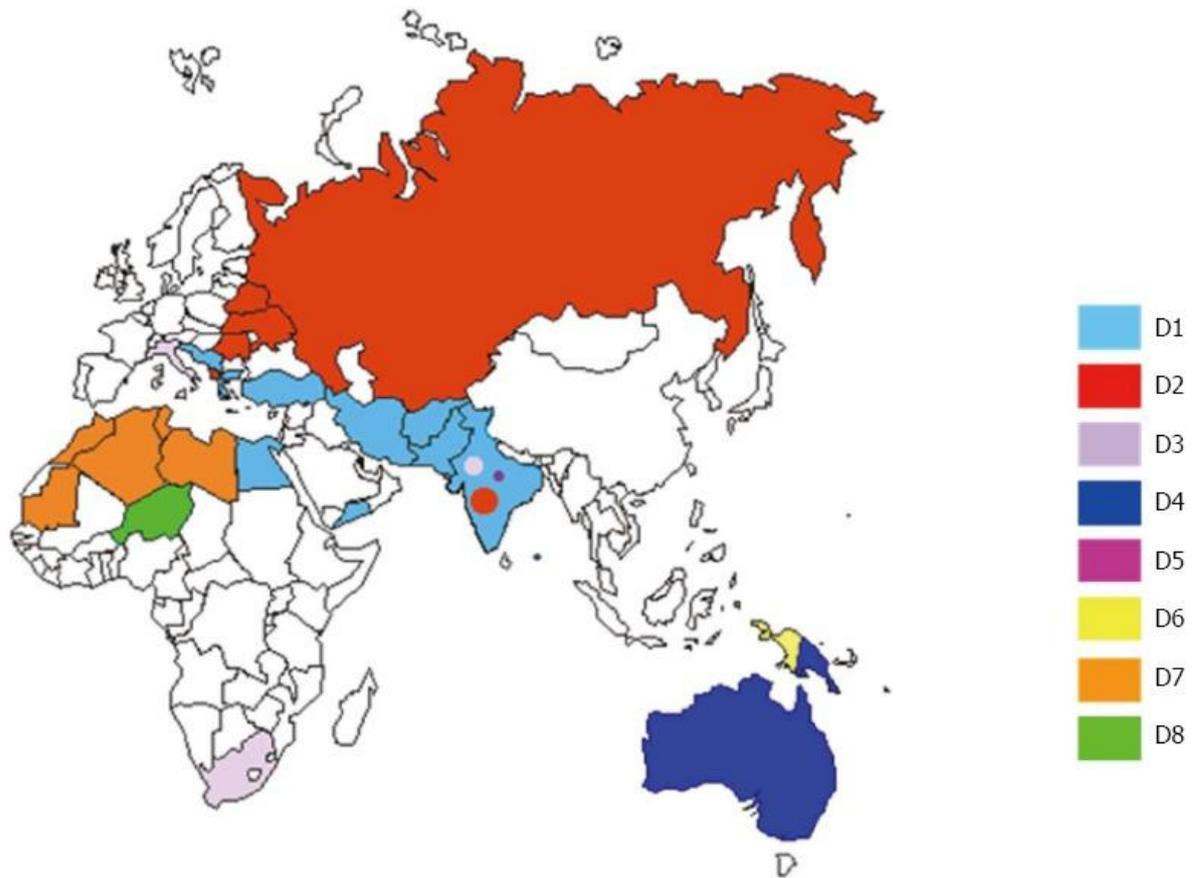


Figura 12. Distribuzione dei sottogenotipi D del virus dell'epatite B.

1.8.2 Filodinamica del genotipo A di HBV

Il genotipo A è il secondo genotipo più diffuso a livello mondiale, ed è diffuso in Africa, Asia, Europa ed America. Gli isolati di genotipo A sono stati ulteriormente classificati in sette gruppi evolutivi distinti che sono i sottogenotipi A1, A2, A3, A4, A5, A6 e A7. Il sottogenotipo A1 è particolarmente diffuso in Africa Meridionale e Orientale (Sud Africa, Uganda, Malawi, Tanzania, Congo e Somalia), nell'Asia del Sud (India, Filippine, Bangladesh e Nepal) ^{106, 117, 131, 227} e in Sud America (Brasile). Il sottogenotipo A2 è il più diffuso in Europa e Nord America ¹¹⁷ ed è stato inoltre isolato in Sud Africa ¹⁰¹. Hannoun ha suggerito che il genotipo A ha avuto origine in Africa e ha ipotizzato un'importazione di HBV-A2 dall'Africa all'Europa dai marinai portoghesi nei secoli XVI e XVII, e l'arrivo di HBV-A1 in Asia a seguito di scambi commerciali e viaggi tra Africa Orientale e Asia Meridionale ¹¹¹.

Più recentemente sono stati descritti i sottogenotipi A3 (nella popolazione dei Pigmei e Bantu che vivono in Camerun e Gabon) ^{109,228}, "tentative A4" (isolato nel Mali) e A5 (in pazienti dalla Nigeria) ¹¹². HBV-A5 è stato isolato anche ad Haiti; probabilmente tale sottogenotipo è stato importato nell'isola caraibica durante il periodo della tratta degli schiavi (tra il XVIII e il XIX secolo) da una regione vicino all'attuale Nigeria (ex golfo del Benin) dove A5 era il sottogenotipo prevalente. Solo di recente HBV-A5 è stato sostituito dal genotipo E e dal sottogenotipo A3, ad oggi i ceppi più diffusi in quella zona ¹¹³. Il sottogenotipo A6 è stato riportato solo in pazienti belgi di origine africana ¹¹⁵. Un nuovo sottogenotipo A7 ("tentative A7") è stato isolato in Camerun ¹¹⁴ (Figura 13).

Secondo un recente lavoro i sottotipi A3, "tentative A4", A5 e "tentative A7" dovrebbero essere classificati all'interno di un unico sottogenotipo chiamato "quasi sottogenotipo-A3" ²²⁹.

Il primo studio sull'origine e sulla dinamica di popolazione di HBV-A, in particolare del sottogenotipo A2 ⁵⁹, ha riportato l'esistenza di un unico ceppo filogeneticamente omogeneo, prevalente tra i soggetti italiani omosessuali HIV-positivi, caratterizzato da un tMRCA (*time of the most recent common ancestor*), ovvero il tempo medio del più recente antenato comune, risalente al 1966, suggerendo che tale sottogenotipo si sia sviluppato in gran parte come il risultato della trasmissione sessuale tra la fine del 1960 e l'inizio del 1980.

È interessante notare che una singola variante di sottogenotipo A2 è stata isolata da pazienti MSM (*men who have sex with men*) che vivono in Olanda ²³⁰, paese caratterizzato da basso livello endemicità per HBV (0,3-0,5 %), in cui la trasmissione sessuale rappresenta la principale via di trasmissione del virus ²³¹. Ulteriori studi condotti in Olanda ^{232, 233} si sono dimostrati particolarmente importanti perché rappresentano il primo tentativo di applicare l'approccio filodinamico al monitoraggio dell'efficacia del programma di vaccinazione mirata alla popolazione MSM, osservando una significativa diminuzione del numero effettivo di infezioni dovute al sottogenotipo A2 tra i soggetti MSM, corroborata da una parallela diminuzione dell'incidenza di infezioni acute da HBV.

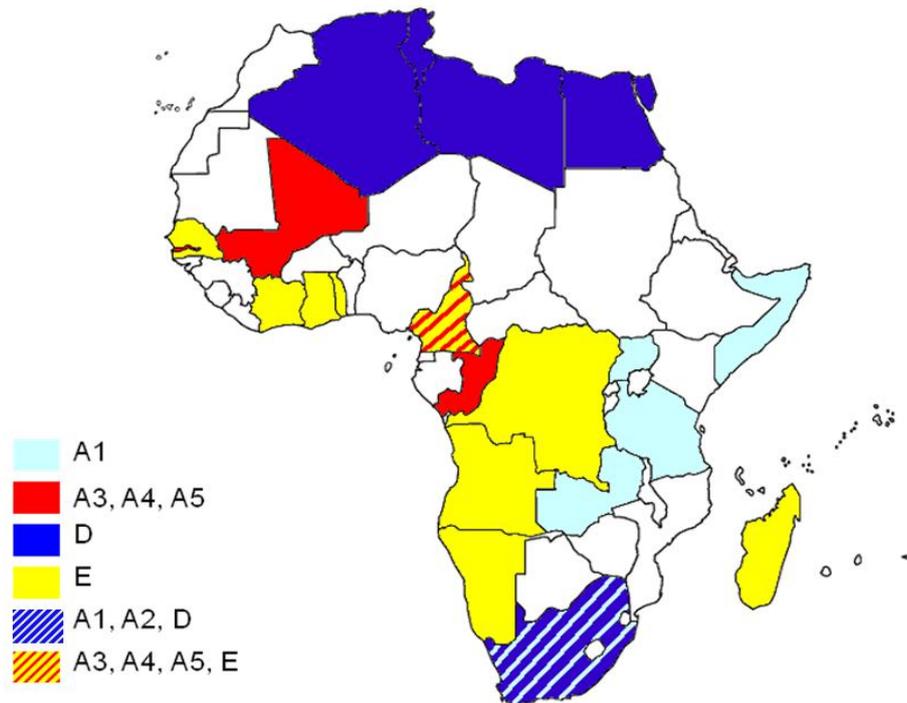


Figura 13. Distribuzione del genotipo A e degli altri genotipi e sottogenotipi di HBV prevalenti in Africa.

E' stato inoltre scoperto che HBV-A2 è molto diffuso e mostra una diversità genetica relativamente bassa tra gli omosessuali in Giappone ²³⁴, suggerendo così una recente e distinta epidemia di questo sottogenotipo a livello mondiale, in questa categoria a rischio.

Al momento mancano studi filogeografici e di distribuzione spaziale del genotipo A. L'unica eccezione è un recente studio condotto da Kramvis ²³⁵, che ha utilizzato un metodo di massima parsimonia per ricostruire la diffusione a livello mondiale del sottogenotipo A1. Secondo tale studio HBV-A1 si sarebbe originato con tutta probabilità in Africa, da dove si sarebbe diffuso in Asia e in America Latina nello stesso arco di tempo. La tratta degli schiavi e la colonizzazione, in particolare la deportazione di schiavi dall'Africa Orientale all'India (fino alla fine del XIX secolo), la colonizzazione belga del Congo (nella prima metà del XX secolo), e la tratta degli schiavi europea (fino al XIX secolo), avrebbero quindi giocato un ruolo fondamentale nella circolazione di tale sottogenotipo a livello mondiale.

3. Scopo dello studio

L'infezione da virus dell'epatite B (HBV), rappresenta uno dei più importanti problemi di Sanità pubblica a livello mondiale, essendo la causa di malattie epatiche acute e croniche, tra cui la cirrosi ed il carcinoma epatocellulare.

La prevalenza di infezione da HBV in Europa è molto variabile. Il livello endemico di questa infezione è diminuito in modo significativo in alcune regioni del Mediterraneo quali Italia, Grecia, Spagna e Francia, mentre in altri paesi come Albania, Bulgaria, Turchia e Israele si hanno ancora livelli di alta endemicità.

Diversi studi hanno finora ricostruito la storia dei genotipi di HBV sulla base della loro filogenesi, ma hanno prodotto risultati contrastanti. La limitazione principale alla ricostruzione della storia dei genotipi di HBV è data da una disconcordanza sul tasso evolutivo del virus, dovuta a diversi metodi di calibrazione utilizzati.

Nel presente progetto di tesi si è voluto condurre uno studio epidemiologico-molecolare mediante l'analisi delle sequenze nucleotidiche del gene P di HBV di pazienti Italiani affetti da epatite cronica di tipo B, allo scopo di determinare i genotipi/sottogenotipi di HBV circolanti e studiarne la distribuzione in relazione a fattori demografici e di rischio e la prevalenza in Italia. Inoltre, al fine di ricostruire la storia epidemiologica e la filodinamica (l'origine, la dinamica di popolazione e le vie di dispersione geografica) dei genotipi e sottogenotipi di HBV maggiormente diffusi in Italia e in Europa, le sequenze di HBV Italiane sono state allineate con sequenze di riferimento ottenute in diverse aree del mondo, recuperate dalle principali banche dati. Lo studio è stato condotto mediante l'utilizzo di nuove metodologie di analisi filogenetica, basate sulla teoria della coalescenza e sull'approccio filogeografico, che ricostruisce i più probabili flussi di migrazione sulla base delle relazioni filogenetiche di ceppi isolati in luoghi diversi.

3. Materiali e metodi

3.1 CASISTICA

Nel presente lavoro sono state analizzate le sequenze nucleotidiche ottenute da 230 soggetti italiani con diagnosi di infezione cronica da virus dell'epatite B, caratterizzati da positività per l'antigene di superficie (HBsAg) per periodi superiori a 6 mesi, in presenza o meno di sintomi clinici.

Di tutti i pazienti erano disponibili i parametri anagrafici ed epidemiologici, quali età, sesso, paese d'origine e coinfezione HBV/HIV.

3.2 METODI

3.2.1 Modalità di raccolta, preparazione e conservazione dei campioni

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 g per 10 minuti a 4°C. Il siero così ottenuto è stato suddiviso in aliquote da 1 ml e quindi congelato a -20°C fino al momento dell'utilizzo.

3.2.2 Estrazione del DNA

L'estrazione del DNA 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 a 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:

- In una provetta sterile da 1,5ml si aliquotano 200 μ l di siero, un uguale volume di tampone di lisi (Lysis Buffer MV1), 20 μ l di Proteinasi K (20 mg/ml) e 4 μ l 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 μ l di Resuspended -Beads e 600 μ l 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 μ l di Wash Buffer MV3 nel pozzetto B,
 - 500 μ l di Wash Buffer MV4 nel pozzetto C,
 - 550 μ l di Wash Buffer MV5 nel pozzetto D,
 - 50 μ l di acqua RNasi free nel pozzetto E.

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

Il DNA, 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.

- L'eluato, recuperato alla fine della procedura, viene raccolto nel pozzetto E e conservato a 4°C fino al momento dell'utilizzo. Per escludere la possibilità di contaminazioni tra i diversi campioni, ogni estrazione di DNA 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.

3.2.3 Amplificazione del DNA virale in regione pre-S/S e P mediante nested-PCR

In seguito all'estrazione di DNA si effettua l'amplificazione tramite nested-PCR. Tale tecnica è una variazione della PCR classica e prevede due reazioni di amplificazione successive utilizzando due diverse coppie di *primer*. La sequenza di DNA bersaglio viene infatti utilizzata per disegnare due coppie di oligonucleotidi, una delle quali (*primers* interni) si appaierà a sequenze interne alla regione già amplificata dalla prima coppia di *primers* (*primers* esterni). In breve: un'aliquota del prodotto della prima PCR (esterna) viene amplificata in presenza di una seconda coppia di *primers* specifici e più interni alla sequenza ottenuta durante la prima fase di amplificazione.

Questa tecnica consente di aumentare di molto sia la specificità che la sensibilità della reazione di amplificazione ed è quindi particolarmente utile per la ricerca di sequenze bersaglio scarsamente rappresentate nel campione in esame (Figura 14).

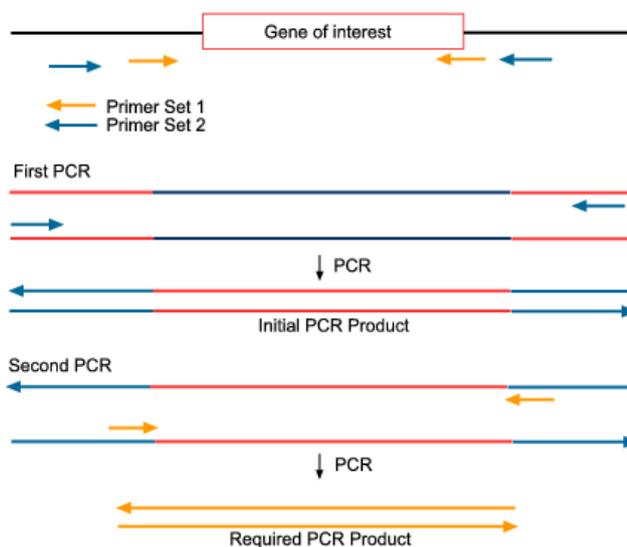


Figura 14. Schema rappresentativo della nested-PCR.

Nel presente lavoro è stata utilizzata una nested-PCR più complessa, che ha previsto l'amplificazione di tre frammenti parzialmente sovrapposti che coprono quasi completamente la regione genomica S (parziale *pre-S1*, *pre-S2* e S completi) e

di conseguenza anche una porzione consistente del gene P di HBV, essendo le ORF di S e P sovrapposte.

Il kit usato per l'amplificazione dei tre frammenti è *GoTaq® Hot Start Polymerase* (Promega, Madison, WI, USA).

Questa reazione di nested-PCR si avvale di due successivi cicli (step) di amplificazione:

Primo step

Questa prima reazione di PCR consente di amplificare 2 diversi frammenti della regione *pre-S/S* e P, rispettivamente di 911 e 907 pb, in presenza delle seguenti coppie di *primers* esterni:

1° frammento (911 pb: nt 2960-655)

senso (HBV-7: nt 2960-2980): 5'-A(AG)T CC(AC) GAT TGG GAC (CT)(AT)C AA-3'

antisenso(ERS-IN3: nt 655-637): 5'-TGA GGC CCA CTC CCA TAG G-3'

2° frammento (907 pb: nt 202-1108)

senso (HBV-9: nt 202-220): 5'-GGC GG(GT) GT(GT) TTT CTT GTT G-3'

antisenso (HBV-6: nt 1108-1085): 5'-AAG TTG GCG A(AG)A A(AG)(AG) (CT)(AG)A AAG CCT-3'

Nel primo step di questa PCR, vengono preparate 2 diverse mix di PCR per ciascun frammento, utilizzando i primers senso e antisenso rispettivi:

Reagente	Volume per campione
H ₂ O sterile	21.7 µl
Buffer 5X colorless	10 µl
dNTP (2.5 mM)	4 µl
MgCl ₂ (25 mM)	3 µl
Primer senso (50 pmoli/µl)	0.5 µl
Primer antisenso (50 pmoli/µl)	0.5 µl
GoTaq Polymerase (5u/µl)	0.3 µl
Volume finale	40.0 µl

A questa mix vengono aggiunti 10 μ L di DNA virale estratto, raggiungendo così un volume totale di 50 μ l, la quale viene poi sottoposta ad amplificazione in un termociclatore così impostato.

Temperatura	Tempo	N° cicli	Processo
94°C	2'	1	Attivazione Taq
94°C	45''	45	Denaturazione DNA
57.5°C	45''		Legame dei primer
72°C	1'30''		Estensione
72°C	10'	1	Estensione finale
4°C	∞		Conservazione

Secondo step

Attraverso il secondo step di PCR vengono amplificati tre diversi frammenti, parzialmente sovrapposti tra loro, il primo dei quali, di 536 pb, interno al primo frammento ottenuto dalla PCR esterna di 911 pb (nt 2960-655), ed i rimanenti, di 476 e 700 pb, interni al secondo frammento di 907 pb prodotto dal primo step della nested-PCR (nt 202-1108).

Allo scopo sono state utilizzate coppie di *primers* differenti per i tre frammenti amplificati:

1° frammento (536 pb: nt 2960-280)

senso (HBV-7: nt 2960-2980): 5'-A(AG)T CC(AC) GAT TGG GAC (CT)(AT)C AA-3'

antisenso (HBV- 4: nt 280-257): 5'-TAG AAA ATT GAG AGA AGT CCA CCA-3'

2° frammento (476 pb: nt 180-655)

senso (HBV-5: nt 180-203): 5'-TAG GAC CCC T(GT)C TCG TGT TAC AGG-3'

antisenso (ERS-IN3: nt 655-637): 5'-TGA GGC CCA CTC CCA TAG G-3'

3° frammento (700 pb: nt 409-1108)

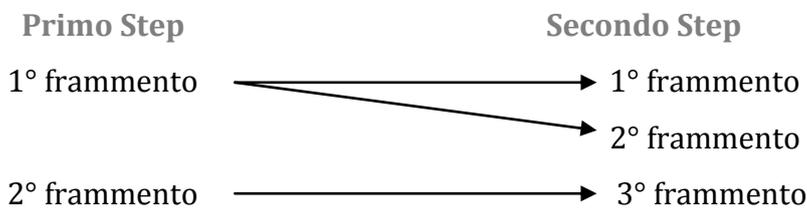
senso (ERS-IN2: nt 409-428): 5'-CAT CCT GCT GCT ATG CCT CA-3'

antisenso (HBV-6: nt 1108-1085): 5'-AAG TTG GCG A(AG)A A(AG)(AG) (CT)(AG)A AAG CCT-3'

Le 3 mix diverse di PCR sono così composte:

	Volume per campione
H ₂ O sterile	25.8 µl
Buffer 5X green	10 µl
dNTP (2.5 mM)	4 µl
MgCl ₂ (25 mM)	3 µl
Primer senso (50 pmoli/µl)	0.5 µl
Primer antisenso (50 pmoli/µl)	0.5 µl
GoTaq Polymerase (5u/µl)	0.2 µl
Volume finale	44.0 µl

A queste vanno aggiunti 6 µL di amplificato del primo step secondo il seguente schema:



L'amplificazione del 1° frammento avviene in termociclatore alle seguenti condizioni:

Temperatura	Tempo	N° cicli	Processo
94°C	2'	1	Attivazione Taq
94°C	40''	45	Denaturazione DNA
55°C	40''		Legame dei primer
72°C	1'20''		Estensione
72°C	10'	1	Estensione finale
4°C	∞		Conservazione

Il profilo di amplificazione del 2° e del 3° frammento è il seguente:

Temperatura	Tempo	N° cicli	Processo
94°C	2'	1	Attivazione Taq
94°C	45''	45	Denaturazione DNA
59°C	45''		Legame dei primer
72°C	1'30''		Estensione
72°C	10'	1	Estensione finale
4°C	∞		Conservazione

Tutte le PCR effettuate nel corso dello studio comprendevano controlli negativi e positivi, nonché un campione bianco (non contenente DNA), ai fini di validare ogni singola corsa.

3.2.4 Rilevamento dei prodotti di PCR

Il rilevamento dei prodotti di amplificazione si ottiene 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 viene incorporato un colorante fluorescente, etidio bromuro (0,01 µg/µl), che si intercala tra le basi dell'acido nucleico.

E' 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.

3.2.5 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:

1. il buffer BPI, che permette il legame del DNA alla resina della colonnina QIAquick, deve avere un pH ≤ 7.5 (color giallo),
2. il buffer PE, con il quale viene eluito il DNA dopo la purificazione, deve essere addizionato di etanolo al 96-100%,
3. 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 BPI,
- si miscela attentamente con la pipetta fino ad ottenere una soluzione omogenea,
- occorre controllare che il colore della miscela sia giallo (simile al colore del BPI senza il campione di PCR), se il colore della miscela è arancione o viola, occorre aggiungere 10 μ L di sodio acetato 3 M (pH5.0) e miscelare. Il colore della miscela virerà verso il giallo,
- 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,
- si scarta l'eluito e si riposiziona la colonna nella medesima provetta,
- per lavare, si aggiungono 0.75mL 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 μ L di buffer EB o di acqua distillata direttamente sopra la membrana della colonnina e si centrifuga per un minuto. In alternativa per aumentare la concentrazione di DNA, si possono aggiungere 30 μ L di buffer EB sulla membrana e si attende un minuto prima di centrifugare,
- si eliminano le colonnine. Il DNA purificato deve essere conservato a 4°C.

3.2.6 Sequenziamento delle regioni genomiche amplificate

3.2.6.1 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.

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

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.

L'utilizzo di un enzima termostabile permette di svolgere l'attività di polimerizzazione del DNA ad alte temperature, minimizzando i problemi causati dalla formazione di strutture secondarie del DNA, di legami aspecifici e da interazioni fra i primer. Inoltre, come nella normale PCR, l'enzima resiste alla denaturazione quando la miscela di sequenziamento è incubata a 96° C, permettendo la realizzazione dei cicli successivi di estensione.

Miscela di sequenza

La miscela di reazione si compone di 3.2 µl di primer, 1.1 µl di BigDye, 3-5 µl di DNA purificato, stimato in base all' intensità rilevata su gel, e acqua fino ad arrivare ad un volume finale di 10 µl.

Il profilo termico di amplificazione per sintetizzare i frammenti di sequenza è il seguente:

Temperatura	Tempo	N° cicli	Processo
96°C	1'30''	1	Attivazione BigDye
95°C	20''	25	Denaturazione DNA
55°C	20''		Legame dei primer
60°C	3'		Estensione dei primer
4°C	∞		Conservazione

3.2.6.2 Purificazione dei prodotti di sequenza

Prima di procedere al sequenziamento dei campioni amplificati è necessaria la purificazione dei prodotti di sequenza da possibili contaminanti, quali *primers*, dNTPs non incorporati e l'enzima DNA-polimerasi in eccesso che potrebbero disturbare la successiva reazione 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), lasciati in ghiaccio per 10 minuti e centrifugati per 30 minuti a 14.000 g.

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

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

3.2.6.3 Sequenziamento

Per il sequenziamento è stato utilizzato il sequenziatore automatico ABI Prism 3130 XL (Applied Biosystems), basato sulla tecnologia dell'elettroforesi capillare. La parte più importante di questo sistema è un'unità intercambiabile composta da sedici capillari in silice fusa; questi, una volta riempiti di polimero, permettono la separazione elettroforetica dei frammenti di DNA marcati.

Si utilizzano 2.1 μ L di campione, dopo averlo precedentemente risospeso in 10 μ L di acqua bidistillata, che si aggiungono a 8 μ L di una soluzione denaturante di formammide.

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, direttamente collegato all'apparecchio.

Il sequenziatore è stato impostato con i seguenti valori: voltaggio 12,2 KV, corrente 67-77 mA, laser 15 mW e temperatura 50°C. Ogni corsa ha una durata di 3,5 ore.

3.2.6.4 Analisi dei dati di sequenza

L'analisi dei dati si esegue utilizzando il *Software Sequence Navigator* (Applied Biosystems Inc., Foster City, California), che permette di allineare le sequenze ottenute dai diversi primer, assemblare una sequenza consenso utilizzabile per le analisi successive, tradurla in aminoacidi e confrontarla con una sequenza di riferimento, per poter effettuare in un secondo momento l'analisi filogenetica.

3.2.7 Creazione dei dataset

Al fine di ricostruire la filogenesi, la dinamica di popolazione e la filogeografia dei genotipi di HBV più diffusi in Italia, abbiamo costruito più dataset corrispondenti ai diversi genotipi/sottogenotipi di HBV. I dataset sono stati costruiti aggiungendo alle sequenze caratterizzate nel presente lavoro, sequenze dello stesso genotipo/sottogenotipo recuperate dalle principali banche dati pubbliche e isolate in diversi Paesi. Di tutte le sequenze, era noto il luogo di isolamento; l'anno di isolamento era noto solo per una parte di essi.

I dataset sono descritti in dettaglio nella sezione Risultati. Gli accession number e i luoghi di campionamento delle sequenze di riferimento utilizzate per la costruzione dei dataset sono riassunti in Appendice 1 e 2.

3.2.8. 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.

3.2.8.1 Discriminazione tra i genotipi mediante strumenti disponibili in Rete

Inizialmente ogni sequenza viene sottoposta ad una iniziale genotipizzazione, utilizzando strumenti disponibili *on line* come l'*Oxford HBV Automated Subtyping Tool v. 1.0* (<http://www.bioafrica.net/rega-genotype/html/subtypinghbv.html>). Questo sito rispetto ad altri disponibili in rete utilizza l'analisi filogenetica per l'analisi delle sequenze e non solo metodi basati sul confronto e la similarità.

In dettaglio, ogni singola sequenza in formato FASTA (un file di testo formato da una prima riga descrittiva identificata dal carattere iniziale ">", seguita dalla sequenza nucleotidica a partire dalla seconda riga) viene inserita nella finestra di immissione della pagina *web* e sottoposta ad analisi filogenetica; è possibile inoltre visualizzare l'albero generato e l'immagine ottenuta dall'analisi di *bootscan*.

Nonostante questo sito venga ormai utilizzato per la sottotipizzazione in numerosi articoli, bisogna ricordare che non è possibile pubblicare nel proprio lavoro l'albero ottenuto con questo sito e che l'analisi non tiene conto delle caratteristiche delle proprie sequenze poiché utilizza un *data set* fisso di sequenze di riferimento. Non sempre l'assegnazione è certa e supportata.

3.2.8.2 Genotipizzazione delle sequenze in studio

La genotipizzazione è stata eseguita su 230 sequenze lunghe 1182 nucleotidi e comprendenti le ORFs codificanti per S, Pre-S2 e un tratto di 59 aminoacidi al C-terminale di Pre-S1, insieme a una ORF sovrapposta codificante per un frammento di 394 aminoacidi della polimerasi virale.

Tutte le sequenze sono state analizzate mediante il software BioEdit v. 7.2.5, contenente il programma ClustalX ²³⁶ che ci ha consentito di allineare le sequenze nucleotidiche in studio con sequenze di riferimento, ottenute dalle principali banche dati, quali la *GenBank*, presso il *National Center for Biotechnology Information* (USA- <http://www.ncbi.nlm.nih.gov/Genbank>).

Per la comparazione delle sequenze italiane abbiamo utilizzato 20 sequenze di riferimento, specifiche per ogni genotipo di HBV.

L'analisi filogenetica è stata effettuata utilizzando i programmi SEQBOOT,

DNADIST, NEIGHBOR e CONSENSE contenuti nel pacchetto *software* PHYLIP (ver. 3.695)²³⁷ e MEGA 6 (*Molecular Evolutionary Genetics Analysis*)²³⁸

In particolare, per la scelta del modello di calcolo delle distanze genetiche, abbiamo utilizzato il programma jModelTest v. 2²³⁹, che consente di scegliere il modello evolutivo più opportuno sulla base di un test denominato “*likelihood ratio test*”.

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 programma MEGA, e il metodo *Maximun-Likelihood* (ML), utilizzando un nuovo algoritmo “hill-climbing”, incluso nel server Phym1 3.0 (<http://www.atgc-montpellier.fr/phym1/>)²⁴⁰.

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 *Maximun-Likelihood* invece 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 75% (dopo 1000 replicati in NJ o 200 in Phym).

La caratterizzazione molecolare del genotipo/sottogenotipo di HBV infettante è stata effettuata mediante l’analisi filogenetica, confrontando ciascuna sequenza del gene P in studio con opportune sequenze di riferimento.

3.2.8.3 Analisi dei ricombinanti con Simplot

SIMPLOT vs 3.5.1 (Stuart C. Ray, 1998) è un programma che si avvale del metodo per la ricerca di ricombinazioni genomiche e permette di studiare sequenze con mosaicismi.

Il programma può utilizzare due distinti *output* grafici. Uno, detto *similarity plot* è rappresentato da una serie di linee che riflettono la similarità (o la distanza) di ogni sequenza (o gruppo di sequenze) di riferimento rispetto alla sequenza *query* da caratterizzare, considerando una finestra scorrevole lungo l'intera sequenza. Il secondo è generato mediante la cosiddetta analisi di *bootscanning*, che permette di calcolare la percentuale di alberi permutati generati considerando sempre una finestra scorrevole che si sposta progressivamente lungo l'intera sequenza. La dimensione della finestra e dello slittamento possono essere variati.

I punti di ricombinazione vengono individuati sia visivamente sia ripetendo l'analisi di *bootscanning* della sequenza *query* con le sequenze di riferimento.

3.2.8.4 Likelihood mapping

La quantità di informazione filogenetica contenuta nei data set di sequenze in studio è stata stimata mediante l'analisi *Likelihood-mapping*²⁴¹. Il *Likelihood-mapping* è un metodo grafico che si basa sui valori di *Maximum Likelihood* per tre tipologie possibili di alberi senza radice, costruiti su un gruppo di quattro sequenze (quartetti), selezionate a random dal dataset. La visualizzazione del contenuto filogenetico del nostro allineamento di sequenze è stata ottenuta con l'analisi di 10,000 quartetti random.

Le tre tipologie di albero di ogni quartetto sono rappresentate come un punto all'interno di un triangolo equilatero. Il triangolo è diviso in 7 regioni (Figura 15). I tre trapezoidi agli angoli (regioni 1 + 2 + 3) rappresentano il *tree-like phylogenetic signal*, in cui si ha una completa risoluzione del segnale filogenetico, ovvero dove un albero è nettamente migliore rispetto agli altri due. Le tre regioni ai lati (4 + 5 + 6) rappresentano il *network-like phylogenetic signal*, in cui è difficile scegliere tra

due tipologie di albero. La regione centrale (7) rappresenta lo *star-like phylogenetic signal*, in cui le tre tipologie di albero sono scarsamente sostenute.

La percentuale dei punti appartenenti alle tre principali aree del triangolo equilatero suggerisce se i dati sono adeguati o non adeguati alla ricostruzione filogenetica.

In particolare, un'analisi di likelihood-mapping che indica una percentuale di punti nella *star-like area* o nella *network-like area* uguale o maggiore del 30%, indica che i dati non sono adatti e affidabili per lo studio filogenetico.

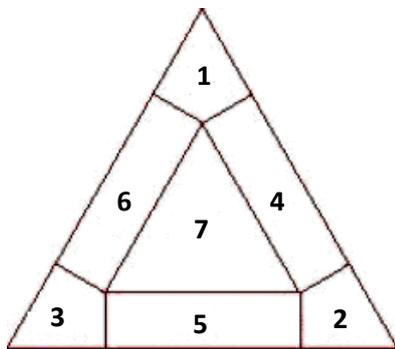


Figura 15. Diagramma di likelihood-mapping

3.2.8.5 Analisi filogenetica mediante i metodi Bayesiani

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. 1.8.0* ²⁴².

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 il modello non parametrico *Bayesian skyline plot* (BSP) 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 H_0 . Solo valori ≥ 6 sono considerati significativi ²⁴³.

Per l'analisi delle nostre sequenze è stato selezionato un modello di sostituzione nucleotidica GTR-I-G (*General-Time Reversible + Invariant sites + Gamma distribution rate*) ²⁴⁴.

Questo tipo di analisi richiede inoltre dei sistemi di inferenza statistica che ne supportino la significatività.

Il programma MrBayes ²⁴⁵ si fonda sulle teorie della statistica Bayesiana ²⁴⁶, che calcola la probabilità a posteriori dell'albero (ovvero la probabilità di ottenere un certo albero in base ai dati forniti), diversamente da quanto avviene nel *Maximum Likelihood*, che ricerca la più alta verosomiglianza dell'albero (ovvero la probabilità di ottenere i dati in base all'albero generato). La probabilità a posteriori è così calcolata, secondo le teorie Bayesiane, a partire dalla probabilità a posteriori dell'analisi filogenetica e dalla *likelihood* dell'albero.

Le probabilità a posteriori stimate vengono approssimate mediante l'algoritmo *Markov Chain Monte Carlo* ²⁴⁷ (MCMC). Inizialmente viene generato uno *starting tree* casuale che viene poi confrontato con un nuovo albero, ottenuto modificando in modo sistematico il precedente, e i relativi parametri; il secondo albero viene quindi accettato o rifiutato sulla base della probabilità a posteriori in uno specifico processo di campionamento. Questo sistema di confronto viene ripetuto a catena fino ad ottenere l'albero con il più alto prodotto di probabilità a posteriore nodale tra tutti quelli generati (definito *consensus tree* o *Maximum clade credibility tree*). In questo modo si ottiene una stima della significatività dei *cluster* dell'albero, fornendo così dei valori di supporto statistico. In particolare clades con una probabilità posteriore ≥ 0.8 vengono considerati significativi.

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 è stata poi validata dal software TRACER v.1.5 (<http://tree.bio.ed.ac.uk/software/tracer/>), il quale verifica la stima dell'effettiva

dimensione del campione (ESS, *effective sample size*), dopo un burn-in del 10%. Solo valori di ESS > 250 sono stati accettati.

L'incertezza delle stime ottenute è stata indicata con intervalli di *Highest Posterior Density* del 95% (HPD).

L'albero datato mostrato è stato ottenuto mediante il programma Tree Annotator, selezionando l'albero con il più alto prodotto di probabilità a posteriori.

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 e i tempi dell'ancestore comune più recente (tMRCA, *time of the most recent common ancestor*).

3.2.9 Analisi filodinamica

Al fine di delineare la storia epidemiologica/evolutiva del virus HBV abbiamo eseguito l'analisi di coalescenza che 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 Beast²⁴², 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.

Sono stati testati due modelli demografici parametrici (*constant population size*, *exponential growth*) e il modello non parametrico *Bayesian skyline plot* (BSP). 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).

L'impiego del modello parametrico è stato utilizzato per stimare il *tasso di crescita* (r). Sulla sua base è stato stimato il *numero riproduttivo di base* (R_0), che rappresenta il numero medio di casi secondari infettati da un singolo caso primario in una popolazione senza difese immunitarie, usando la formula $R_0=rD+1$. D è la durata dell'infezione nei portatori. Inoltre, è stato calcolato anche il *tempo di raddoppiamento dell'epidemia* $\lambda=\ln(2)/r$.

3.2.10 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 sequenze della nostra casistica in aggiunta a sequenze di riferimento estratte dal database *GenBank* per le quali era disponibile l'anno del campionamento ed era nota l'area geografica.

Inoltre per ricostruzione filogeografica le sequenze sono state sottoposte ad analisi mediante il programma *Spread* ²⁴⁸ che riproduce visivamente la migrazione virale nelle regioni studiate usando Google Earth (<https://www.google.com/earth/>).

3.2.11 Analisi statistica

L'analisi statistica dei dati ottenuti è stata condotta mediante pacchetti *software* quali EPI-INFO, versione 6 e SPSS versione 11.5.

Per la valutazione delle variabili qualitative sono stati impiegati i test di significatività (Chi quadrato) ed il test di Fisher. Le variabili quantitative (comprese le età e le distanze genetiche) sono state espresse mediante medie (\pm deviazione standard) e/o mediane (con campo di variazione-range).

La significatività statistica è stata valutata mediante test OpenEpi e test non parametrici quali il test di Mann-Whitney. 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 CARATTERISTICHE DEMOGRAFICHE DELLA POPOLAZIONE IN STUDIO

Dei 230 soggetti in studio, 39 (i cui campioni provenivano da uno studio retrospettivo effettuato presso gli Istituti di Malattie Infettive e Tropicali dell'Università degli Studi di Milano e dell'Università degli Studi di Pavia nel decennio 1993-2003) presentavano coinfezione HBV/HIV-1 e 191 (i cui campioni sono stati ottenuti presso la Sezione di Virologia del Dipartimento Di Sanità Pubblica-Microbiologia-Virologia e la Clinica di Malattie Infettive e Tropicali dell'Università degli Studi di Milano nel periodo 2002-2007) erano monoinfetti con HBV.

Dei 39 soggetti con coinfezione HBV/HIV, 29 (74.4%) erano maschi e 10 femmine (25.6%) con età media di 42.6 anni (range 23-80 anni).

Per quanto riguarda la nazionalità, 34 (87.2%) erano di nazionalità italiana e 5 (12.8%) di nazionalità extraeuropea.

Tabella 3. Caratteristiche anagrafiche dei 230 pazienti HBV positivi inclusi nello studio.

		HIV-1 positivi	HIV-1 negativi	p	Totale
Genere	M	29 (74.4%)	121 (63.4%)	0.1 (n.s)	150 (65.2%)
	F	10 (25.6%)	70 (36.7%)		80 (34.8%)
Età (DS)		42.6	49.2	<0.001	45.9
Nazionalità	Italiana	34 (87.2%)	188 (98.4%)	0.3 (n.s.)	222 (96.5%)

Dei 191 soggetti con sola infezione cronica da HBV, 121 (63.4%) erano maschi con età media di 49.2 anni, (range 13-84 anni) e la maggioranza (98.4%) era di nazionalità italiana (Tabella 3). I fattori di rischio erano noti solo per 92 soggetti (40%): per 34 pazienti (37%) il fattore di rischio era la tossicodipendenza, per 43 (46.7%) il contatto sessuale (di questi 21 erano MSM -*man who have sex with man*), per 14 soggetti (15.2%) la trasmissione iatrogena e per 1 paziente (1,1%) la trasmissione parenterale.

4.2 CARATTERIZZAZIONE E DISTRIBUZIONE DEL GENOTIPO VIRALE

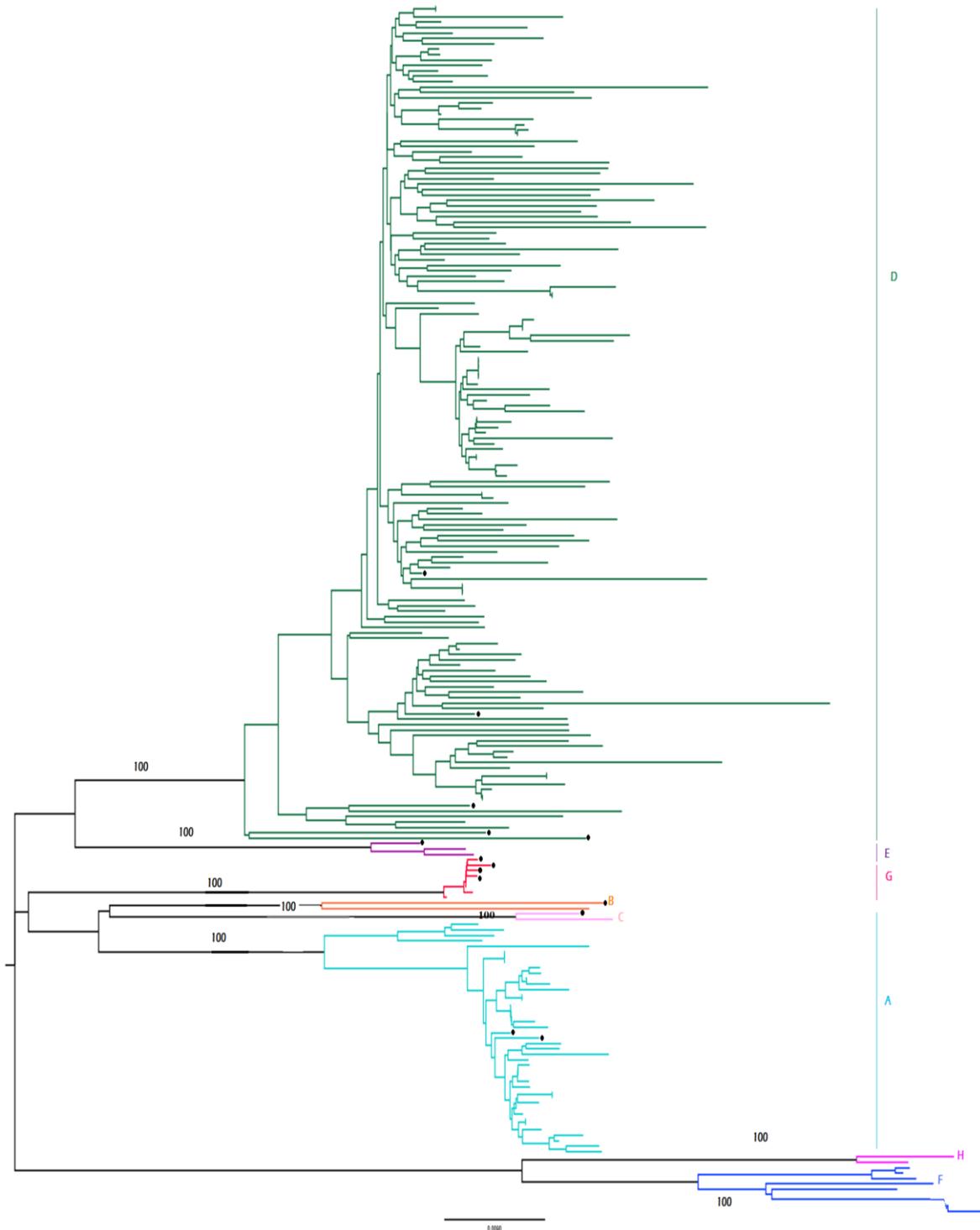


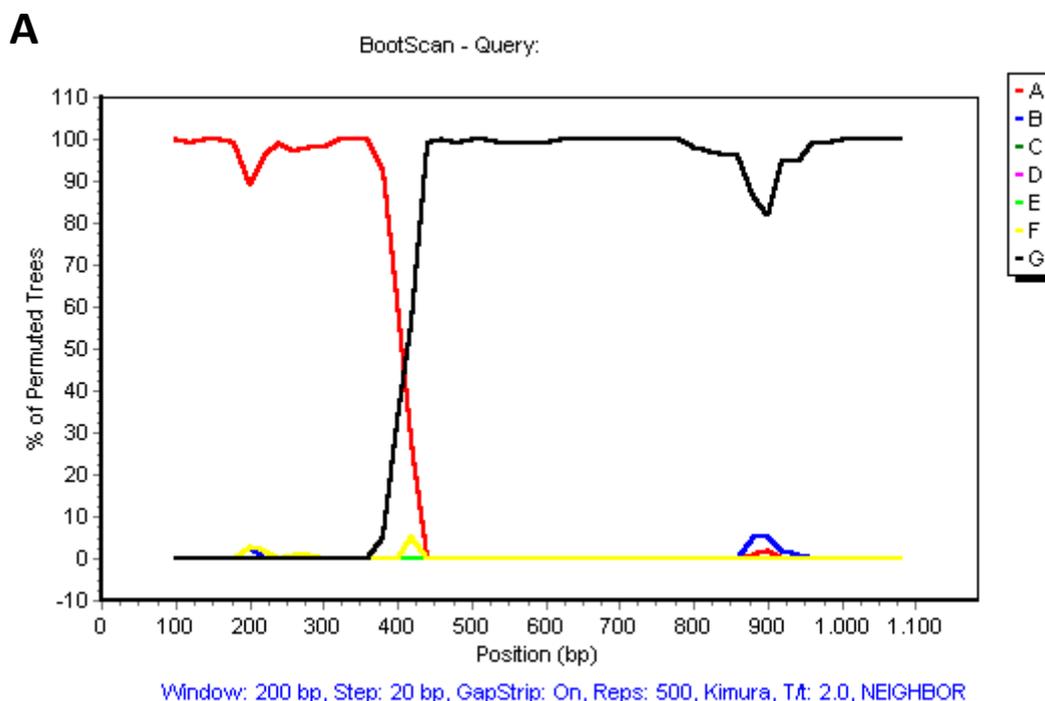
Figura 16. Albero filogenetico del gene *P* comprendente 230 pazienti italiani e 20 sequenze di riferimento genotipo-specifiche.

In Figura 16 è rappresentato l'albero ottenuto dall'analisi delle 230 sequenze italiane del gene P di HBV in comparazione con 20 sequenze di riferimento dei principali genotipi di HBV.

Tale albero, costruito mediante metodo di *Neighbor Joining* (NJ) e modello *Kimura 2 Parametri* (K2P), ha mostrato 8 *clade* (gruppi monofiletici) significativi (*bootstrap*=100), corrispondenti agli 8 genotipi virali (A-H).

L'analisi filogenetica ha permesso la determinazione del genotipo per 227 sequenze: 166 (72.2%) ceppi sono risultati di genotipo D, 43 (18.7%) di genotipo A, 8 (3.5%) di genotipo F, 6 (2.6%) di genotipo G, 1 (0.4%) di genotipo B, 1 (0.4%) di genotipo C e 2 (0.9%) di genotipo E.

Poiché per 3 sequenze (1.3%) non è stata possibile l'assegnazione mediante l'albero filogenetico ad un genotipo virale noto, è stata condotta un'ulteriore analisi di ricombinazione. Tale analisi, condotta mediante programma *Simplot*, ha consentito di evidenziare la presenza di 3 sequenze ricombinanti che mostravano un pattern A/G in 1 caso, A/D nel secondo e D/A nel terzo (Figura 17A-B-C).



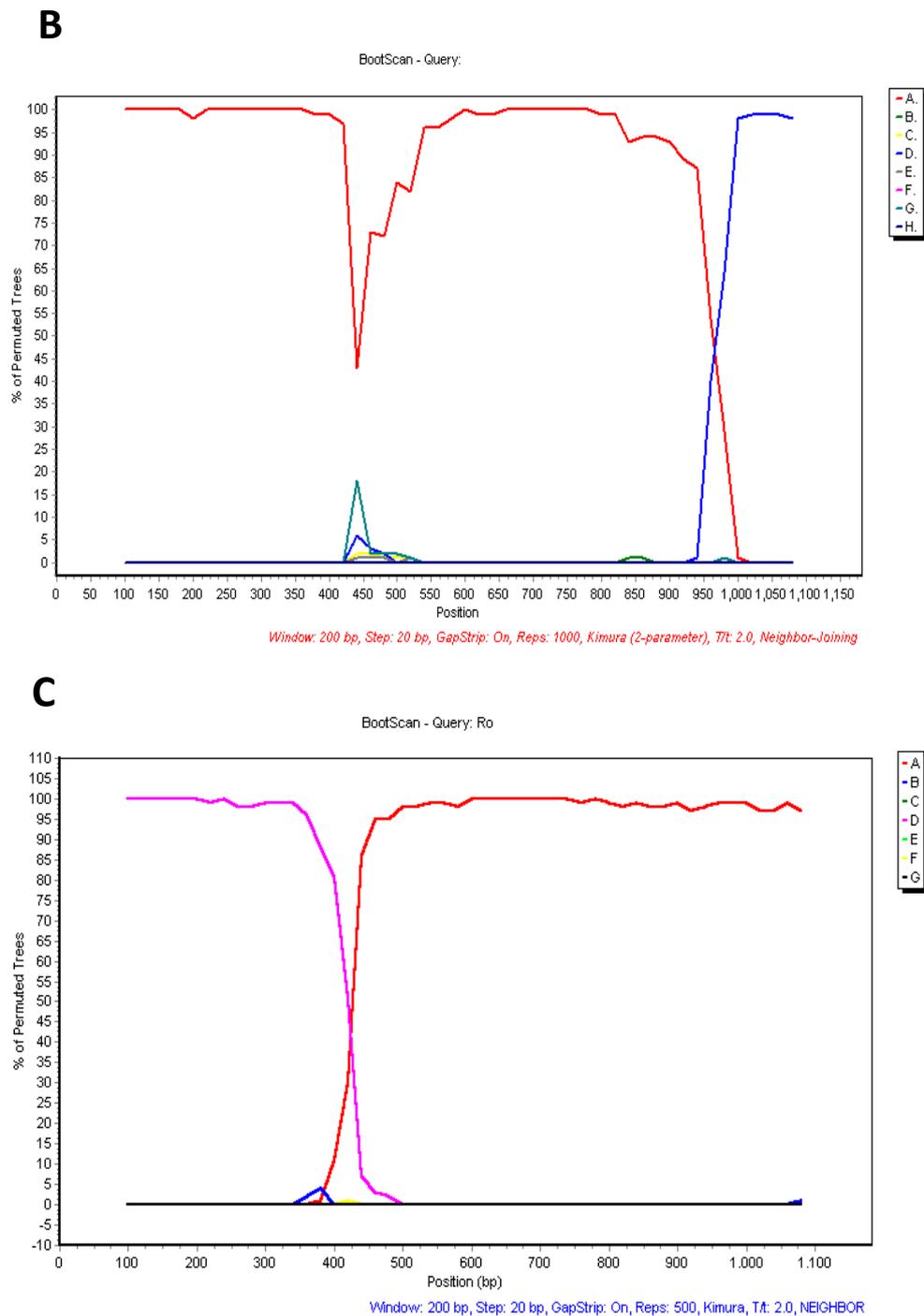


Figura 17. Analisi di bootscanning delle sequenze ricombinanti. A) ricombinante con pattern A/G; B) ricombinante con pattern A/D; C) ricombinante con pattern D/A.

4.2.1 Distribuzione dei sottogenotipi di HBV nella popolazione in studio

È stato possibile caratterizzare il sottogenotipo di HBV in 227 pazienti dei 230 totali, dovendo escludere i 3 ceppi ricombinanti. Dei 166 isolati di genotipo D, 16 (9.6%) sono risultati di sottogenotipo D1, 15 (9%) D2, 133 (80.1%) D3 e 2 (1.3%) D4. Dei 43 soggetti con infezione da HBV di genotipo A all'analisi filogenetica, 3 (7%) mostravano il sottogenotipo A1 e 40 (93%) il sottogenotipo A2. Per quanto riguarda gli 8 ceppi di genotipo F, 3 (37.5%) appartenevano al sottogenotipo F1 e 5 (62.5%) al sottogenotipo F2 (Figura 18).

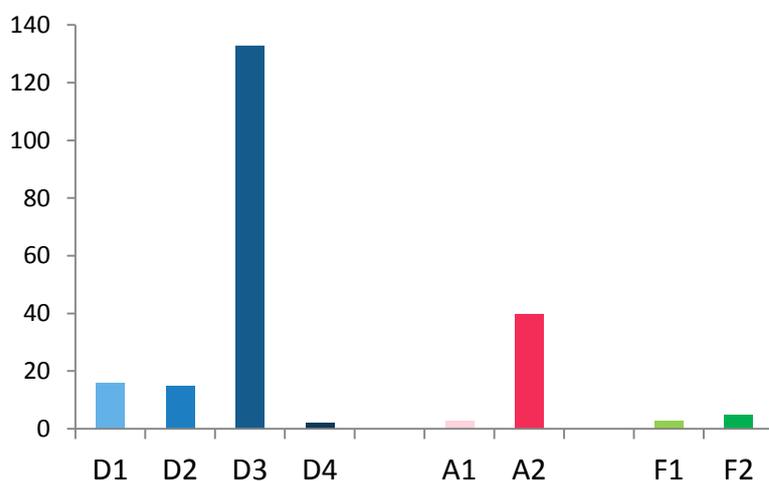


Figura 18. Distribuzione dei sottogenotipi di HBV nella popolazione in studio.

4.2.2 Associazione tra fattore di rischio e genotipo/sottogenotipo virale

Nei pazienti con coinfezione HBV/HIV-1, il genotipo D è risultato essere il più diffuso tra i soggetti che facevano uso di droghe per via iniettiva rispetto a quelli che non mostravano tale fattore di rischio (prevalenza nei tossicodipendenti del 69.2%, $p=0.04$). Il genotipo A risulta invece maggiormente associato ai soggetti MSM rispetto a coloro che non riportavano l'omosessualità come fattore di rischio (66.7%, $p=0.016$). Dei 4 soggetti infetti con HBV di genotipo G, 2 (50%) riferivano

comportamenti omosessuali, 1 (25%) comportamenti eterosessuali a rischio ed 1 (25%) era un soggetto tossicodipendente (Figura 19).

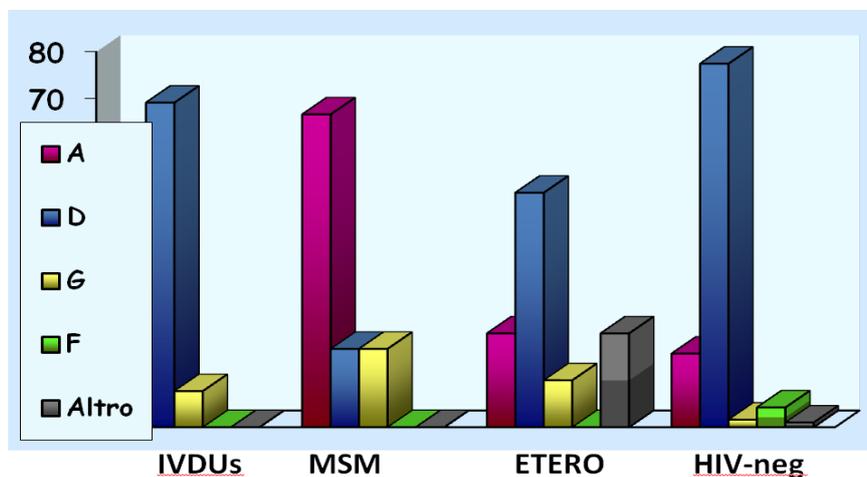


Figura 19. Prevalenza dei genotipi più rappresentati di HBV (A, D, G, F) rispetto al fattore di rischio nei soggetti HIV-1 positivi e HIV-1 negativi.

Tra i pazienti con infezione singola da HBV il fattore di esposizione era noto, come già detto, solo in una minoranza dei casi.

Il sottogenotipo D3 risulta essere prevalente nei tossicodipendenti (69.2%), essendo virtualmente l'unico sottogenotipo D identificato nei genotipizzati. Il sottogenotipo A2 è il più frequente tra gli MSM (66.7%) (Figura 20).

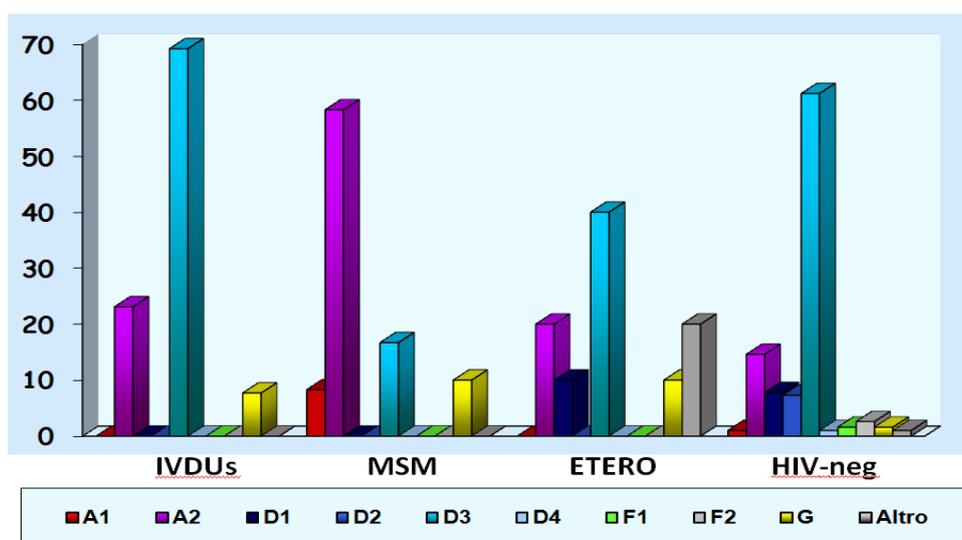


Figura 20. Prevalenza dei sottogenotipi di HBV rispetto al fattore di rischio nei soggetti HIV-1 positivi e HIV-1 negativi.

4.3 ANALISI DEL GENOTIPO D

4.3.1 Creazione dei dataset

Al fine di ricostruire la filodinamica di HBV-D in Italia e la filogeografia di questo genotipo a livello globale abbiamo costruito due dataset:

1) il primo dataset comprende 79 sequenze di sottogenotipo D3, che rappresenta il sottogenotipo di HBV più frequente in Italia, utilizzato per delineare la storia epidemiologica/evolutiva del virus HBV nel nostro Paese;

2) il secondo dataset, utilizzato per la ricostruzione filogeografica dei sottogenotipi D di HBV, è stato costruito aggiungendo alle sequenze di HBV-D caratterizzate nel presente lavoro, sequenze dello stesso genotipo, recuperate dalle principali banche dati pubbliche e isolate in diversi Paesi, ottenendo in questo modo un dataset complessivo di 312 isolati. Di tutte le sequenze, era noto il luogo di isolamento, mentre l'anno di isolamento era noto solo per una parte di essi. Per questa ragione, al fine di ottenere una filogenesi in scala spazio-temporale, abbiamo applicato un valore di ER, da noi stimato in precedenza su un dataset indipendente, (4.4×10^{-4} sostituzioni/sito/anno)²²³ alla ricostruzione dell'albero filogeografico.

Gli accession number e i luoghi di campionamento delle sequenze di riferimento utilizzate per la costruzione del dataset sono riassunti in Appendice 1.

4.3.2 Likelihood mapping

Il metodo di Likelihood mapping è stato applicato ai due dataset di genotipo D in studio.

Valutando i 10,000 quartetti random, questa analisi ha mostrato che la porzione di quartetti random che rientrano nella zona centrale della mappa è del 17,3% per il dataset D3 (Figura 21, Pannello A) e del 2,9 % per il dataset filogeografico contenete le sequenze di tutti i sottogenotipi D (Figura 21, Pannello B), mentre quelli che si distribuiscono negli angoli sono il 72,1% e 92,7% rispettivamente.

Questo indica che i due allineamenti contengono sufficiente informazione filogenetica e sono quindi adatti alla ricostruzione filogenetica.

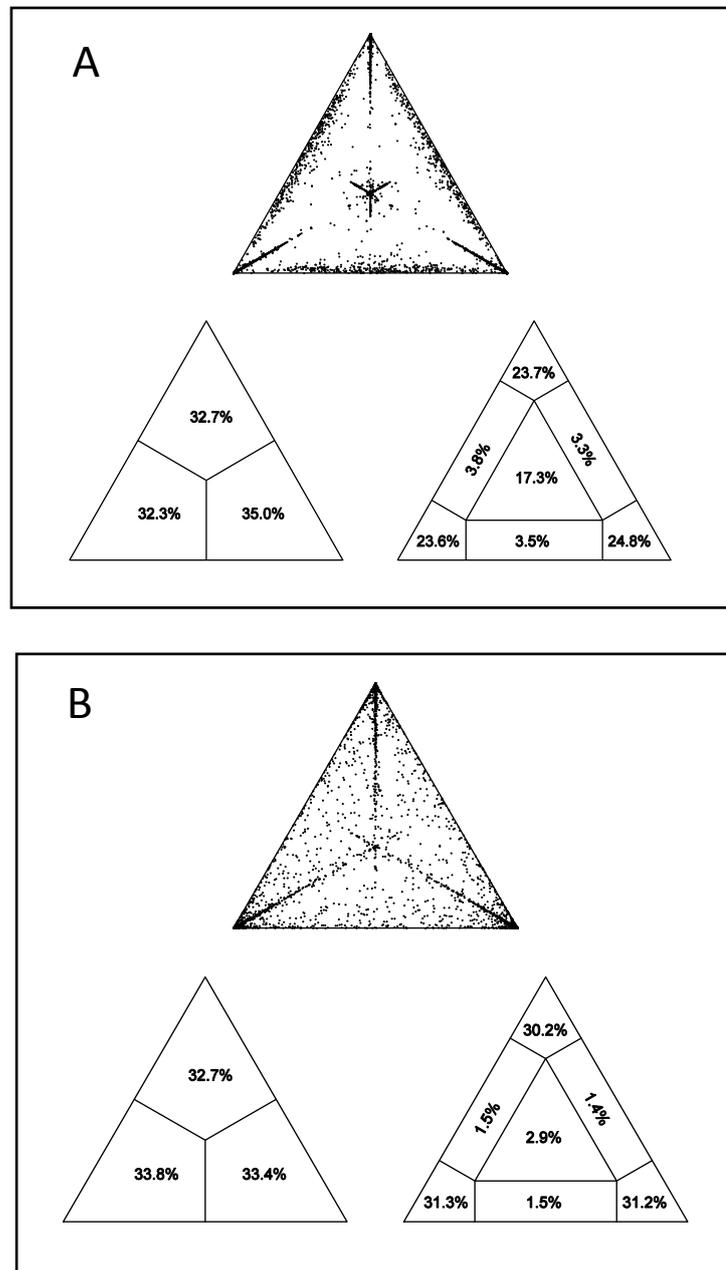


Figura 21. Likelihood mapping applicato ai dataset in studio, contenenti sequenze di HBV-D3 (Pannello A) e sequenze di HBV di tutti i sottogenotipi D (Pannello B). I punti all'interno del triangolo rappresentano le probabilità posteriori delle possibili tipologie di alberi senza radice per ogni quartetto. I numeri indicano le percentuali dei punti nelle diverse aree del triangolo.

4.3.3 Analisi filodinamica

Tutti gli isolati HBV-D3 italiani sono stati inclusi in un sotto-dataset per condurre l'analisi MCMC della dinamica di popolazione, al fine di delineare la storia epidemiologica/evolutiva del virus in Italia.

Come riportato nella sezione Materiali e Metodi, sono stati confrontati due modelli demografici parametrici (*constant size, exponential growth*) e uno non parametrico, (*Bayesian skyline*), sulla base del *Bayes Factor* ottenuto.

La comparazione dei modelli di coalescenza, effettuata sulla base del *Bayes Factor* (BF), mostra che sia il modello non-parametrico BSP ($2\ln\text{BF}=83.4$ a favore dello skyline) che quello di crescita esponenziale ($2\ln\text{BF}=36.3$) sono egualmente adatte ai dati e migliori rispetto al modello di crescita di popolazione costante.

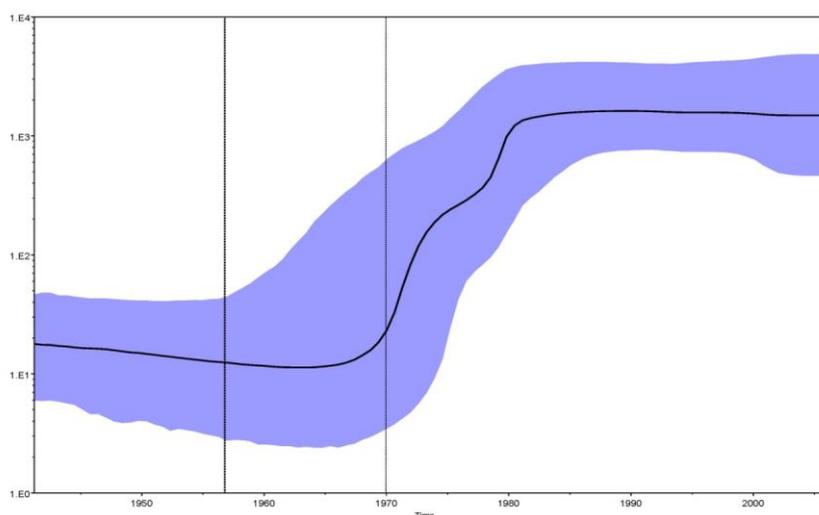


Figura 22. Skyline plot bayesiano rappresentante la stima del numero effettivo di infezioni da HBV-D3 (asse y: scala \log_{10}) nel tempo (asse x: anni). La linea continua rappresenta la mediana e l'area colorata l'intervallo di credibilità.

L'analisi dello skyline plot condotta sul dataset di sequenze italiane D3 (Figura 22) suggerisce che il sottogenotipo D3 sia presente in Italia dagli anni '40 e sia rimasto ad un livello costante di endemia fino agli anni '60. In seguito si è assistito ad una crescita del numero effettivo di infezioni in due fasi: la prima si sarebbe verificata intorno alla fine degli anni '60, con una riduzione a metà degli anni '70, e la seconda si sarebbe verificata intorno agli anni '80.

L'impiego del modello parametrico ha permesso la stima del *tasso di crescita esponenziale* ($r=0.138$; 95% HPD 0.039-0.228), sulla cui base, abbiamo potuto calcolare alcuni parametri epidemiologici fondamentali per la descrizione dell'epidemia verificatasi in Italia (Tabella 4). In particolare, abbiamo stimato il *numero riproduttivo di base* (R_0), variabile da 1.69 a 5.14, in relazione alla diversa possibile durata dell'infettività (rispettivamente 5, 10, 15, 20 o 30 anni). Inoltre, è stato possibile calcolare il *tempo di raddoppiamento dell'epidemia* ($\lambda=\ln(2)/r$) in 5 anni (intervallo di credibilità 17.8-3).

Tabella 4. Parametri epidemiologici dell'epidemia sostenuta di HBV-D3 in Italia.

	Medio	Lower	Upper
R	0.138	0.039	0.228
R₀: D=5	1.69	1.195	2.14
D=10	2.38	1.39	3.28
D=15	3.07	1.58	4.42
D=20	3.76	1.78	5.56
D=30	5.14	2.17	7.84
λ	5	17.8	3

Tra i 79 isolati D3, sono presenti 29 sequenze, provenienti da soggetti IVDUs, caratterizzate da una serie di mutazioni, di cui la più nota è la sostituzione a livello del gene S T125M, descritta da più Autori in soggetti IVDUs di tutto il mondo.

Sono state quindi effettuate due analisi indipendenti: una in cui dal dataset completo di sequenze italiane di sottogenotipo D3 sono state escluse le sequenze con la mutazione T125M del ceppo associato alla tossicodipendenza, l'altra solo su quest'ultime.

Lo skyline condotto sul dataset di isolati italiani privo di sequenze di ceppo IVDUs (Figura 23) mostra un livello costante di infezioni fino agli anni '50. Successivamente si evidenzia una crescita del numero effettivo di infezioni fino agli anni '70, quando la curva raggiunge il plateau.

L'analisi dello skyline delle sole sequenze IVDUs (Figura 24) mostra invece una crescita esponenziale molto tardiva (primi anni '70) che si conclude nei primi anni '80 quando la curva raggiunge il plateau.

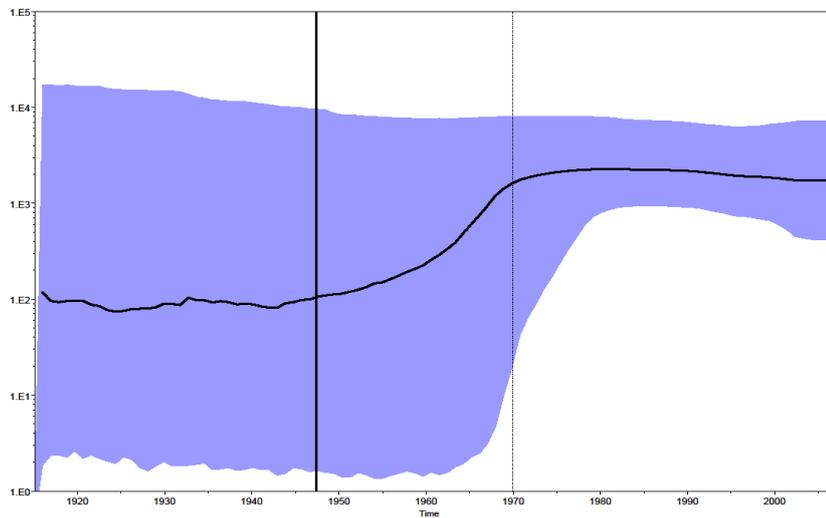


Figura 23. Skyline plot realizzato sulle sequenze D3 italiane escludendo dal dataset gli isolati T125M. In ordinata: numero delle infezioni effettive nel tempo ($Ne(t)$), in ascissa: tempo (in anni).

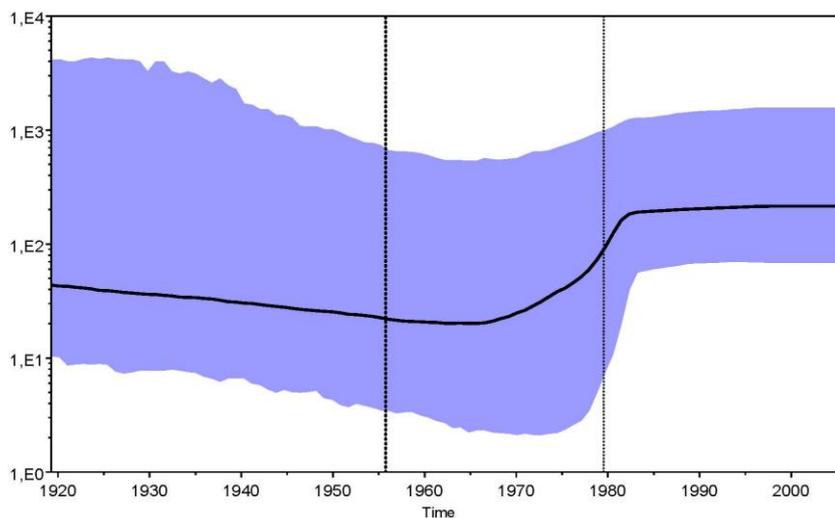


Figura 24. Skyline plot realizzato sulle sequenze D3 T125M del ceppo IVDUs. In ordinata: numero delle infezioni effettive nel tempo ($Ne(t)$), in ascissa: tempo (in anni).

4.3.4 Analisi filogeografica

Nel corso dell'analisi filogeografica del genotipo D, le 312 sequenze sono state classificate, dal punto di vista spaziale, in gruppi corrispondenti alle località: Albania (AL): n=59; Asia Centrale (AC): n=10; Egitto/Israele (EI): n=5; Estremo Oriente (EO): n=8; Grecia (GR): n=4; India (IN): n=26; Iran(IR): n=13; Italia (IT): n=63; Russia (RU): n=53; Serbia (RS): n=3; Sud Africa (SA): n=6; Spagna (SP): n=4; Tunisia (TN): n=41; Turchia (TR): n=17.

L'analisi dell'albero filogeografico (Figura 25) mostra che i 312 isolati di genotipo D segregano in *clade* significativi, corrispondenti ai genotipi D maggiormente presenti nelle aree geografiche considerate.

In particolare, HBV-D5 e D7 formano due gruppi altamente significativi ($pp=1$) che divergono dalla radice dell'albero prima degli altri sottogenotipi. Gli isolati appartenenti ai tre sottogenotipi maggiori (D1, D2 e D3) formano tre *clade* ($pp=0.52$ per D1 e $pp=1$ per D2 e D3) e condividono due nodi significativi: il primo comune a tutti e tre i sottogenotipi ($pp=0.89$) e il secondo comune ai sottogenotipi D1 e D2 ($pp=0.88$).

Sulla base della nostra ricostruzione filogeografica, la località più probabile della radice dell'albero è risultata l'India (probabilità a posteriori, $pp=0.30$) seguita dall'Asia Centrale ($pp=0.22$). Gli isolati provenienti dal *clade* D5 comprendono ceppi di origine indiana, mentre quelli del *clade* D7 hanno un'origine tunisina. Gli altri *clade* includono isolati provenienti da diverse parti del mondo. All'interno del *clade* D1 sono presenti vari isolati provenienti da India, Asia Centrale (Kazakhstan e Uzbekistan), Cina, Vietnam, Tunisia, Iran, Egitto, Europa Orientale (Turchia e Bulgaria) e Italia. La località più probabile del MRCA del sottogenotipo D1 è risultata la Turchia ($pp=0.40$).

Il *clade* D3 include numerosi isolati italiani, qualche ceppo albanese e del nord-est europeo, 3 isolati provenienti dalla Serbia, 4 dall'India, un isolato proveniente dalla Mongolia e un gruppo di 6 isolati sudafricani che segregano significativamente dagli altri ceppi. L'Asia Centrale ($pp=0.54$) è la località più probabile del MRCA per il sottogenotipo D3. Anche gli isolati D3 segregano in modo significativo in base all'origine geografica.

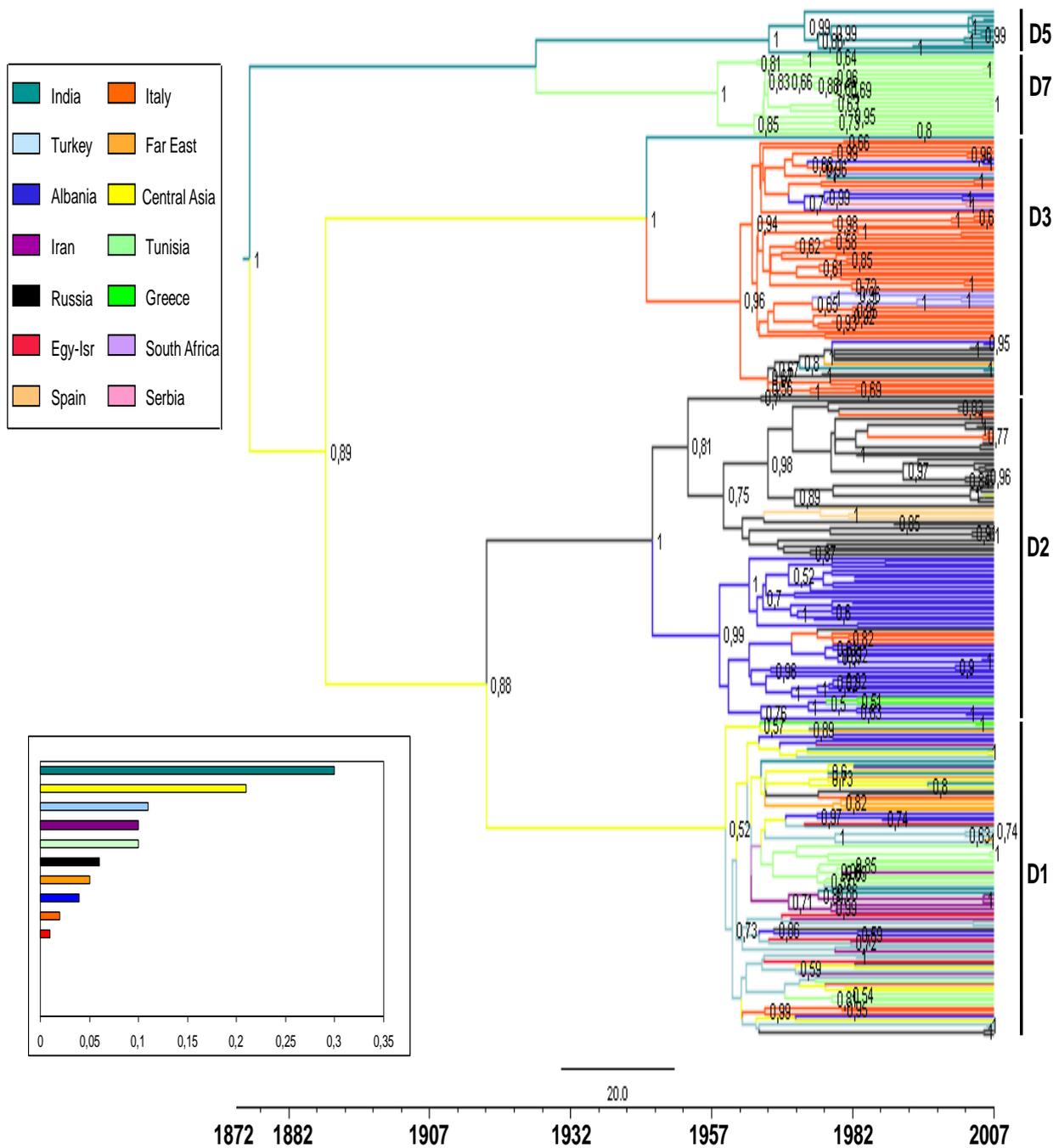


Figura 25. Albero datato delle 312 sequenze del gene P di HBV di genotipo D, di differente provenienza geografica. La scala ai piedi dell'albero rappresenta il tempo (in anni) prima dell'ultimo anno di campionamento (2007). Nella legenda sono indicati i Paesi e i rispettivi colori. Il grafico in alto rappresenta le probabilità posteriori delle diverse località della radice.

Il *clade* corrispondente al sottogenotipo D2 contiene in gran parte ceppi provenienti da Russia, Bielorussia ed Estonia e un ampio gruppo contenente isolati albanesi, qualche ceppo italiano e pochi isolati greci e turchi. Le località più probabili del MRCA per il sottogenotipo D2 sono l'Asia Centrale (pp =0.33) e la Russia (pp =0.30). Tutti gli isolati D2 tendono a segregare in modo significativo in base all'origine geografica. In particolare, sono presenti 2 *cluster* significativi: il gruppo D2A (pp =0.99), il cui MRCA ha un'origine più probabile in Albania (pp =0.96), che include tutti i ceppi albanesi e qualche ceppo italiano; il gruppo D2B (pp =0.81), il cui MRCA ha un'origine più probabile in Russia, che include tutti i ceppi provenienti dal Nord-est Europeo e qualche ceppo italiano.

Nella Tabella 5 sono riportati i tMRCA stimati per i principali *clade* descritti.

Tabella 5. tMRCA, date e località di origine dei principali sottogenotipi e clades italiani e stranieri di genotipo D.

Nodo	tMRCA	Anni	Località	PP
Radice	128	1880	India	0.3
D1-D3	86	1922	Asia Centrale	0.33
D1-D2	68	1940	Asia Centrale	0.43
D1	48	1960	Turchia	0.40
D2	52	1956	Asia Cen.+ Russia	0.63
D2A	45	1963	Albania	0.96
D2B	43	1965	Russia	0.96
D3	65	1943	Asia Cen.+India	0.56
D5	40	1968	India	0.95
D7	52	1956	Tunisia	0.99

La stima media dei tMRCA della radice dell'albero è di 128 anni fa (intervallo di credibilità del 95% HPD, tra 66 e 210 anni fa), suggerendo che l'origine dei sottogenotipi D risalga al 1880 (95% HPD, tra 1798 e 1942). I sottogenotipi D5 e D7 sono stati i primi a divergere dalla radice dell'albero, anche se la loro diffusione è relativamente recente (40 e 52 anni fa, corrispondenti agli anni 1968 e 1956). Sulla base delle nostre stime, il progenitore comune ai sottogenotipi D1, D2 e D3 è esistito circa 86 anni fa (circa 1922), mentre il progenitore comune dei

sottogenotipi D1e D2 è esistito circa 68 anni fa (1940). Anche in questo caso i tre sottogenotipi si sono diffusi in tempi recenti: 65 anni fa il sottogenotipo D3 (1943), 52 anni fa il D2 (1956) e 48 anni fa il D1(1960). Sulla base del tMRCA stimato l'origine dei due *cluster* D2 era 1963 per il D2A (95% HPD, tra 1943 e 1981) e il 1965 per D2B (95% HPD, tra 1945 e 1981).



Figura 26. *Flussi significativi di migrazione nello spazio del genotipo D. Sono mostrati i flussi supportati da un Bayes Factor (BF) >6. La ricostruzione spaziale è stata effettuata tramite il programma SPREAD. La potenza del supporto statistico è indicata dal colore delle linee (da rosso scuro = debole, a rosso chiaro = forte). La mappa è stata prodotta utilizzando un'immagine satellitare ottenuta con Google Earth.*

L'analisi filogeografica bayesiana ha mostrato diversi flussi significativi, supportati da un Bayes factor (BF) >6, tra le diverse località (Figura 26). Sulla base di tale analisi possiamo affermare che il genotipo D, dopo essersi originato in India nel 1872, si è diffuso in Asia Centrale nel 1915. È interessante notare che l'India risulta collegata solo con l'Asia Centrale (BF = 27,3), mentre l'Asia Centrale è collegata a cinque diverse località geografiche (Russia, Estremo Oriente, Turchia, Iran e India), suggerendo un ruolo centrale nella diffusione globale di HBV- D.

Altre località con diversi flussi significativi sono la Russia (con Italia, Spagna e Asia Centrale), la Turchia (con Asia Centrale, Iran ed Egitto/Israele) e l'Italia (con Russia, Albania e Sud Africa).

4.4 ANALISI DEL GENOTIPO A

4.4.1 Creazione dei dataset

38 delle 43 sequenze del gene P di HBV-A da noi caratterizzate, di cui era disponibile l'intera lunghezza di 1182 nt, sono state allineate con sequenze di riferimento rappresentative di tutti i sottogenotipi di HBV-A precedentemente descritti (A1, A2, A3, "tentative A4", A5, A6, "tentative A7"). Quest'ultime comprendono: 2 sequenze, entrambe di sottogenotipo A1, ottenute da campioni provenienti da Pemba (Zanzibar) caratterizzati precedentemente nel nostro laboratorio, 73 sequenze provenienti da isolati caratterizzati in collaborazione con il Dipartimento di Medicina sperimentale e Chirurgia dell'Università di Roma "Tor Vergata" e 247 sequenze recuperate dalle principali banche dati pubbliche (<http://www.ncbi.nlm.nih.gov/genbank/>). Per i soggetti italiani inclusi nello studio della filodinamica del genotipo A le caratteristiche anagrafiche e il sottogenotipo virale sono riassunte in Tabella 6.

In particolare le sequenze sono state raggruppate in 4 sottogenotipi: A1, A2, A3_{qs} (qs = quasi sottogenotipo, comprendente l'ex A3, "tentative A4", A5, "tentative A7") e A6.

Sono stati costruiti due dataset distinti: il primo (dataset A2) composto da 165 sequenze (36 delle quali provenienti dai nostri isolati) di HBV di sottogenotipo A2; il secondo (dataset A1/A3_{qs}/A6) contenente un totale di 195 isolati (2 dei quali provenienti dai nostri campioni) di tutti i sottogenotipi A, ad eccezione dell'A2.

Di tutte le sequenze era noto il luogo di isolamento, mentre l'anno di isolamento era noto solo per una parte di essi. Gli accession number e i luoghi di

campionamento delle sequenze di riferimento utilizzate per la costruzione dei dataset sono riassunti in Appendice 2.

Tabella 6. *Caratteristiche anagrafiche e sottogenotipo di HBV dei soggetti italiani infetti inclusi nello studio della filodinamica del genotipo A.*

		n soggetti (%) Roma "Tor Vergata"	n soggetti (%) Milano
Genere	M	59 (80.8%)	31 (81.6%)
Nazionalità	Italiana	51 (79.7%)	34 (89.5%)
	Straniera	13(20.3%)	4 (10.5%)
Coinfezione	HIV	25 (34.2%)	13 (34.2%)
	HCV	2 (2.7%)	-
Sottogenotipo	A1	19 (26%)	2 (5.3%)
	A2	53 (72,6%)	36 (94.7%)
	A3	1 (1.4%)	-

4.4.2 Likelihood mapping

Il metodo di Likelihood mapping è stato applicato ai due dataset di genotipo A in studio. Valutando i 10,000 quartetti random, questa analisi ha mostrato che la porzione di quartetti random che rientrano nella zona centrale della mappa è del 29 % per il dataset A2 (Figura 27, Pannello A), e del 6,3% per il dataset A1/A3_{QS}/A6 (Figura 27, Pannello B), mentre quelli che si distribuiscono negli angoli sono il 67% e 87,5% rispettivamente. Questo indica che i due allineamenti contengono sufficiente informazione filogenetica e sono quindi adatti alla ricostruzione filogenetica.

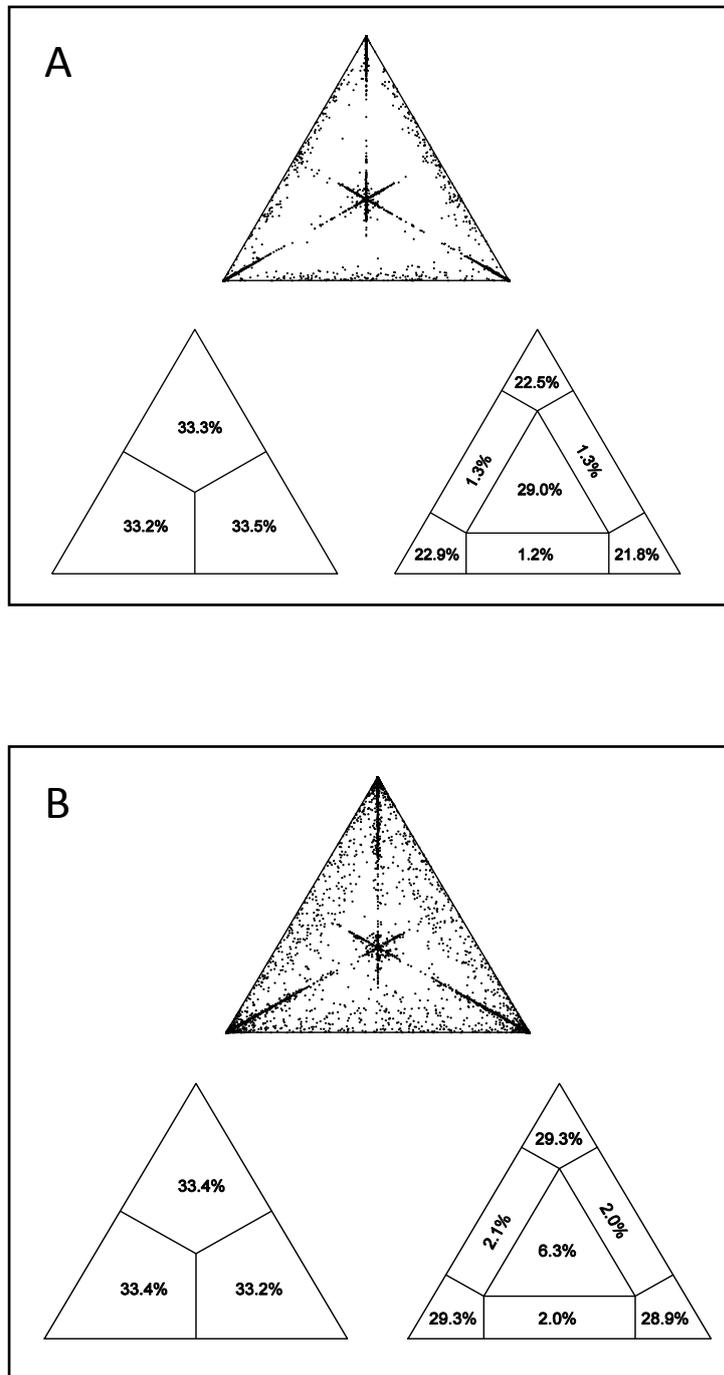


Figura 27. Likelihood mapping applicato ai dataset in studio contenenti sequenze di HBV-A2 (Pannello A) e sequenze di HBV di sottogenotipo A1/A3_{qs}/A6 (Pannello B).

4.4.3 Stima dell'evolutionary rate dei sottogenotipi di HBV-A

Gli evolutionary rate (ER) delle sequenze del gene P di genotipo A sono stati stimati separatamente per i due dataset (A2 e A1/A3_{QS}/A6) mediante il test BF, comparando due modelli di orologio molecolare, lo “*strict*” e il “*relaxed*”.

Considerate le diverse scale temporali e i diversi evolutionary rate tra HBV-A2 e gli altri sottogenotipi A, sono stati utilizzati due diversi approcci di calibrazione per i due dataset in studio. L'approccio utilizzato per il dataset A1/A3_{QS}/A6 è basato su dati storici/antropologici e consiste nella calibrazione dell'età del clade haitiano A5 con il periodo in cui il commercio degli schiavi ad Haiti ha raggiunto il picco massimo (circa 224-316 anni fa). La calibrazione del dataset HBV-A2 è invece basata sulla data di campionamento degli isolati. In entrambe i casi è stato selezionato un modello di crescita di popolazione complesso (BSP) con orologio “rilassato”. I valori di $2\ln BF$ risultanti da ciascun confronto eseguito sono stati riportati nella Tabella 7.

Tabella 7. Valori di $2\ln BF$ risultanti dal confronto tra diversi orologi molecolari e modelli demografici mediante il test BF.

Prior	Confronti	HBV-A2	HBV-A1/A3 _{QS} /A6
Modelli di orologio molecolare	Strict vs relaxed	97.8	139.8
	Constant vs BSP	334.2	229.7
Modelli di coalescenza	Constant vs Exponential	278	130.8
	Exponential vs BSP	56.3	98.9

L'utilizzo di una calibrazione esterna per il clade haitiano A5 ha determinato una stima di evolutionary rate per il dataset A1/A3_{QS}/A6 di 3.0×10^{-5} sostituzioni/sito/anno, con un intervallo di credibilità compreso tra $1,9 \times 10^{-5}$ e 4.4×10^{-5} sostituzioni/sito/anno. Diversamente, la stima di ER di HBV-A2 determinata mediante sequenze eterocrone, ha generato un valore medio di 1.1

$\times 10^{-4}$ sostituzioni/sito/anno con un intervallo di credibilità compreso tra 3.7×10^{-5} e 1.8×10^{-4} sostituzioni/sito/anno.

4.4.4 Analisi filogenetica e filodinamica del sottogenotipo A2

La Figura 28 mostra l'albero datato MCC (*maximum clade credibility*) costruito sulle 165 sequenze HBV-A2. Le sequenze sono mescolate tra di loro, indipendentemente dalla località di campionamento. Solo gli isolati italiani e rumeni tendono a segregare dalle altre sequenze formando un ampio clade significativo (*posterior probability*=0,91) nell'albero bayesiano ottenuto con il programma MrBayes (dato non mostrato).

Un totale di 8 sottoclades altamente significativi ($pp > 0.9$), comprendenti da 3 a 8 isolati, sono stati identificati. Sei di questi, incluso il clade più numeroso, sono italiani, mentre i due rimanenti, entrambi costituiti da 3 isolati, sono uno di origine rumena e l'altro contiene isolati provenienti dalla Germania. Tra i sottogruppi italiani, uno è costituito da tre soggetti HIV-1 positivi e il più numeroso (8 isolati) comprendeva 4 sequenze di soggetti HIV-1 positivi.

L'analisi dei tMRCA ha mostrato che la radice dell'albero è datata 110 anni fa (con un intervallo di credibilità tra 24 e 188 anni fa), il che suggerisce che il progenitore comune ai ceppi A2 attualmente in circolazione esisteva nei primi decenni del 1900. L'età media dei sottogruppi italiani varia da circa 56 a 15 anni fa (1956-1998).

In particolare, i due sottogruppi comprendenti pazienti HIV-1 positivi presentano stime di tMRCA di circa 35 anni. Il sottogruppo comprendente le donne rumene, presenta un tMRCA di circa 40 anni suggerendo un'origine che risale agli anni '70 (Tabella 8).

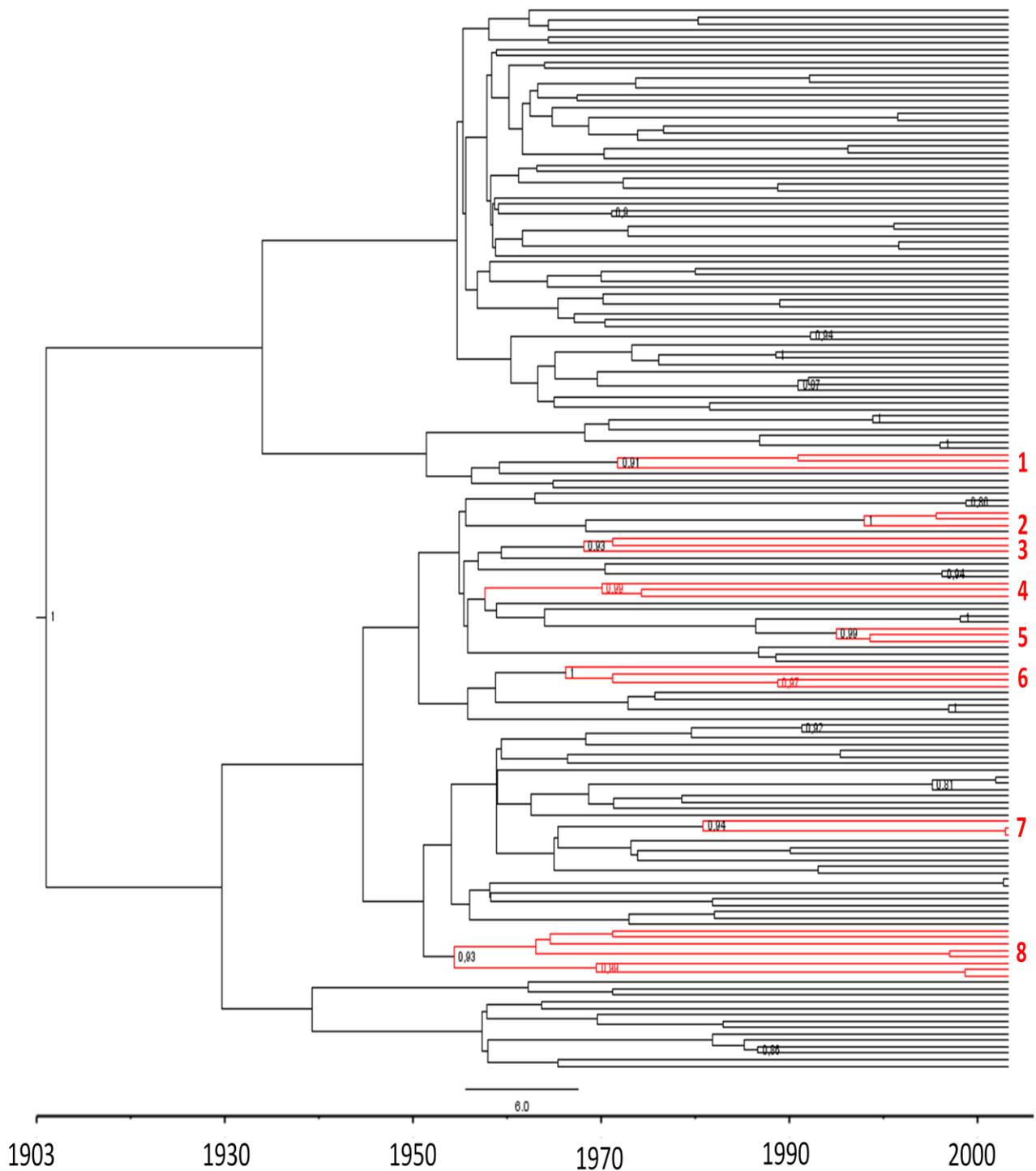


Figura 28. Albero bayesiano MCC (*maximum clade credibility*) costruito sulle 165 sequenze del gene *P* di sottogenotipo A2. La scala ai piedi dell'albero rappresenta il tempo in anni. I numeri indicati sui nodi rappresentano le posterior probability. I clade significativi #1-8 (posterior probabilità > 0.9) descritti nel testo, sono evidenziati in rosso.

Tabella 8. *tMRCA*, intervalli di credibilità (95%HPD), date di origine, n° di isolati e caratteristiche dei clade significativi ($pp > 0.9$) di sottogenotipo A2.

Nodo	tMRCA	Min	Max	Anno	Min	Max	N°isolati	Catteristiche
Root	110	24	188	1903	1989	1825		
#1	34	19	75	1979	1994	1938	3	DE
#2	21.9	0.4	45	1991.1	2012.6	1968	3	
#3	36.6	0.1	51	1976.4	2012.9	1962	3	HIV+
#4	39.8	5	64	1973.2	2008	1949	3	Ro
#5	29.5	0.1	48	1983.5	2012.9	1965	3	
#6	45.5	9	73	1967.5	2004	1940	4	
#7	14.9	0.1	42	1998.1	2012.9	1971	3	
#8	56.6	12.7	81	1956.4	2000.3	1932	8	
#8bis	34	0.02	42	1979	2012.98	1971	3	HIV+

Lo skyline plot ottenuto sul dataset di isolati HBV-A2 (Figura 29), mostra che il numero effettivo di infezioni dovute a tale sottogenotipo è rimasto costante fino al 1970 quando ha cominciato a crescere e ha raggiunto il plateau a metà degli anni '90.

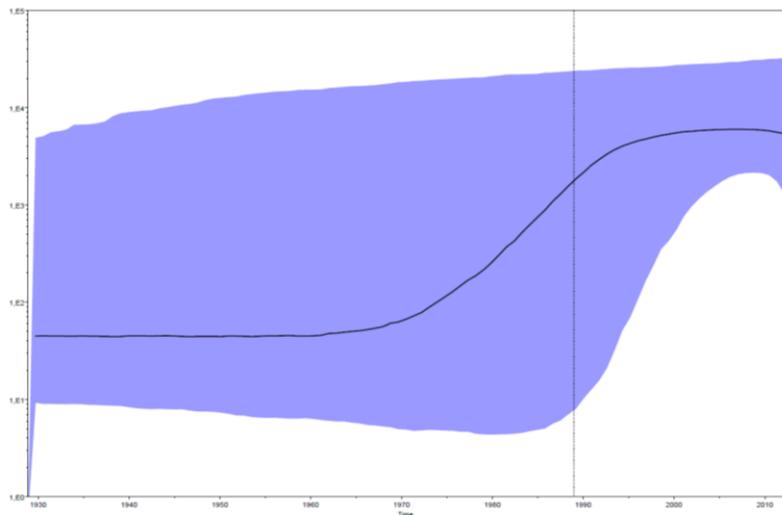


Figura 29. Skyline plot realizzato sul data set di isolati A2. In ordinata: numero delle infezioni effettive nel tempo ($N_e(t)$), in ascissa: tempo (in anni).

4.4.5 Analisi filogenetica e filogeografica dei sottogenotipi A1, A3_{QS} e A6

Nel corso dell'analisi filogeografica del genotipo A, le sequenze sono state classificate dal punto di vista spaziale in gruppi corrispondenti alle località: Africa Centro-occidentale (WAF): n=35; Africa Orientale (EAF): n=20; America Latina (LA): n=15; Asia (AS): n=31; Haiti (HA): n=52; Italia (IT): n=21; Sud Africa (SA): n=21.

La Figura 30 mostra l'albero datato MCC, costruito sul dataset contenente le sequenze di HBV-A di tutti i sottogenotipi, ad eccezione dell'A2, e le località più probabili di ciascun nodo interno in diversi colori.

Tre clade altamente significativi sono presenti: il maggiore (p=0.91) corrispondente al sottogenotipo A1, il secondo (p=0.97), comprendente sequenze A3_{QS} e il terzo (p=1), quello meno numeroso, composto da 4 isolati di pazienti di origine belga-africana classificati come HBV-A6.

All'interno del clade A3_{QS} si distinguono diversi gruppi monofiletici: il clade altamente significativo (pp=1) comprendente 20 isolati A5 provenienti da Haiti (impiegato per la calibrazione) e separato dalle altre sequenze A5 Africane (pp=0.98), il clade "tentative A4" (pp=0.99) e il clade "tentative A7" (pp=0.81). Il *cluster* prima identificato come sottogenotipo A3 si suddivide in 2 gruppi con *posterior probability* di 0.63 e 0.73. Tutte le sequenze incluse nel clade A3_{QS} derivano da isolati ottenuti in diversi paesi dell'Africa Centro-occidentale (Camerun, Nigeria, Gabon, Guinea, Mali, Gambia), con l'eccezione degli isolati Haitiani.

Il clade corrispondente al sottogenotipo A1, si suddivide in due grandi gruppi supportati da *posterior probability* elevate. Il primo (pp=0.78), che abbiamo chiamato "Africano", comprende solo sequenze campionate in questo continente (soprattutto nei paesi dell'Africa Sud-orientale), formando diversi sotto-clade significativi e caratterizzati da una forte struttura geografica. I 2 isolati provenienti dall'Isola di Pemba sono inclusi in questo gruppo e a loro volta formano un sottogruppo altamente significativo (pp=1) con gli isolati provenienti da Malawi e Tanzania. Il secondo grande gruppo A1 (pp=1), soprannominato "Cosmopolita",

Gli isolati haitiani appartenenti al sottogenotipo A1 sono suddivisi in due clade principali: uno più numeroso ($pp=0.87$) costituito da 22 sequenze, che ha come outgroup un isolato proveniente dalla Somalia, e un secondo più piccolo ($pp=0.99$) di sole 7 sequenze.

Gli isolati provenienti dal Sud America, in particolare Colombia e Brasile, formano un clade monofiletico significativo ($pp=0.7$) ulteriormente suddiviso in due gruppi: il primo gruppo ($pp=0.98$) costituito da isolati Colombiani ed Argentini e una sequenza Haitiana, il secondo ($pp=0.93$) costituito da sequenze brasiliane.

Tutti gli isolati italiani segregano in un unico clade altamente significativo ($pp=1$). La maggior parte dei ceppi asiatici tendono a clusterizzare in un unico gruppo supportato da una bassa *posterior probability* ($pp=0.52$).

In Tabella 9 sono riportate le stime dei tMRCA e le date corrispondenti dei principali clade descritti. Il tMRCA medio della radice dell'albero è di 1057 anni fa (95% HPD=530-1631) suggerendo un'origine dei sottogenotipi A attualmente circolanti nel periodo dell'Alto Medioevo (data media 955 d.C.).

Le stime medie dei tMRCA dei principali tre clade A1, A3_{QS} e A6 sono di circa 666, 640 e 296 anni fa rispettivamente. In particolare l'A1 e l'A6 condividono un nodo datato 818 anni fa (1275-403).

Sulla base delle nostre stime, i clade "Cosmopolita" ed "Africano" di sottogenotipo A1 si sono separati nel tardo 1400; il sottogruppo asiatico si è originato in media nel 1648 (95% HPD=1506-1822), mentre il gruppo monofiletico europeo (Italia) solo più tardi, nel 1710 (95% HPD=1606-1859).

I tre clade haitiani (due A1 e uno A5) hanno stime medie di tMRCA molto simili tra loro che variano tra i 254 e i 294 anni fa, suggerendo così diversi eventi di penetrazione del virus ad Haiti tra la fine del XVII e l'inizio del XVIII secolo. Anche i due clade dell'America Latina hanno mostrato stime di tMRCA analoghe a quelle dei clade haitiani (273 e 239 anni fa), suggerendo molteplici ingressi di HBV-A1 in Sud America nel corso del XVIII secolo.

L'analisi filogeografica (Tabella 9 e Figura 30) ha mostrato che la località più probabile della radice dell'albero è l'Africa Centro-occidentale (location spp = 0.66 vs 0.29 per il Sud Africa), che è anche la località più probabile per i clade A3_{QS} e A6.

Al contrario, la località più probabile per il clade A1 (spp=0.81) e i suoi principali sottogruppi “Africano” e “Cosmopolita” (spp=0.8 e 0.7 rispettivamente) è risultata l’Africa Orientale.

Tabella 9. *tMRCA, intervalli di credibilità (95%HPD), date di origine, località e posterior probability della località della radice e dei principali clade significativi.*

Nodo	tMRCA	Min	Max	Anno	Max	Min	Località	PP Località
Radice	1057	530	1631	955	1482	381	WAF	0.66
A3_{qs}	640	353	945	1372	1659	1067	WAF	0.85
A1/A6 MRCA	818	403	1275	1195	1610	738	WAF	0.64
A1	666	326	953	1346	1686	1059	EAF	0.81
<i>A1 Cosmo</i>	519	277	735	1493	1735	1277	EAF	0.74
<i>A1 Afr</i>	516	269	755	1496	1743	1257	EAF	0.8
ASIA	364	190	506	1648	1822	1506	AS	0.83
Haiti (A5)*	294	230	308	1718	1782	1704	HA	1
Haiti1 (A1)*	254	131	324	1758	1881	1688	HA	1
Haiti2 (A1)	263	129	358	1749	1883	1654	HA	1
LA (Brasile)	239	120	309	1774	1893	1704	LA	1
LA (Colombia)	273	127	374	1740	1886	1639	LA	0.68
Europa	302	153	406	1710	1859	1606	IT	1
A6	296	111	488	1716	1901	1524	WAF	1

L’analisi di migrazione con Spread (Figura 31) ha identificato legami significativi (BF>6) tra Africa Centro-occidentale e Africa Orientale (BF=43.3), tra Africa Orientale e Sud Africa (BF>>100), tra Sud Africa e Asia (BF=22.3), tra Africa Orientale e Haiti (BF=22.1) e tra Haiti e America Latina (BF=60.7).



Figura 31. *Flussi significativi di migrazione nello spazio del genotipo A. Sono mostrati i flussi supportati da un Bayes Factor (BF) >6.*

5. Discussione

Questo lavoro di tesi rappresenta uno studio della prevalenza e distribuzione dei diversi genotipi/sottogenotipi di HBV in Italia, della storia epidemiologica dell'infezione negli ultimi decenni, in particolare della dinamica di popolazione, e della diffusione geografica a livello globale di questo virus mediante un metodo di analisi basato sulla teoria della coalescenza.

Dati di letteratura indicano come il genotipo D di HBV sia largamente presente in Europa, essendo responsabile di più del 70-80% delle infezioni in diversi Paesi dell'Est (Russia, regione Baltica, Bielorussia, Ungheria, Serbia)^{88,130,249} e del bacino del Mediterraneo (Croazia, Albania, Italia, Grecia, Turchia, Nord Africa)^{128,130,165, 166, 224, 250}. Al contrario, una bassa prevalenza di questo genotipo è stata trovata nei Paesi Occidentali del Mediterraneo, quali Spagna e Francia^{251, 252}.

La distribuzione e l'epidemiologia molecolare dei sottogenotipi D in Europa non è ancora chiara ed è presumibilmente complessa. Diversi studi recenti hanno mostrato che il sottogenotipo D1 sarebbe frequente in Medio Oriente, in alcuni Paesi dell'Est del Mediterraneo (Grecia e Turchia)^{220, 250} e nel Nord Africa (Tunisia, Marocco Algeria), dove sono stati evidenziati anche altri sottogenotipi, quali D2, D3 e D4^{129,131,167}. Diversamente, il sottogenotipo D2 risulta ampiamente diffuso nei Paesi dell'Europa Orientale, dove rappresenta più dell'80% degli isolati di genotipo D (Russia e Bielorussia). Il sottogenotipo D3 è invece prevalente in Italia e Serbia^{130,224} ed è probabilmente presente in diversi paesi occidentali, in particolare tra i soggetti tossicodipendenti (IVDUs)¹¹⁷.

In particolare, per quanto riguarda la distribuzione dei genotipi tra i pazienti italiani, i nostri dati mostrano una netta prevalenza del genotipo D (72.2%) e in particolare del sottogenotipo D3 (>80%), seguito dal genotipo A (18.7%).

Per quanto riguarda il genotipo A, principalmente osservato nei pazienti omosessuali maschi (66,7%, p=0.016), è stata riscontrata la presenza di due dei sottogenotipi principali: A1 (in una minoranza dei casi) e A2 (in più del 90% dei casi).

L'analisi filodinamica delle sequenze D3 in studio, è stata eseguita per valutare la storia epidemiologica di HBV nel periodo corrispondente alla diffusione del sottogenotipo virale maggiormente prevalente in Italia.

L'analisi della curva skyline del sottogenotipo D3 suggerisce una duplice crescita del numero effettivo di infezioni, una attorno alla fine degli anni '60 e la seconda negli anni '80. Un'ulteriore caratterizzazione di queste due fasi di crescita effettuata sulle sequenze che includevano o non includevano la mutazione T125M ha confermato i dati precedenti.

In particolare, lo skyline condotto sul dataset privo delle sequenze di ceppo IVDUs mostra un livello costante di endemia fino agli anni '50, con una successiva crescita esponenziale tra gli anni '50 e gli anni '70, quando la curva raggiunge un plateau che perdura fino ad oggi, nonostante un'apparente tendenza (non significativa) alla diminuzione negli ultimi 20 anni. L'analisi della curva delle sole sequenze IVDUs mostra, invece, una crescita esponenziale più tardiva (fine anni '60) che si conclude nei primi anni '80, quando la curva raggiunge il plateau.

Ciò suggerisce come la diffusione in Italia del sottogenotipo predominante tra le infezioni cosiddette "*community acquired*" sia dovuta a due fenomeni distinti. Il primo sembra corrispondere al periodo in cui si faceva largo uso di pratiche trasfusionali di sangue e derivati in assenza di metodi validi di screening dei donatori (con ciò si giustifica anche il plateau raggiunto negli anni '70, quando sono stati introdotti i test per HBsAg nella selezione di donatori); il secondo corrisponderebbe al lasso di tempo compreso tra gli anni '70 e '80 quando la circolazione del clade tipico dei soggetti tossicodipendenti è imputabile alla rilevante diffusione della tossicodipendenza in quegli anni.

Ulteriori studi sono tuttavia necessari per confermare la riduzione del numero di infetti negli ultimi decenni, probabilmente in relazione a diversi fattori, tra cui l'introduzione nel 1997 della vaccinazione anti-epatite B, obbligatoria per alcune categorie di soggetti.

In conclusione, i dati prodotti dal presente lavoro di tesi indicano che in Italia l'ampia diffusione del sottogenotipo D3 si sarebbe verificata nel periodo compreso tra gli anni '50 e '70, sostenuta principalmente dal largo impiego di trasfusioni di

sangue non sicure, con l'eccezione del ceppo prevalente nei tossicodipendenti, la cui crescita esponenziale è avvenuta in epoca successiva.

L'analisi filogeografica suggerisce che l'origine dei sottogenotipi D1, D2, D3, D5 e D7, deve essere fatta risalire ai primi decenni del '900 in un'area inclusa nel subcontinente Indiano. E' noto che in India coesistono tutti i sottogenotipi menzionati, con prevalenze variabili a seconda della zona geografica: all'est predomina il sottogenotipo D3, scarsamente presente a ovest (dove è prevalente il sottogenotipo D2) e a nord (dove è prevalente il D1) ²⁵³, seguito da D2, D5 e D1¹³¹. Considerata la grande eterogeneità dei ceppi di genotipo D presente in India, è assai probabile che essa rappresenti il crogiolo in cui tale genotipo si è originato e diversificato. La via di diffusione verso il resto del Vecchio Mondo è meno chiara, anche se dall'analisi condotta nel presente studio possono essere tratti alcuni punti fermi.

I sottogenotipi D1 e D2, strettamente correlati tra loro, sembrano avere un'origine comune nei territori dell'ex-Unione Sovietica ma, mentre D2 risulta il ceppo largamente prevalente nel Nord Est, D1 sembra essersi diffuso in Asia Centrale e Turchia e aver quindi invaso il bacino del Mediterraneo. Ciò può giustificare come l'Albania, che rappresenta una porta tra l'area Mediterranea e l'Est Europa, mostri sia un'elevata prevalenza di D2, che una presenza significativa di D1. Va inoltre sottolineato che i ceppi albanesi tendono a segregare tra loro tanto nel D2 quanto nel D1.

Più oscura sembra essere invece la strada seguita dal sottogenotipo D3. Sulla base dell'indagine qui riportata, l'Italia risulta essere il luogo di origine di questo ceppo, anche se il dato ottenuto potrebbe essere stato influenzato dalla maggiore numerosità di isolati Italiani appartenenti a tale sottogenotipo. Nell'albero filogeografico la presenza di un isolato indiano nel clade D3 suggerirebbe comunque un'origine indiana anche di questa variante virale, che, secondo le nostre stime si sarebbe diffusa in Italia circa 40-50 anni fa. La relativa scarsità di sequenze D3 provenienti da altre aree geografiche rende difficoltosa l'indagine.

Quindi, sulla base della nostra ricostruzione filogeografica possiamo tentare di formulare un'ipotesi sull'origine e diffusione del genotipo D: i nostri dati suggeriscono che alla fine del XIX sec., il progenitore del genotipo D ha lasciato l'India e raggiunto l'Asia Centrale dove i sottogenotipi si sono separati. Nella prima decade del XX secolo dall'Asia Centrale si sono successivamente diffusi in Europa e nell'area del Mediterraneo attraverso due direzioni principali:

- in direzione sud-occidentale, principalmente dovuta alla diffusione del sottogenotipo D1, attraverso il Medio Oriente (dove molto probabilmente il D1 si è separato) fino a raggiungere il nord Africa e il sud-est del Mediterraneo;
- in direzione nord ovest, strettamente associata al sottogenotipo D2, attraverso l'ex Unione Sovietica e Albania, fino a raggiungere l'Europa Orientale e il Mediterraneo.

Una terza epidemia di HBV in Eurasia non ha ancora spiegazioni ed è rappresentata dalla diffusione del sottogenotipo D3.

Ulteriori studi si rendono quindi necessari per poter completare l'analisi filogeografica utile alla ricostruzione dell'origine di tale ceppo.

Il secondo genotipo più diffuso in Italia è risultato HBV-A, in particolare il sottogenotipo A2. In uno studio recente abbiamo osservato un'associazione significativa di questo sottogenotipo con la trasmissione sessuale, in particolare tra gli omosessuali maschi ²²⁴. Diversa sembra essere la storia del sottogenotipo A1 (vedi di seguito). Per tale motivo abbiamo inoltre voluto studiare la filodinamica di questo genotipo.

La difficoltà principale nel ricostruire la scala temporale dell'evoluzione di HBV è data dalla discordanza sui tassi di sostituzione virale, dovuta a risultati contrastanti ottenuti in diversi studi. La maggior parte di queste controversie sembra essere legata all'uso improprio di differenti metodi di calibrazione. In particolare, l'utilizzo di eventi recenti per la calibrazione dell'orologio molecolare (come il campionamento eterocrono) ha prodotto alti tassi di sostituzione di HBV che supportano un'evoluzione relativamente recente del virus nell'uomo. Al contrario, l'impiego di una calibrazione basata su dati storici o paleontologici, ha

prodotto tassi evolutivi più bassi ²⁵⁴ che suggeriscono un'antica coesistenza del virus con l'uomo ²⁵⁵.

Diverse evidenze sono a sostegno di un'antica introduzione di HBV come agente infettivo nella specie umana, come gli studi sul genotipo F nelle tribù aborigene del Sud America, che sostengono un'introduzione precolombiana di questo virus nelle Americhe ²⁵⁶, o quelli di prevalenza del sottogenotipo C2 in una piccola tribù delle Isole Andamane (gli Jarawa), rimasta in completo isolamento per migliaia di anni ²¹⁹.

I tassi di sostituzione di HBV tendono a cambiare nel corso del tempo, in relazione ad alcuni fattori quali la modalità di diffusione e le caratteristiche demografiche delle popolazioni infette. Genotipi virali presenti da centinaia di anni come infezioni "endemiche" in popolazioni confinate e trasmessi soprattutto nella prima infanzia (ad esempio HBV-F o HBV-C2), si evolvono a un ritmo inferiore rispetto a genotipi definiti "epidemici", largamente diffusi solo di recente (nel XX secolo), attraverso una trasmissione altamente efficiente in popolazioni ad elevata densità (genotipi D1-3 in Eurasia e genotipo E in Africa) ²⁵⁵.

Il genotipo A di HBV, in particolare, rappresenta un modello appropriato per studiare gli effetti dei diversi metodi di calibrazione sulla ricostruzione cronologica di HBV, data la complessa storia epidemiologica e le diverse scale temporali di dispersione dei suoi sottogenotipi. In particolare, mentre il sottogenotipo A2 ha le caratteristiche di un ceppo "epidemico" recentemente diffuso nelle popolazioni adulte, soprattutto attraverso la trasmissione sessuale ^{224,233}, i sottogenotipi A1, A3_{QS} e A6, strettamente correlati alle popolazioni africane e a quelle zone interessate da migrazioni relativamente recenti dall'Africa (come India o America Latina), conservano la caratteristica di ceppi prevalentemente "endemiche" ^{166,235,257}.

Dei due dataset costruiti, quello che comprende i ceppi «endemiche» (sottogenotipi A1, A3_{QS} e A6) è stato analizzato fissando come punto di calibrazione la radice del clade haitiano A5, che è stata stimata tra la fine del XVII e la fine del XVIII secolo²⁵⁸. Questa calibrazione ci ha permesso di ottenere un valore di tasso di sostituzione di 3.0×10^{-5} sostituzioni/sito/anno, più basso, come atteso, rispetto a quello stimato

per HBV-A2, utilizzando un approccio di calibrazione diverso (1.1×10^{-4} sostituzioni/sito/anno).

Sulla base della nostra ricostruzione spazio-temporale l'antenato comune dei sottogenotipi A attualmente circolanti si sarebbe originato in Africa Centro-occidentale in media 1057 anni fa, ovvero intorno all'anno 1000 dC.

Intorno al XIV secolo due clade molto significativi segregano tra loro: un clade comprendente la maggior parte dei sottogenotipi dell'Africa Centro-occidentale (A3_{QS}) e un clade composto da isolati di sottogenotipo A1. Un terzo piccolo clade più divergente, composto da 3 isolati A6 di origine belga-africana, si separa solo più tardi, tra il XVII e il XVIII secolo.

Per quanto riguarda il luogo di origine, l'Africa Centro-occidentale risulta essere il più probabile sia per il sottogenotipo A3_{QS} che per il sottogenotipo A6, mentre per HBV-A1 l'origine risulta essere l'Africa Orientale. E' difficile ricostruire gli eventi che possono aver influenzato le prime divergenze del genotipo A in Africa nel periodo medioevale. Probabilmente la diffusione di HBV-A dall'Africa Occidentale a quella Orientale (tra il 1000 e la metà del 1300) è correlata alla tarda espansione delle popolazioni Bantu dalle regioni africane dell'ovest a quelle del sud-est. Tale cultura origina dalla regione che rappresenta l'odierno confine sud-occidentale tra Nigeria e Camerun ²⁵⁹, che coincide approssimativamente con la zona di origine di HBV-A, e successivamente si è diffusa verso sud-est, dove secondo il nostro studio ha avuto origine il sottogenotipo A1.

Il sottogenotipo A1 si ramifica ulteriormente in due gruppi altamente significativi intorno al XV secolo: uno ("Africano") contenente le sequenze dell'Africa Orientale e Meridionale, e l'altro ("Cosmopolita") contenente gli isolati provenienti da Asia, Europa, Africa e America. Gli isolati asiatici (principalmente provenienti da India, Bangladesh, Giappone, Filippine e Nepal) formano un unico clade, con un basso valore di significatività, il cui MRCA stimato suggerisce un'origine nel XVI secolo circa. I frequenti contatti tra le popolazioni bantu e quelle asiatiche, già a partire dal XII secolo (come dimostrano le influenze arabe e persiane nella cultura swahili ²⁶¹), dovuti principalmente a commerci lungo la costa dell'Africa Orientale e alla tratta degli schiavi attraverso l'Oceano Indiano nel XV-XVII secolo,

rappresentarono importanti vie di diffusione del ceppo “Cosmopolita” HBV-A1 verso il subcontinente indiano ¹¹¹. La presenza di schiavi africani in India e in tutto il mondo arabo nel XV secolo ²⁶² e la tratta degli schiavi portoghese, responsabile della migrazione di migliaia di schiavi dall’Africa Orientale verso subcontinente indiano e Filippine spagnole nel XVI secolo (periodo corrispondente al tMRCA del clade asiatico) possono aver giocato un ruolo fondamentale ²⁶³.

Con l’eccezione di un’unica sequenza sudafricana inclusa nel clade asiatico, gli unici isolati africani del ceppo “Cosmopolita” sono quelli ottenuti da pazienti provenienti dalla Somalia, e la posizione di outgroup della maggior parte degli isolati somali all’interno di questo sottoclade suggerisce che questo Paese abbia svolto un ruolo centrale nell’origine di tale ceppo. Tra il XVI e il XVIII secolo, la città di Mogadiscio è stata non solo uno dei principali porti di imbarco per la tratta degli schiavi dalla costa Swahili all’Asia, ma anche una meta frequente degli schiavi Bantu, impiegati i lavori più duri nelle piantagioni. Questo potrebbe spiegare la presenza prevalente del sottogenotipo A1 “Cosmopolita” in Somalia ²⁶².

Tutti gli altri isolati dell’Africa Orientale e Meridionale rientrano invece nel clade “Africano”, formando gruppi significativi che si originarono a partire dalla metà del 1400. È interessante notare che i due isolati di Pemba (un’isola dell’arcipelago di Zanzibar) formano un gruppo monofiletico altamente significativo con due sequenze ottenute in Malawi e Tanzania e risalenti al XVIII secolo. È infatti ben noto che Zanzibar rappresentava uno dei principali centri coinvolti nella tratta degli schiavi dal tempo del Sultanato dell’Oman nella metà del XVIII secolo ²⁶⁴.

Il ruolo centrale svolto dal commercio degli schiavi nella diffusione del genotipo A di HBV è suggerito anche dalla sua elevata prevalenza nelle comunità afroamericane dell’America Latina ²⁶⁵. La popolazione di Haiti è costituita principalmente dai discendenti degli schiavi africani che arrivarono durante il periodo del colonialismo francese. Il numero di schiavi è passato da poche migliaia a circa mezzo milione tra la fine del 1600 e il 1791, quando la rivoluzione degli schiavi di Haiti ha portato alla sconfitta dei francesi ²⁵⁸, periodo in cui ad Haiti arrivarono i genotipi HBV-A5 e HBV-A1 ¹¹³.

Sorprendentemente, A1 è l'unico sottogenotipo trovato nelle comunità nere nel resto del Sud America e forma due clade che riflettono la suddivisione tra le ex colonie spagnole e quelle portoghesi nel periodo compreso tra l'inizio e la fine del XVIII secolo. Questo dato, apparentemente, contraddice l'assenza di HBV-A1 in Africa Occidentale, che rappresentava la principale fonte di schiavi venduti in America Latina (95% del totale nei primi anni del 1700), in quanto l'importazione di schiavi dall'Africa Sud-orientale è stata ostacolata da alti tassi di mortalità dovuta ai lunghi tempi di attraversamento (Figura 32) ²⁵⁸. In un recente studio condotto in Brasile, Lago e colleghi ²⁶⁶ suggeriscono che la principale fonte di ingresso di HBV-A in Brasile sia rappresentata dall'Africa Sud-orientale ipotizzando che il genotipo A sia stato importato a metà del XIX secolo. La nostra stima di tMRCA suggerisce invece che l'origine del clade latino-americano si sia verificata, con tutta probabilità, prima di questo periodo (intorno al XVIII secolo).

Una possibile spiegazione della distribuzione dei sottogenotipi A di HBV tra gli abitanti di origine africana dell'America Latina è data dal traffico di schiavi portoghese e francese dall'Africa Sud-orientale nel corso del XVII e XVIII secolo. Un'ipotesi alternativa è che il sottogenotipo A1 fosse presente in Africa Centro-occidentale, ma di recente sia stato sostituito da HBV-E, che attualmente rappresenta il genotipo di HBV più diffuso in quest'area. La limitata diversità genetica e la scarsa diffusione di HBV-E tra gli afro-americani suggeriscono che tale genotipo si sia diffuso solo recentemente in seguito ad un uso improprio di aghi durante le campagne di vaccinazione di massa tra il 1920 e il 1960 ²⁶⁷, sostituendo il sottogenotipo meno prevalente A1 nella parte orientale del continente.

All'interno della nostra casistica un gruppo di pazienti italiani presenta il sottogenotipo A1. Curiosamente, la maggior parte degli isolati italiani forma un clade altamente significativo che suggerisce un *effetto del fondatore* caratterizzato da un'unica introduzione di HBV-A1 in Italia seguita da una circolazione autoctona. A differenza di altri paesi europei come Belgio, Francia e Gran Bretagna, l'Italia non è stata coinvolta nelle migrazioni storiche dall'Africa, avendo una storia di colonialismo relativamente breve: la regione costiera Swahili della Somalia

divenne un protettorato italiano alla fine del 1880 e successivamente (durante l'epoca fascista) una colonia italiana fino al 1941. Sono quindi necessari ulteriori studi per chiarire l'origine di HBV-A1 in Italia.

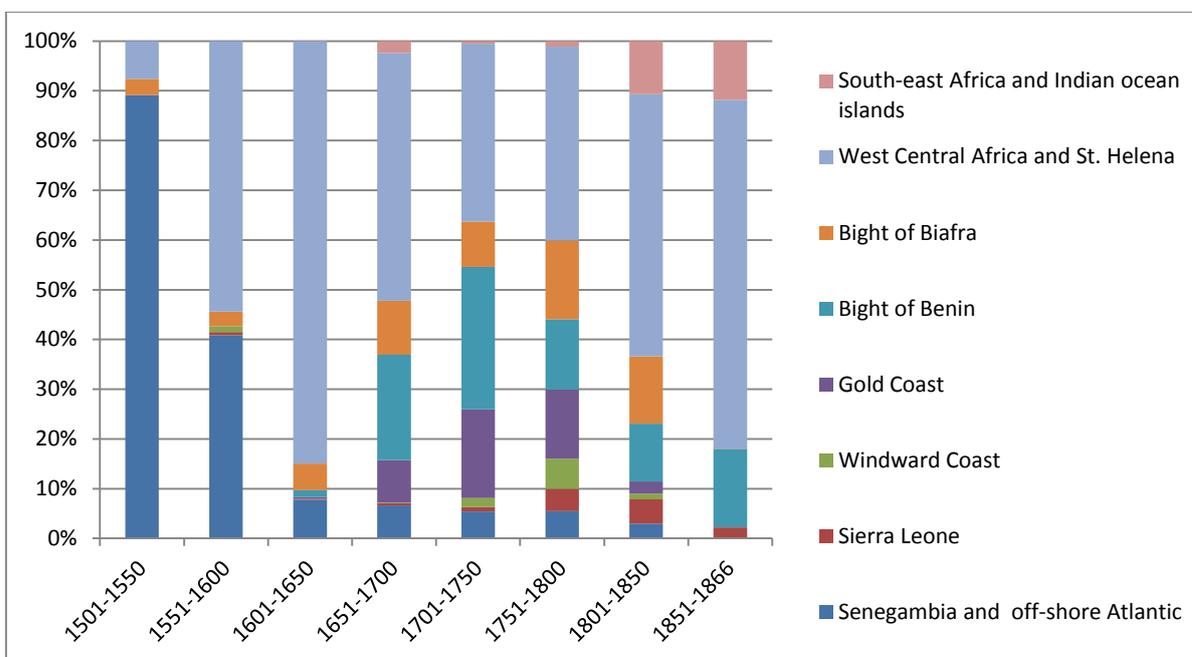


Figura 32. *Origini della tratta degli schiavi transatlantica tra il XVI e il XIX secolo. I dati sono disponibili presso "The Trans-Atlantic Slave Trade Database" (<http://www.slavevoyages.org/tast/index.faces>).*

L'analisi separata della filogenesi di HBV-A2 in scala temporale suggerisce un'origine più recente dei ceppi attualmente in circolazione, in particolare nei primi decenni del XX secolo. Gli isolati provenienti da tutta Europa e dagli Stati Uniti sono mescolati tra di loro, e sono privi di una distinta struttura geografica. Una serie di piccoli gruppi, soprattutto italiani, presentano un tMRCA compreso tra il 1970 e il 1990, e l'analisi filodinamica ha mostrato una crescita esponenziale del numero effettivo di infezioni in quegli stessi anni. Questi dati, in linea con quelli ottenuti in studi precedenti ⁵⁹, suggeriscono una recente diffusione di HBV-A2 tra i soggetti ad alto rischio di infezione, in particolare MSM ^{231, 233, 268, 269} e recentemente anche eterosessuali ²³³. La frequenza relativamente elevata di

pazienti HIV-positivi nei clade italiani, e la presenza di un gruppo di giovani donne rumene in cui il fattore di rischio è sconosciuto, suggerisce il ruolo centrale della trasmissione sessuale in Italia e in Romania. Sulla base di queste osservazioni, possiamo ipotizzare che HBV-A2, largamente diffuso come risultato di comportamenti sessuali ad alto rischio, è il primo sottogenotipo A ad essersi recentemente distinto nei paesi Occidentali (Europa e USA).

In conclusione, questo studio suggerisce che i genotipi di HBV si diffusero in Europa in tempi diversi e attraverso diverse vie di trasmissione. In particolare, alcuni genotipi (come HBV-D o HBV-A2), che definiamo "epidemic", si sono diffusi di recente (negli ultimi 100 anni), soprattutto attraverso la trasmissione sessuale, in particolare quella promiscua, e la trasmissione parenterale. Quest'ultima comprende l'esposizione iatrogena (come le trasfusioni di sangue non sicure o l'uso di siringhe non sterili) e la tossicodipendenza. Al contrario altri ceppi (come HBV-A1), che definiamo "endemic", si diffusero in tempi più antichi e attualmente circolano in individui di particolari origini geografiche o etnie, trasmettendosi prevalentemente attraverso il contatto intrafamiliare.

Appendice A. Accession number, località, anno di isolamento e genotipo/sottogenotipo delle 197 sequenze incluse nel dataset di HBV-D utilizzato per lo studio filogeografico.

BIELORUSSIA (BR):

EU414135- Belarus, 2005, D1
 EU414136- Belarus, 2005, D1
 EU414137- Belarus, 2005, D1
 EU414138- Belarus, 2005, D2
 EU414139- Belarus, 2005, D2
 EU414140- Belarus, 2005, D2
 EU414141- Belarus, 2005, D3
 EU414142- Belarus, 2005, D3
 EU414143- Belarus, 2005, D3

EGITTO-ISRAELE (EI):

EGITTO:

AB104709- Egypt, D1
 AB104711- Egypt, D1
 AB104712- Egypt, D1
 AB104710- Egypt, D1

ISRAELE:

L27106- Israele, D

ESTONIA-LITUANIA (ES):

ESTONIA:

EU594434- Estonia, D3
 EU594405- Estonia, D2
 AY653895- Estonia, D
 EU594414- Estonia, D2
 EU594413- Estonia, D2
 EU594410- Estonia, D2
 EU594409- Estonia, D2
 EU594408- Estonia, D2
 EU594407- Estonia, D2
 EU594406- Estonia, D2
 EU594405- Estonia, D2
 EU594404- Estonia, D2
 EU594403- Estonia, D2
 EU594402- Estonia, D2
 EU594401- Estonia, D2
 EU594400- Estonia, D2

LITUANIA:

X02496- Latvia, D1

ESTREMO ORIENTE (EO):

CINA:

GU357846- China, 2006, D1
 FJ562309- China, 2007, D1
 FJ386590- China, 2008, D1
 GQ377589- China, 2007, D1
 AF280817- China, D1

MONGOLIA:

DQ111987- Mongolia, D3

TIBET:

AY057948- Tibet, D

VIETNAM:

AF121239- Vietnam, D

GRECIA (GR):

AY341335- Greece, D2
 X80924- Greece, D
 X80926- Greece, D1
 X80925- Greece, D2

INDIA (IN):

GQ205377- India, 2007, D5
 GQ205378- India, 2007, D5
 GQ205379- India, 2007, D5
 GQ205380- India, 2007, D5
 GQ205381- India, 2008, D5
 GQ205382- India, 2008, D5
 GQ205383- India, 2008, D5
 GQ205384- India, 2008, D5
 GQ205385- India, 2008, D5
 GQ205386- India, 2008, D5
 GQ205387- India, 2008, D5
 GQ205388- India, 2008, D5
 GQ205389- India, 2008, D5
 DQ315777- India, 2003, D3
 DQ315776- India, 2000, D3

AB246347- India, D1
 AY161150- India, D
 AY161157- India, D
 AY161159- India, D
 AY161161- India, D1
 AY373430- India, D3
 AY945307- India, D1
 DQ315777- India, D3
 EF103276 - India, D
 DQ315778- India, D1
 DQ315779- India, D5
 AB090269- India, D2
 AB090268- India, D2

IRAN (IN):

GU456684- Iran, 2007, D1
 GU456683- Iran, 2007, D1
 GU456682- Iran, 2007, D1
 GU456681- Iran, 2007, D1
 GU456680- Iran, 2007, D1
 GU456679- Iran, 2007, D1
 GU456677- Iran, 2007, D1
 GU456635- Iran, 2007, D2
 AY741794- Iran, D1
 AY741795- Iran, D1
 AY741796- Iran, D1
 AY741797- Iran, D1
 AY741798- Iran, D1

KAZAKHSTAN- UZBEKISTAN (KU):**KAZAKHSTAN:**

EU594396- Kazakhstan, D1
 EU594428- Kazakhstan, D2

UZBEKISTAN:

AB188244- Uzbekistan, D1
 AB222709- Uzbekistan, D
 AB222710- Uzbekistan, D1
 AB222711- Uzbekistan, D1
 AB222712- Uzbekistan, D1
 AB222713- Uzbekistan, D1
 AB222710- Uzbekistan, D1

POLONIA (PL):

GQ477459- Polonia, 2006, D1
 GQ477458- Polonia, 2006, D1
 GQ477457- Polonia, 2006, D2
 GQ477456- Polonia, 2006, D2
 GQ477455- Polonia, 2006, D2
 GQ477454- Polonia, 2006, D2
 GQ477453- Polonia, 2006, D2
 GQ477452- Polonia, 2006, D2
 Z35716- Polonia, D2

RUSSIA (RU):

AB126581- Russia, 2000, D1
 AB126581- Russia, D1
 AB205127- Russia, D2
 AB205128- Russia, D2
 EU594382- RussiaD3
 EU594433- Russia, D2
 EU594436- Russia, D3
 EU594382- Russia, D3
 EU594433- Russia, D2
 EU594432- Russia, D2
 EU594431- Russia, D2
 EU594430- Russia, D2
 EU594429- Russia, D2
 EU594427- Russia, D2
 EU594426- Russia, D2
 EU594425- Russia, D2
 EU594424- Russia, D2
 EU594423- Russia, D2
 EU594422- Russia, D2
 EU594421- Russia, D2
 EU594420- Russia, D2
 EU594419- Russia, D2
 EU594418- Russia, D2
 EU594417- Russia, D2
 EU594416- Russia, D2
 EU594415- Russia, D2
 EU594399- Russia, D2
 EU594398- Russia, D2

SPAGNA (SP):

AJ627220- Spain, D2
 AJ627222- Spain, D2
 AJ627223- Spain, D2

SUDAFRICA (SA):

AY233291- South Africa, D3
AY233292- South Africa, D3
AY233293- South Africa, D3
AY233294- South Africa, D3
AY233295- South Africa, D3
AY233296- South Africa, D3

TUNISIA (TN):

FJ904444- Tunisia, 2006, D7
FJ904447- Tunisia, 2006, D7
FJ904442- Tunisia, 2005, D7
FJ904438- Tunisia, 2006, D7
FJ904437- Tunisia, 2006, D7
FJ904435- Tunisia, 2006, D7
FJ904433- Tunisia, 2005, D7
FJ904430- Tunisia, 2005, D7
FJ904428- Tunisia, 2006, D7
FJ904425- Tunisia, 2006, D7
FJ904417- Tunisia, 2006, D7
FJ904416- Tunisia, 2006, D7
FJ904414- Tunisia, 2006, D7
FJ904413- Tunisia, 2006, D7
FJ904410- Tunisia, 2006, D7
FJ904409- Tunisia, 2006, D7
FJ904408- Tunisia, 2006, D7
FJ904406- Tunisia, 2006, D7
FJ904405- Tunisia, 2006, D7
FJ904403- Tunisia, 2006, D7
FJ904401- Tunisia, 2006, D7
FJ904400- Tunisia, 2006, D7
FJ904398- Tunisia, 2006, D7
FJ904396- Tunisia, 2006, D7
FJ904394- Tunisia, 2005, D7

FJ904446- Tunisia, 2006, D1
FJ904445- Tunisia, 2006, D1
FJ904443- Tunisia, 2005, D1
FJ904432- Tunisia, 2006, D1
FJ904429- Tunisia, 2006, D1
FJ904427- Tunisia, 2006, D1
FJ904426- Tunisia, 2006, D1
FJ904424- Tunisia, 2006, D1
FJ904421- Tunisia, 2006, D1
FJ904420- Tunisia, 2006, D1
FJ904415- Tunisia, 2006, D1
FJ904412- Tunisia, 2006, D1
FJ904402- Tunisia, 2006, D1
FJ904399- Tunisia, 2005, D1
FJ904431- Tunisia, D1
FJ904418- Tunisia, D1

TURCHIA (TR):

AY721605- Turkey, 2004, D1
AY721609- Turkey, 2004, D1
AY721611- Turkey, 2004, D1
AY796030- Turkey, 2004, D1
AF121239- Turkey, 1993, D1
AF121240- Turkey, 1994, D1
AF121241- Turkey, 1995, D1
AF121242- Turkey, 1998, D1
AY796031- Turkey, 2004, D2
AY721606- Turkey, D1
AY721607- Turkey, D1
AY721608- Turkey, D1
AY721612- Turkey, D1
AY796032- Turkey, D
AY661792- Turkey, D
AY661793- Turkey, D
M32138- Turkey, D1

Appendice B. *Accession number e località delle sequenze di riferimento incluse nei 2 dataset di sequenze di genotipo A di HBV.*

Accession Number	Country	Accession Number	Country
EU366129	Argentina	FN545836	Camerun
AB116085	Bangladesh	FN545835	Camerun
AB116083	Bangladesh	FN545834	Camerun
AB116082	Bangladesh	FN545833	Camerun
AB116084	Bangladesh	FN545832	Camerun
GQ331048	Belgio	FN545831	Camerun
GQ331047	Belgio	FN545830	Camerun
GQ331046	Belgio	FN545829	Camerun
EU859952	Belgio	FN545828	Camerun
GU563558	Belgio	FN545826	Camerun
GU563554	Belgio	AB194952	Camerun
GU563546	Belgio	AB194951	Camerun
FJ349224	Belgio	AM494718	Repubblica Centrafrican:
FJ349222	Belgio	JQ023660	Colombia
EU859953	Belgio	JQ023661	Colombia
EU859949	Belgio	JQ023662	Colombia
EU859948	Belgio	JQ023663	Colombia
EU859942	Belgio	DQ020002	Congo
GU563557	Belgio	GQ486663	Repubblica Ceca
GU563555	Belgio	GQ486544	Repubblica Ceca
GU563562	Belgio	GQ486306	Repubblica Ceca
HM101118	Brasile	AB263406	Repubblica Ceca
HM101118	Brasile	AJ344115	Francia
HM101129	Brasile	AJ309371	Francia
HM101124	Brasile	GQ486727	Francia
HM772996	Brasile	HE576988	Francia
HM772995	Brasile	HE576989	Francia
HM772997	Brasile	AM184126	Gabon
HM772994	Brasile	AM184125	Gabon
HM101119	Brasile	AY934764	Gambia
HM101122	Brasile	AY934763	Gambia
AM180624	Camerun	AF143305	Germania
AB194952	Camerun	AF143304	Germania
AB194951	Camerun	AF143303	Germania
AB194950	Camerun	AF143302	Germania
FN545840	Camerun	AF143300	Germania
FN545839	Camerun	AF143299	Germania
FN545838	Camerun	AF143298	Germania

FN545837	Camerun	GQ486568	Germania
GQ486712	Germania	FJ692568	Haiti
GQ486631	Germania	FJ692567	Haiti
GQ486712	Germania	FJ692566	Haiti
GQ161813	Guinea	FJ692565	Haiti
FJ692613	Haiti	FJ692564	Haiti
FJ692612	Haiti	FJ692563	Haiti
FJ692611	Haiti	FJ692562	Haiti
FJ692610	Haiti	FJ692561	Haiti
FJ692609	Haiti	FJ692560	Haiti
FJ692608	Haiti	FJ692559	Haiti
FJ692607	Haiti	FJ692558	Haiti
FJ692605	Haiti	FJ692557	Haiti
FJ692604	Haiti	AY161140	India
FJ692603	Haiti	EF103278	India
FJ692602	Haiti	AY373429	India
FJ692601	Haiti	AB116087	India
FJ692600	Haiti	AB116086	India
FJ692599	Haiti	AY161142	India
FJ692598	Haiti	AY161140	India
FJ692597	Haiti	AY161138	India
FJ692596	Haiti	AF418682	India
FJ692595	Haiti	AF418676	India
FJ692594	Haiti	AY373428	India
FJ692593	Haiti	AY373432	India
FJ692592	Haiti	AB246335	India
FJ692591	Haiti	AB453989	Giappone
FJ692590	Haiti	AB453988	Giappone
FJ692589	Haiti	AB453987	Giappone
FJ692588	Haiti	AB453986	Giappone
FJ692585	Haiti	AB116091	Giappone
FJ692584	Haiti	AP007263	Giappone
FJ692583	Haiti	AB076679	Malawi
FJ692582	Haiti	AB076678	Malawi
FJ692581	Haiti	AM180623	Mali
FJ692580	Haiti	AB116089	Nepal
FJ692579	Haiti	AB116088	Nepal
FJ692578	Haiti	AB116090	Nepal
FJ692577	Haiti	FN545825	Nigeria
FJ692576	Haiti	FJ692556	Nigeria
FJ692575	Haiti	FJ692554	Nigeria
FJ692574	Haiti	AY934774	Filippine
FJ692572	Haiti	AB241115	Filippine
FJ692571	Haiti	AB241114	Filippine

FJ692569	Haiti	AY233286	South Africa
AB116092	Filippine	AY233285	South Africa
M57663	Filippine	AY233284	South Africa
Z35717	Polonia	AY233283	South Africa
GQ477503	Polonia	AY233282	South Africa
GQ477498	Polonia	AY233281	South Africa
GQ477492	Polonia	AY233279	South Africa
GQ477467	Polonia	AY233278	South Africa
GQ477461	Polonia	AY233277	South Africa
GQ486583	Polonia	AY233276	South Africa
GQ486490	Polonia	AY233275	South Africa
GQ486454	Polonia	AY233274	South Africa
GQ486567	Polonia	AF297623	South Africa
GQ486426	Polonia	AY233280	South Africa
GQ486425	Polonia	U87742	South Africa
GQ477488	Polonia	AF297625	South Africa
GQ477486	Polonia	AF297621	South Africa
GQ477460	Polonia	AB246336	South Africa
GQ486856	Polonia	AY934772	Uganda
GQ486408	Polonia	DQ020003	United Arab Emirates
GQ486616	Polonia	S50225	United Kingdom
EU414132	Polonia	GQ486409	United Kingdom
GQ486391	Polonia	AY034878	USA
GQ486416	Polonia	X02763	USA
FM199981	Rwanda	V00866	USA
FM199980	Rwanda	AY152726	USA
FM199979	Rwanda	AB064314	USA
FM199977	Rwanda	GQ486571	USA
FM199976	Rwanda	GQ486474	USA
FM199974	Rwanda	GQ486550	USA
JQ687533	Serbia	GQ486598	USA
JQ687529	Serbia	GQ486642	USA
AY934771	Somalia	GQ486836	USA
AY934770	Somalia		
AY934769	Somalia		
AY934768	Somalia		
AY934767	Somalia		
AY934766	Somalia		
AY934765	Somalia		
AY903452	South Africa		
AY233290	South Africa		
AY233289	South Africa		
AY233288	South Africa		
AY233287	South Africa		

6. Bibliografia

1. **Ganem D.**, Prince A. Hepatitis B virus infection - natural history and clinical consequences. *N Engl J Med* 2004; 350: 1118-1129.
2. **Dane D.S.**, Cameron C.H., Briggs M. Virus-like particle in serum of patients with Australia antigen associated hepatitis B. *Lancet* 1970; 1(7649): 695-698.
3. **Piazza M.**, Epatite virale acuta e cronica. *VIII Edizione*. Ghedini, Napoli. 2005.
4. **Patient R.**, Hourieux C., Sizaret P.Y., Trassard S., Sureau C., Roingeard P. Hepatitis B virus subviral envelope particle morphogenesis and intracellular trafficking. *J Virol* 2007; 81: 3842-3851.
5. **Huovila A.P.**, Eder A.M., Fuller S.D. Hepatitis B surface antigen assembles in a post-ER, pre-Golgi compartment. *J Cell Biol* 1992; 118: 1305-1320.
6. **Robinson W.S.**, Lutwick L.I. The virus of hepatitis, type B. *N Engl J Med* 1976; 295: 1168-1175.
7. **Bruns M.**, Miska S., Chassot S., Will H. Enhancement of hepatitis B virus infection by noninfectious subviral particles. *J Virol* 1998; 72: 1462-1468.
8. **Schaefer S.** Hepatitis B virus taxonomy and hepatitis B virus genotypes. *World J Gastroenterol* 2007; 13(1): 14-21.
9. **Kann M.**, Gerlich W.H. Hepatitis B. In: *Microbiology and microbial infections. Virology*. B.W.J. Mahy, L. Collier. Eds. *Arnold; London* 1998; 745-775.
10. **Summers J.**, Mason W.S. Replication of the genome of a hepatitis-like virus by reverse transcription of an RNA intermediate. *Cell* 1990; 29: 403-415.
11. **Schädler S.**, Hildt E. HBV life cycle: entry and morphogenesis. *Viruses* 2009; 1: 185-209.
12. **Ganem D.**, Schneider R.J. Hepadnaviridae: the viruses and their replication. In: *Fields Virology*, 4th ed., Knipe D.M, Howley P.M, Eds; Lippincott Williams: Wilkins, Philadelphia, USA 2001; 2: 2923-2969.
13. **Locarnini S.**, McMillan J., Bartholomeusz A. The hepatitis B virus and common mutants. *Seminars in liver disease* 2003; 23(1): 5-20
14. **Klinkert M.**, Theilmann L., Praff E., Schaller H. Pre-S1 antigens and antibodies early in the course of acute hepatitis B virus infection. *J Virol* 1996; 58: 522-525.

15. **Gerlich W.H.**, Heermann K.H., Lu X. Functions of hepatitis B surface proteins. *Arch Virol Suppl* 1992; 4: 129-132.
16. **Bruss V.**, Ganem D. The role of envelope proteins in hepatitis B virus assembly. *Proc Natl Acad Sci USA* 1991; 88: 1059-1063.
17. **Mangold C.M.**, Streeck R.E. Mutational analysis of the cysteine residues in the hepatitis B virus small envelope protein. *J Virol* 1993; 67: 4588-4597.
18. **Khan N.**, Guarnieri M., Ahn S.H., Li J., Zhou Y., Bang G., Kim K.H., Wands J.R., Tong S. Modulation of hepatitis B virus secretion by naturally occurring mutations in the S gene. *J Virol* 2004; 78(7): 3262-3270.
19. **Okamoto H.**, Imai M., Tsuda F., Tanaka T., Miyakawa Y., Mayumi M. Point mutation in the S gene of hepatitis B virus for a d/y or w/r subtypic change in two blood donors carrying a surface antigen of compound subtype adyr or adwr. *J Virol* 1987; 61: 3030-3034.
20. **Norder H.**, Couroucé A.M., Magnius L.O. Molecular basis of hepatitis B virus serotype variations within the four major subtypes. *J Gen Virol* 1992; 73: 3141-3145.
21. **Couroucé-Pauty A.M.**, Lemaire J.M., Roux J.F. New hepatitis B surface antigen inside the ad category. *Vox Sang* 1978; 27: 533-549.
22. **Magnius L.O.**, Kaplan L., Vyas G.N., Perkins H.A. A new virus specified determinant of hepatitis B surface antigen. *Acta Pathol Microbiol Scand [B]* 1975; 83: 295-297.
23. **Norder H.**, Ebert J.W., Fields H.A., Mushahwar I.K., Magnius L.O. Complete sequencing of a gibbon hepatitis B virus genome reveals a unique genotype distantly related to the chimpanzee hepatitis B virus. *Virology* 1996; 218: 214-223.
24. **Wallace W.**, Carman W. Surface variation of HBV: scientific and medical relevance. *Viral Hepat* 1997; 3: 5-16.
25. **Bruss V.**, Gerlich W.H. Formation of transmembranous Hepatitis B e antigen by cotraslation in vitro processing of the viral precore protein. *Virology* 1988; 163: 268-275.
26. **Milich D.R.**, Sallberg M, Maruyama T. The humoral immune response in acute and chronic hepatitis B virus infection. *Springer Semin Immunopathol* 1995; 17: 14-66.
27. **Yen T.S.B.**, Hepadnaviral X protein: review of recent progress. *J Biomed Sci* 1996; 3: 20-30.

28. **Warris G**, Siddiqui A. Regulatory mechanisms of viral hepatitis B and C. *J Biosci* 2003; 3: 311-321.
29. **Kay A.**, Zoulim F. Hepatitis B virus genetic variability and evolution. *Virus Res* 2007; 127: 164-176.
30. **Will H.**, Reiser W., Weimer T. et al. Replication strategy of human hepatitis B virus. *J Virol* 1987; 61: 904-911.
31. **Block T.M.**, Guo H., Guo J.T. Molecular virology of hepatitis B virus for clinicians. *Clin Liver Dis.* 2007; 11(4): 685-706.
32. **Kidd-Ljunggren K.**, Miyakawa Y., Kidd A.H. Genetic variability in hepatitis B viruses. *J Gen Virol* 2002; 83: 1267-1280.
33. **The hepatitis B virus Page:** providing current information on the hepatitis B virus. Replication. *Robert's HBV Page:* <http://www.hepatitisbviruspage.com>
34. **Tutleman J.S.**, Pourcel C., Summers J. Formation of the pool of covalently close circular viral DNA in hepadnavirus-infected cells. *Cell* 1986; 47: 451-460.
35. **Ferrari C.**, Missale G., Boni C. et al. Immunopathogenesis of hepatitis B. *J Hepat* 2003; 39: S36-S42.
36. **Chisari F.V.** Hepatitis B virus immunopathogenesis. *Annu Rev Immunol* 1995; 13: 29-60.
37. **Guidotti L.G.**, Chisari F.V. To kill or to cure: options in host defense against viral infection. *Curr Opin Immunol* 1996; 8(4): 478-483.
38. **Chisari F.V.** Cytotoxic T cells and viral hepatitis. *J Clin Invest* 1997; 99: 1472-1477.
39. **Lee W.M.** Medical progress: hepatitis B virus infection. *N Engl J Med* 1997; 337 (24): 1733-1745.
40. **McMahon B.J.**, Alward W.L.M., Hall D.B., et al. Acute Hepatitis B virus infection: relation of age to the clinical expression of disease and subsequent development of the carrier state. *J Infect Dis* 1985; 151: 599-603.
41. **Mahoney F.J.** Update on Diagnosis, Management, and Prevention of Hepatitis B Virus Infection. *Clin Microbiol Reviews* 1999; 12(2): 351-366.

42. **Milich D.R.**, Jones J.E., Hughes J.L. et al. Is a function of the secreted hepatitis B e antigen to induce immunologic tolerance in utero? *Proc Natl Acad Sci USA* 1990; 87: 6599-6604.
43. **Tang B.**, Kruger W.D., Chen G., Shen F., Lin W.Y., Mboup S., London W.D., Evams AA. Hepatitis B viremia is associated with increased risk of hepatocellular carcinoma in chronic carriers. *J Med Virol* 2004, 72: 35-40.
44. **Colgrove R.**, Simon G., Ganem D. Transcriptional activation of homologous and heterologous genes by the hepatitis B virus X gene product in cells permissive for viral replication. *J Virol* 1988; 63: 4019-4026.
45. **European Association for the Study of the Liver.** EASL clinical practice guidelines: management of chronic hepatitis B. *J Hepatol* 2009; 50: 227-242.
46. **Kew M.C.** Epidemiology of chronic hepatitis B virus infection, hepatocellular carcinoma, and hepatitis B virus-induced hepatocellular carcinoma. *Pathol Biologie* 2010; 58: 273-277.
47. **Ioannou G.N.** Chronic hepatitis B infection: a global disease requiring global strategies. *Hepatology* 2013; 58(3):839-43.
48. **Dény P.**, Zoulim F. Hepatitis B virus: from diagnosis to treatment. *Pathol Biol* 2010; 58(4): 245-53.
49. **Gerlich W.H.** Medical virology of hepatitis B: how it began and where we are now. *Virology* 2013; 10:239.
50. **World Health Organization.** Childhood diseases in Africa, *Fact Sheet N109*. Geneva, Switzerland: World Health Organization 1996; <http://www.who.int/inf-fs/en/fact109html>
51. **Botha J.F.**, Ritchie M.J.J., Dusheiko G.M., Mouton H.W.K., Kew M.C. Hepatitis B virus carrier state in Black Children in Ovamboland: role of perinatal and horizontal transmission. *Lancet* 1984; 2: 1210-1212.
52. **Lok A.S.**, McMahon B.J. Chronic hepatitis B. *Hepatology* 2007; 45: 507-539.
53. **Alter M.J.** Epidemiology of hepatitis B in Europe and worldwide. *J Hepatol* 2003; 39: S64-69.

54. **Kao J.H.** Molecular Epidemiology of Hepatitis B Virus. *Korean J Intern Med* 2011; 26(3):255-61.
55. **Villeneuve J.P.** The natural history of chronic hepatitis B virus infection. *J Clin Virol* 2005; 34(Suppl1): S139-142.
56. **Schutte K.**, Bornschein J., Malfertheiner P. Hepatocellular carcinoma-epidemiological trends and risk factors. *Dig Dis* 2009; 27: 80-92.
57. **Chang M.H.**, You S.L., Chen C.J., Liu C.J., Lee C.M., Lin S.M. et al. Decreased incidence of hepatocellular carcinoma in hepatitis B vaccines: a 20-year follow-up study. *J Natl Cancer Inst* 2009; 101: 1348-1355.
58. **Medley G.F.**, Lindop N.A., Edmunds W.J., Nokes D.J. Hepatitis B virus endemicity: heterogeneity, catastrophic dynamics and control. *Nat Med* 2001; 7: 619-624.
59. **Zehender G.**, De Maddalena C., Giambelli C., Milazzo L., Schiavini M., Bruno R., Tanzi E., Galli M. Different evolutionary rates and epidemic growth of hepatitis B virus genotypes A and D. *Virology* 2008; 380: 84-90.
60. **Hoofnagle J.** Therapy of acute and chronic hepatitis B. *Adv Intern Med* 1994; 39: 241-275.
61. **Yuen M.F.**, Lai C.L. Treatment of chronic hepatitis B: Evolution over two decades. *J Gastroenterol Hepatol* 2011; 26 Suppl 1:138-143.
62. **Wong D.K.**, Cheung A.M., O'Rourke K. et al. Effect of alpha-interferon treatment in patients with hepatitis B e antigen-positive chronic hepatitis B. A meta-analysis. *Ann Intern Med* 1993; 119: 312-323.
63. **Marcellin P.**, Bonino F., Lau G.K. et al. Sustained response of hepatitis B e antigen-negative patients 3 years after treatment with peginterferon alpha-2a. *Gastroenterology* 2009; 136(7): 2169-2179.
64. **Yuen M.F.**, Seto W.K., Chow D.H. et al. Long-term lamivudine therapy reduces the risk of long-term complications of chronic hepatitis B infection even in patients without advanced disease. *Antivir Ther* 2007; 12(8): 1295-1303.
65. **Matsumoto A.**, Tanaka E., Rokuhara A. et al. Efficacy of lamivudine for preventing hepatocellular carcinoma in chronic hepatitis B: A multicenter retrospective study of 2795 patients. *Hepatol Res* 2005; 32(3): 173-184.

66. **Lai C.L.**, Yuen M.F. Chronic hepatitis B--new goals, new treatment. *N Engl J Med* 2008; 359(23): 2488-2491.
67. **Hadziyannis S.J.**, Tassopoulos N.C., Heathcote E.J. et al. Long-term therapy with adefovir dipivoxil for HBeAg-negative chronic hepatitis B for up to 5 years. *Gastroenterology* 2006; 131(6): 1743-1751.
68. **Chang T.T.**, Gish R.G., de Man R. et al. A comparison of entecavir and lamivudine for HBeAg-positive chronic hepatitis B. *N Engl J Med* 2006; 354: 1001-1010.
69. **Lai C.L.**, Shouval D., Lok A.S. et al. Entecavir *versus* lamivudine for patients with HBeAg-negative chronic hepatitis B. *N Engl J Med* 2006; 354: 1011-1020.
70. **Seto W.K.**, Lai C.L., Fung J., Yuen J., Wong D.K.H., Yuen M.F. A three-year study on viral suppression and resistance profile for treatment-naïve CHB patients receiving continuous entecavir treatment. *Hepatol Int* 2010; 4: 58.
71. **Chang T.T.**, Lai C.L., Kew Yoon S. et al. Entecavir treatment for up to 5 years in patients with hepatitis B e antigen-positive chronic hepatitis B. *Hepatology* 2010; 51: 422-430.
72. **Tenney D.J.**, Rose R.E., Baldick C.J. et al. Long-term monitoring shows hepatitis B virus resistance to entecavir in nucleoside-naïve patients is rare through 5 years of therapy. *Hepatology* 2009; 49: 1503-14.
73. **Suzuki Y.**, Suzuki F., Kawamura Y. et al. Efficacy of entecavir treatment for lamivudine-resistant hepatitis B over 3 years: histological improvement or entecavir resistance? *J Gastroenterol Hepatol* 2009; 24: 429-35.
74. **Leung N.** Treatment of HBeAg-positive chronic hepatitis B with nucleos(t)ide analogues. *Liver Int* 2011; 31 Suppl 1: 85-89.
75. **Gish R.G.**, Chang T.T., Lai C.L. et al. Loss of HBsAg antigen during treatment with entecavir or lamivudine in nucleoside-naïve HBeAg-positive patients with chronic hepatitis B. *J Viral Hepat* 2010; 17: 16-22.
76. **Lai C.L.**, Gane E., Liaw Y.F. et al. Telbivudine *versus* lamivudine in patients with chronic hepatitis B. *N Engl J Med* 2007; 357: 2576-2588.
77. **Hynicka L.M.**, Yunker N., Patel P.H. A review of oral antiretroviral therapy for the treatment of chronic hepatitis B. *Ann Pharmacother* 2010; 44(7-8): 1271-1286.

78. **Liaw Y.F.**, Gane E., Leung N. et al. 2-Year GLOBE trial results: telbivudine is superior to lamivudine in patients with chronic hepatitis B. *Gastroenterology* 2009; 136: 486–495.
79. **Fleischer R.D.**, Lok A.S. Myopathy and neuropathy associated with nucleos(t)ide analog therapy for hepatitis B. *J Hepatol* 2009; 51(4): 787-791.
80. **Snow-Lampart A.**, Chappell B.J., Curtis M. et al. Resistance surveillance for up to 144 weeks in HBeAg-positive and HBeAg-negative hepatitis B patients treated with tenofovir DF showed no relationship between virologic breakthrough and emergence of genotypic changes in HBV polymerase. *Hepatology* 2009; 50 (Suppl.): 532A.
81. **Yuen M.F.**, Han K.H., Um S.H. et al. Antiviral activity and safety of LB80380 in hepatitis B e antigen-positive chronic hepatitis B patients with lamivudine-resistant disease. *Hepatology* 2010; 51: 767–776.
82. **Lai C.L.**, Ahn S.H., Lee K.S. et al. Phase IIb multicentred randomised trial of besifovir (LB80380) versus entecavir in Asian patients with chronic hepatitis B. *Gut* 2013; doi: 10.1136/gutjnl-2013-305138.
83. **Echevarría J.M.**, Avellón A. Hepatitis B virus genetic diversity. *J Med Virol* 2006; 78: S36-S42.
84. **Schaefer S.** Hepatitis B virus: significance of genotypes. *J Viral Hepatitis* 2005; 12: 111-124.
85. **Okamoto H.**, Tsuda F., Sakugawa H. et al. Typing hepatitis B virus by homology in nucleotide sequence: comparison of surface antigen subtypes. *J Gen Virol* 1988; 69(Pt 10): 2575-2583.
86. **Yu H.**, Yuan Q., Ge S.X. et al. Molecular and phylogenetic analyses suggest an additional hepatitis B virus genotype “I”. *PloS One* 2010; 5(2): e9297.
87. **Arankalle V.A.**, Gandhe S.S., Borkakoty B.J., Walimbe A.M., Biswas D., Mahanta J. A novel HBV recombinant (genotype I) similar to Vietnam/Laos in a primitive tribe in Eastern India. *J Viral Hepat* 2010; 17: 501-510.
88. **Olinger C.M.**, Jutavijittum P., Hubschen J.M. et al. Possible new hepatitis B virus genotype, South-East Asia. *Emerg Infect Dis* 2008; 14: 1777-1780.

89. **Tran T.T.**, Trinh T.N., Abe K. New complex recombinant genotype of hepatitis B virus identified in Vietnam. *J Virol* 2008; 82: 5657-5663.
90. **Kurbanov F.**, Tanaka Y., Kramvis A., Simmonds P., Mizokami M. When should "I" consider a new hepatitis B virus genotype? *J Virol* 2008; 82: 8241-8242.
91. **Tatematsu K.**, Tanaka Y., Kurbanov F. et al. A genetic variant of hepatitis B virus divergent from known human and ape genotypes isolated from a Giapponeese patient and provisionally assigned to new genotype J. *J Virol* 2009; 83: 10538-10547.
92. **Mulders M.N.**, Venard V., Njyou M. et al. Low genetic diversity despite hyperendemicity of hepatitis B virus genotype E throughout West Africa. *J Infect Dis* 2004; 190 (2): 400-408.
93. **Abdo A.A.**, Al-Jarallah B.M., Sanai F.M., Hersi A.S., Al-Swat K. et al. Hepatitis B genotypes: relation to clinical outcome in patients with chronic hepatitis B in Saudi Arabia. *World J Gastroenterol* 2006; 12: 7019-7024.
94. **Stuyver L.**, De Gendt S., Van Geyt C., Zoulim F., Fried M., Schinazi R.F., Rossau R. A new genotype of hepatitis B virus: complete genome and phylogenetic relatedness. *J Gen Virol* 2000; 81: 67-74.
95. **Sanchez L.V.**, Maldonado M., Bastidas-Ramirez B.E., Norder H., Panduro A. Genotypes and S-gene variability of Mexican hepatitis B virus strains. *J Med Virol* 2002; 68(1): 24-32.
96. **Lindh M.** HBV genotype G - an odd genotype of unknown origin. *J Clin Virol* 2005; 34: 315-316.
97. **Arauz-Ruiz P.**, Norder H., Robertson B.H., Magnus L.O. Genotype H: a new Amerindian genotype of hepatitis B virus revealed in Central America. *J Gen Virol* 2002; 83: 2059-2073.
98. **Lusida M.I.**, Nugrahaputra V.E., Soetjipto, Handajani R. et al. Novel subgenotypes of hepatitis B virus genotypes C and D in Papua, Indonesia. *J Clin Microbiol* 2008; 46: 2160-2166.
99. **Nurainy N.**, Muljono D.H., Sudoyo H., Marzuki S. Genetic study of hepatitis B virus in Indonesia reveals a new subgenotype of genotype B in east Nusa Tenggara. *Arch Virol* 2008; 153: 1057-1065.

100. **Mulyanto**, Depamede S.N., Surayah K., Tsuda F., Ichiyama K., Takahashi M., Okamoto H. A nationwide molecular epidemiological study on hepatitis B virus in Indonesia: identification of two novel subgenotypes, B8 and C7. *Arch Virol* 2009; 154: 1047-1059.
101. **Theджа** M.D., Muljono D.H., Nurainy N. et al. Ethnogeographical structure of hepatitis B virus genotype distribution in Indonesia and discovery of a new subgenotype, B9. *Arch Virol* 2011; 156(5): 855-68.
102. **Mulyanto**, Depamede S.N., Surayah K., Tjahyono A.A., Jirintai, Nagashima S., Takahashi M., Okamoto H. Identification and characterization of novel hepatitis B virus subgenotype C10 in Nusa Tenggara, Indonesia. *Arch Virol* 2010; 155: 705-715.
103. **Mulyanto**, Depamede S.N., Wahyono A., Jirintai, Nagashima S., Takahashi M., Okamoto H. Analysis of the full-length genomes of novel hepatitis B virus subgenotypes C11 and C12 in Papua, Indonesia. *J Med Virol* 2011; 83: 54-64.
104. **Mulyanto**, Pancawardani P., Depamede S.N. et al. Identification of four novel subgenotypes (C13-C16) and two inter-genotypic recombinants (C12/G and C13/B3) of hepatitis B virus in Papua province, Indonesia. *Virus Res* 2012; 163(1):129-40.
105. **Kimbi G.C.**, Kramvis A., Kew M.C. Distinctive sequence characteristics of subgenotype A1 isolates of hepatitis B virus from South Africa. *J Gen Virol* 2004; 85: 1211-1220.
106. **Kramvis A.**, Weitzmann L., Owiredu W.K., Kew M.C. Analysis of the complete genome of subgroup A' hepatitis B virus isolates from South Africa. *J Gen Virol* 2002; 83(Pt 4): 835-839.
107. **Alvarado-Mora M.V.**, Botelho L., Gomes-Gouvêa M.S. et al. Detection of Hepatitis B virus subgenotype A1 in a Quilombo community from Maranhão, Brasile. *Virol J* 2011; 8:415.
108. **Owiredu W.K.**, Kramvis A., Kew M.C. Molecular analysis of hepatitis B virus genomes isolated from black African patients with fulminant hepatitis B. *J Med Virol* 2001; 65: 485-492.
109. **Kurbanov F.**, Tanaka Y., Fujiwara K., Sugauchi F., Mbanya D., Zekeng L., Ndembi N., Ngansop C., Kaptue L., Miura T., Ido E., Hayami M., Ichimura H., Mizokami M. A new subtype (subgenotype) Ac (A3) of hepatitis B virus and recombination between genotypes A and E in Camerun. *J Gen Virol* 2005; 86: 2047-2056.

110. **Hubschen J.M.**, Mugabo J., Peltier C.A. et al. Exceptional genetic variability of hepatitis B virus indicates that Rwanda is east of an emerging African genotype E/A1 divide. *J Med Virol* 2009; 81: 435-440.
111. **Hannoun C.**, Soderstrom A., Norkrans G., Lindh M. Phylogeny of African complete genomes reveals a West African genotype A subtype of hepatitis B virus and relatedness between Somali and Asian A1 sequences. *J Gen Virol* 2005; 86: 2163-2167.
112. **Olinger C.M.**, Venard V., Njayou M., Oyefolu A.O., Maiga I., Kemp A.J., Omilabu S.A., le Faou A., Muller C.P. Phylogenetic analysis of the precore/core gene of hepatitis B virus genotypes E and A in West Africa: new subtypes, mixed infections and recombinations. *J Gen Virol* 2006; 87: 1163-1173.
113. **Andernach I.E.**, Nolte C., Pape J.W., Muller C.P. Slave trade and hepatitis B virus genotypes and subgenotypes in Haiti and Africa. *Emerg Infect Dis* 2009; 15: 1222-1228.
114. **Hubschen J.M.**, Mbah P.O., Forbi J.C., Otegbayo J.A., Olinger C.M., Charpentier E., Muller C.P. Detection of a new subgenotype of hepatitis B virus genotype A in Cameroon but not in neighbouring Nigeria. *Clin Microbiol Infect* 2011; 17(1): 88-94.
115. **Pourkarim M.R.**, Lemey P., Amini-Bavil-Olyae S. et al. Novel hepatitis B virus subgenotype A6 in African-Belgian patients. *J Clin Virol* 2010; 47(1): 93-6.
116. **Sugauchi F.**, Orito E., Ichida T., Kato H., Sakugawa H., Kakumu S., Ishida T., Chutaputti A., Lai C.L., Ueda R., Miyakawa Y., Mizokami M. Hepatitis B virus of genotype B with or without recombination with genotype C over the precore region plus the core gene. *J Virol* 2002; 76: 5985-5992.
117. **Norder H.**, Couroucé A.M., Coursaget P., Echevarria J.M., Leef S.D., Mushahwar I.K., Robertson B.H., Locarnini S., Magnius L.O. Genetic diversity of hepatitis B virus strains derived worldwide: genotypes, subgenotypes, and HBsAg subtypes. *Intervirology* 2004; 47: 289-309.
118. **Sakamoto T.**, Tanaka Y., Orito E., Co J., Clavio J., Sugauchi F., Ito K., Ozasa A., Quino A., Ueda R., Sollano J., Mizokami M. Novel subtypes (subgenotypes) of hepatitis B virus genotypes B and C among chronic liver disease patients in the Philippine. *J Gen Virol* 2006; 87: 1873-1882.

119. **McMahon B.J.** The influence of hepatitis B virus genotype and subgenotype on the natural history of chronic hepatitis B. *Hepatol Int* 2009; 3(2): 334-342.
120. **Osiowy C., Larke B. Giles E.** Distinct geographical and demographic distribution of hepatitis B virus genotypes in the Canadian Arctic as revealed through an extensive molecular epidemiological survey. *J Viral Hepat* 2010.
121. **Chan H.L., Tsui S.K., Tse C.H., Ng E.Y., Au T.C., Yuen L., Bartholomeusz A, Leung K.S., Lee K.H., Locarnini S., Sung J.J.** Epidemiological and virological characteristics of 2 subgroups of hepatitis B virus genotype C. *J Infect Dis* 2005; 191: 2022-2032.
122. **Tanaka Y., Orito E. Yuen M.F., Mukaide M., Sugauchi F., Ito K., Ozasa A., Sakamoto T., Kurbanov F., Lai C.L., Mizokami M.** Two subtypes (subgenotypes) of hepatitis B virus genotype C: A novel subtyping assay based on restriction fragment length polymorphism. *Hepatol Res* 2005; 33: 216-224.
123. **Sugauchi F., Mizokami M., Orito E., Ohno T., Kato H., Suzuki S., Kimura Y., Ueda R., Butterworth L.A., Cooksley W.G.** A novel variant genotype C of hepatitis B virus identified in isolates from Australian Aborigines: complete genome sequence and phylogenetic relatedness. *J Gen Virol* 2001; 82: 883-892.
124. **Cavinta L., Cao G., Schaefer S.** Description of a new hepatitis B virus C6 subgenotype found in the Papua province of Indonesia and suggested renaming of a tentative C6 subgenotype found in the Philippine as subgenotype C7. *J Clin Microbiol* 2009; 47: 3068-3069.
125. **Ciccozzi M., Ciccaglione A.R., Lo Presti A., Equestre M., Cella E., Ebranati E., Gabanelli E., Villano U., Bruni R., Yalcinkaya T., Tanzi E., Zehender G.** Evolutionary dynamics of HBV-D1 genotype epidemic in Turkey. *J Med Virol* 2014; 86(1): 109-16.
126. **Mohebbi S.R., Amini-Bavil-Olyae S., Zali N., Noorinayer B., Derakhshan F., Chiani M. et al.** Molecular epidemiology of hepatitis B virus in Iran. *Clin Microbiol Infect* 2008; 14: 858-866.
127. **Amini-Bavil-Olyae S., Alavian S.M., Adeli A., Sarrami-Forooshani R., Sabahi F., Sabouri E., Tavangar H.R., Azizi M., Mahboudi F.** Hepatitis B virus genotyping, core promoter, and

precore/core mutations among Afghan patients infected with hepatitis B: a preliminary report. *J Med Virol* 2006; 78: 358-364.

128. **Meldal B.H.**, Moula N.M., Barnes I.H., Boukef K., Allain J.P. A novel hepatitis B virus subgenotype D7 in Tunisian blood donors. *J Gen Virol* 2009; 90: 1622-1628.
129. **Lazarevic I.**, Cupic M., Delic D., Svirtlih N.S., Simonovic J., Jovanovic T. Distribution of HBV genotypes, subgenotypes and HBsAg subtypes among chronically infected patients in Serbia. *Arch Virol* 2007; 152(11): 2017-2025.
130. **Ezzikouri S.**, Chemin I., Chafik A., Wakrim L., Nourlil J. et al. Genotype determination in Moroccan hepatitis B chronic carriers. *Infect Genet Evol* 2008; 8: 306-312.
131. **Banerjee A.**, Kurbanov F., Datta S., Chandra P.K., Tanaka Y., Mizokami M., Chakravarty R. Phylogenetic relatedness and genetic diversity of hepatitis B virus isolates in Eastern India. *J Med Virol* 2006; 78: 1164-1174.
132. **Abdou Chekaraou M.**, Brichtler S., Mansour W. et al. A new hepatitis B virus subgenotype D (D8) strain, resulting from recombination between genotypes D and E, is circulating in Niger along with HBV/E strains. *J Gen Virol* 2010; 91: 1609-1620.
133. **Ghosh S.**, Banerjee P., Deny P. et al. New HBV subgenotype D9, a novel D/C recombinant, identified in patients with chronic HBeAg-negative infection in Eastern India. *J Viral Hepat* 2013; 20(3): 209-18.
134. **Kato H.**, Fujiwara K., Gish R.G., Sakugawa H., Yoshizawa H., Sugauchi F., Orito E., Ueda R., Tanaka Y., Kato T., Miyakawa Y., Mizokami M. Classifying genotype F of hepatitis B virus into F1 and F2 subtypes. *World J Gastroenterol* 2005; 11: 6295-6304.
135. **Norder H.**, Arauz-Ruiz P., Blitz L., Pujol F.H., Echevarria J.M., Magnius L.O. The T(1858) variant predisposing to the precore stop mutation correlates with one of two major genotype F hepatitis B virus clades. *J Gen Virol* 2003; 84: 2083-2087.
136. **Naumann H.**, Schaefer S., Yoshida C.F., Gaspar A.M., Repp R., Gerlich W.H. Identification of a new hepatitis B virus (HBV) genotype from Brasile that expresses HBV surface antigen subtype adw4. *J Gen Virol* 1993; 74 (Pt 8): 1627-1632.

137. **Huy T.T.**, Ushijima H., Sata T., Abe K. Genomic characterization of HBV genotype F in Bolivia: genotype F subgenotypes correlate with geographic distribution and T (1858) variant. *Arch Virol* 2006; 151: 589-597.
138. **Roman S.**, Panduro A. HBV endemicity in Mexico is associated with HBV genotypes H and G. *World J Gastroenterol* 2013; 19(33): 5446-53.
139. **Hussain M.**, Lok A.S.F. Mutations in the hepatitis B virus polymerase gene associated with antiviral treatment for hepatitis B. *J Viral Hepatitis* 1999; 6: 183-194.
140. **Locarnini S.** Molecular virology of hepatitis B. *Seminars in liver disease* 2004; 24 Suppl 1: 3-10.
141. **Kaneko S.**, Miller R.H. Heterogeneity of the core gene sequence in a patient chronically infected with hepatitis B virus. *J Infect Dis* 1989; 160: 903-904.
142. **Mimms L.** Hepatitis B virus escape mutants: "pushing the envelope" of chronic hepatitis B virus infection. *Hepatology* 1995; 21: 884-887.
143. **Seddigh-Tonekaboni S.**, Waters J.A., Jeffers F. et al. Effect of variation in the common "a" determinant on the antigenicity of hepatitis B surface antigen. *J Med Virol* 2000; 60: 113-121.
144. **Melegari M.**, Bruno S., Wands J.R. Properties of hepatitis B virus pre-S1 deletion mutants. *Virology* 1994; 199: 292-300.
145. **Casemann W.H.**, Meyer M., Kekulé A.S. et al. A trans-activator function is generated by integration of hepatitis B virus preS7S sequences in human hepatocellular carcinoma DNA. *Procl Natl Acad Sci USA* 1990; 87: 2970-2974.
146. **Ching-Lung L.**, Dienstag J., Schiff E., Nancy W.Y., Aktins M., Hunt C., Brown N., Woessner M., Boehme R., Condreay L. Prevalence and clinical correlates of YMDD variants during lamivudine therapy for patients with chronic hepatitis B. *Clin Inf Dis* 2003; 36: 687-696.
147. **Liu X.**, Schinazi R.F. Hepatitis B virus resistance to lamivudine and its clinical implications. *Antiv Chem Chemoter* 2002; 13: 143-155.
148. **Bartholomeusz A.**, Locarnini S. Mutations in the hepatitis B virus polymerase gene that are associated with resistance to famciclovir and lamivudine. *Int Antiviral News* 1997; 5: 123-124.

149. **Pichoud C.**, Seigneres B., Wang Z. et al. Transient selection of hepatitis B virus polymerase gene mutant associated with a decreased replication capacity and famciclovir resistance. *Hepatology* 1999; 29: 230-237.
150. **Alter J.M.** Epidemiology and prevention of hepatitis B. *Seminars in liver disease* 2003.
151. **Yang J.**, Xing K., Deng R., Wang J., Wang X. Identification of Hepatitis B virus putative intergenotype recombinants by using fragment typing. *J Gen Virol* 2006; 87: 2203-2215.
152. **Simmonds P.**, Midgley S. Recombination in the genesis and evolution of hepatitis B virus genotypes. *J Virol* 2005; 79: 15467-15476.
153. **Owiredu W.K.**, Kramvis A., Kew M.C. Hepatitis B virus DNA in serum of healthy black African adults positive for hepatitis B surface antibody alone: possible association with recombination between genotypes A and D. *J Med Virol* 2001; 64: 441-454.
154. **Cui C.**, Shi J., Hui L., Xi H., Zhuoma., Quni, Tsedan, Hu G. The dominant hepatitis B virus genotype identified in Tibet is a C/D hybrid. *J Gen Virol* 2002; 83: 2773-2777.
155. **Suwannakarn K.**, Tangkijvanich P., Theamboonlers A., Abe K., Poovorawan Y. A novel recombinant of hepatitis B virus genotypes G and C isolated from a Thai patient with hepatocellular carcinoma. *J Gen Virol* 2005; 86: 3027-3030.
156. **Bekondi C.**, Olinger C.M., Boua N., Talarmin A., Muller C.P., Le Faou A., Venard V. Repubblica Centrafricana is part of the West-African hepatitis B virus genotype E crescent. *J Clin Virol* 2007; 40: 31-37.
157. **Garmiri P.**, Loua A., Haba N., Candotti D., Allain J.P. Deletions and recombinations in the core region of hepatitis B virus genotype E strains from asymptomatic blood donors in Guinea, west Africa. *J Gen Virol* 2009; 90: 2442-2451.
158. **Laoi B.N.**, Crowley B. Molecular characterization of hepatitis B virus (HBV) isolates, including identification of a novel recombinant, in patients with acute HBV infection attending an Irish hospital. *J Med Virol* 2008; 80: 1554-1564.
159. **Hannoun C.**, Norder H., Lindh M. An aberrant genotype revealed in recombinant hepatitis B virus strains from Vietnam. *J Gen Virol* 2000; 81: 2267-2272.
160. **Magiorkinis E.N.**, Magiorkinis G.N., Paraskevis D.N., Hatzakis A.E. Re-analysis of a human hepatitis B (HBV) isolate from an East African wild born Pan troglodytes schweinfurthii:

evidence for interspecies recombination between HBV infecting chimpanzee and human. *Gene* 2005; 349: 165-171.

161. **Soubiran G.**, Le Bras M., Marini P., Sekou H. High HBsAg and anti-delta carrier rate among asymptomatic Africans living on the campus of the University of Niamey, Niger. *Trans R Soc Trop Med Hyg* 1987; 81: 998-1000.
162. **Cenac A.**, Pedroso M.L., Djibo A., Develoux M., Pichoud C., Lamothe F., Trepo C., Warter A. Hepatitis B, C, and D virus infections in patients with chronic hepatitis, cirrhosis, and hepatocellular carcinoma: a comparative study in Niger. *Am J Trop Med Hyg* 1995; 52: 293-296.
163. **Mamadou S.**, Laouel Kader A., Rabiou S., Aboubacar A., Soumana O., Garba A., Delaporte E., Mboup S. Prevalence of the HIV infection and five other sexually-transmitted infections among sex workers in Niamey, Niger. *Bull Soc Pathol Exot* 2006; 99: 19-22.
164. **Khelifa F.**, Thibault V. Characteristics of hepatitis B viral strains in chronic carrier patients from north-east Algeria. *Pathol Biol (Paris)* 2009; 57: 107-113.
165. **Hubschen J.M.**, Andernach I.E., Muller C.P. Hepatitis B virus genotype E variability in Africa. *J Clin Virol* 2008; 43: 376-380.
166. **Kramvis A.**, Kew M.C. Epidemiology of hepatitis B virus in Africa, its genotypes and clinical associations of genotypes. *Hepatol Res* 2007; 37: S9-S19.
167. **Lin C.L.**, Kao J.H. The clinical implications of hepatitis B virus genotype: Recent advances. *J Gastroenterol Hepatol* 2011; 26 Suppl 1:123-130.
168. **Yang H.I.**, Yeh S.H., Chen P.J. et al. REVEAL-HBV Study Group. Associations between hepatitis B virus genotype and mutants and the risk of hepatocellular carcinoma. *J Natl Cancer Inst* 2008; 100: 1134-1143.
169. **Kao J.H.**, Chen P.J., Lai M.Y. et al. Hepatitis B genotypes correlate with clinical outcomes in patients with chronic hepatitis B. *Gastroenterology* 2000; 118(3): 554-559.
170. **Ni Y.H.**, Chang M.H., Wang K.J. et al. Clinical relevance of hepatitis B virus genotype in children with chronic infection and hepatocellular carcinoma. *Gastroenterology* 2004, 127: 1733-1738.

171. **Yin J.**, Zhang H., Li C. et al. Role of hepatitis B virus genotype mixture, subgenotypes C2 and B2 on hepatocellular carcinoma: compared with chronic hepatitis B and asymptomatic carrier state in the same area. *Carcinogenesis* 2008; 29: 1685-1691.
172. **Chen D.S.** From hepatitis to hepatoma: lessons from type B viral hepatitis. *Science* 1993; 262: 369-370.
173. **Chu C.J.**, Hussain M., Lok A.S. Hepatitis B virus genotype B is associated with earlier HBeAg seroconversion compared with hepatitis B virus genotype C. *Gastroenterology* 2002; 122(7): 1756-1762.
174. **Kao J.H.**, Chen P.J., Lai M.Y., Chen D.S. Hepatitis B virus genotypes and spontaneous hepatitis B e antigen seroconversion in Taiwanese hepatitis B carriers. *J Med Virol* 2004; 72: 363-369
175. **Sanchez-Tapias J.M.**, Costa J., Mas A. et al. Influence of hepatitis B virus genotype on the long-term outcome of chronic hepatitis B in western patients. *Gastroenterology* 2002; 123(6):1848-1856.
176. **Thakur V.**, Guptan R.C., Kazim S.N. et al. Profile, spectrum and significance of HBV genotypes in chronic liver disease patients in the Indian subcontinent. *J Gastroenterol Hepatol* 2002; 17: 165-170.
177. **Yuasa R.**, Takahashi K., Dien B.V. et al. Properties of hepatitis B virus genome recovered from Vietnamese patients with fulminant hepatitis in comparison with those of acute hepatitis. *J Med Virol* 2000; 61: 23-28.
178. **Fung S.K.**, Lok A.S. Hepatitis B virus genotypes: do they play a role in the outcome of HBV infection? *Hepatology* 2004; 40(4): 790-792.
179. **Casey J.L.**, Niro G.A., Engle R.E. et al. Hepatitis B virus (HBV)/hepatitis D virus (HDV) coinfection in outbreaks of acute hepatitis in the Peruvian Amazon Basin: the roles of HDV genotype III and HBV genotype F. *J Infect Dis* 1996; 174: 920-926.
180. **Flodgren E.**, Bengtsson S., Knutsson M. et al. Recent high incidence of fulminant hepatitis in Samara, Russia: molecular analysis of prevailing hepatitis B and D virus strains. *J Clin Microbiol* 2000; 38: 3311-3316.

181. **Kao J.H.**, Wu N.H, Chen P.J, et al. Hepatitis B genotypes and the response to interferon therapy. *J Hepatol* 2000; 33(6): 998-1002.
182. **Lin C.L.**, Kao J.H. Hepatitis B viral factors and treatment responses in chronic hepatitis B. *J Formos Med Assoc* 2013; 112(6):302-11.
183. **Erhardt A.**, Göbel T., Ludwig A., Lau G.K., Marcellin P., van Bömmel F., Heinzl-Pleines U., Adams O., Häussinger D. Response to antiviral treatment in patients infected with hepatitis B virus genotypes E-H. *J Med Virol.* 2009; 81(10): 1716-1720.
184. **Chien R.N.**, Yeh C.T., Tsai S.L., Chu C.M., Liaw Y.F. Determinants for sustained HBeAg response to lamivudine therapy. *Hepatology* 2003; 38(5): 1267-1273.
185. **Yuen M.F.**, Wong D.K., Sablon E., Yuan H.J., Sum S.M., Hui C.K., Chan A.O., Wang B.C., Lai C.L. Hepatitis B virus genotypes B and C do not affect the antiviral response to lamivudine. *Antivir Ther* 2003; 8(6): 531-534.
186. **Chan H.L.**, Wong M.L., Hui A.Y., Chim A.M., Tse A.M., Hung L.C., Chan F.K., Sung J.J. Hepatitis B virus genotype has no impact on hepatitis B e antigen seroconversion after lamivudine treatment. *World J Gastroenterol* 2003; 9(12): 2695-2697.
187. **Kao J.H.**, Liu C.J., Chen D.S. Hepatitis B viral genotypes and lamivudine resistance. *J Hepatol* 2002; 36(2): 303-304.
188. **Buti M.**, Cotrina M., Valdes A. et al. Is hepatitis B subtype testing useful in predicting virological response and resistance to lamivudine? *J Hepatol* 2002; 36(3): 445-446.
189. **Westland C.**, Delaney W 4th, Yang H. et al. Hepatitis B virus genotypes and virologic response in 694 patients in phase III studies of adefovir dipivoxil. *Gastroenterology* 2003; 125(1): 107-116.
190. **Lurie Y.**, Manns M.P., Gish R.G. et al. The efficacy of entecavir is similar regardless of disease-related baseline subgroups in treatment of nucleoside-naïve, HBeAg (+) and HBeAg (-) patients with chronic hepatitis B. *J Hepatol* 2005; 42 (Suppl. 2): 184.
191. **Hou J.**, Yin Y.K., Xu D. et al. Telbivudine versus lamivudine in Chinese patients with chronic hepatitis B: Results at 1 year of a randomized, double-blind trial. *Hepatology* 2008; 47(2):447-454.

192. **Hsieh T.H.**, Tseng T.C., Liu C.J., Lai M.Y., Chen P.J., Hsieh H.L., Chen D.S., Kao J.H. Hepatitis B virus genotype B has an earlier emergence of lamivudine resistance than genotype C. *Antivir Ther* 2009; 14(8):1157-1163.
193. **Solomon R.**, Van Raden M., Kaslow R. et al. Association of hepatitis B surface antigen and core antibody with acquisition and manifestations of the acquired immunodeficiency syndrome. *Am J Public Health* 1990; 80: 1475-1478.
194. **Laure F.**, Zagury D., Saimont A. et al. Hepatitis B virus DNA sequences in lymphoid cells from patients with AIDS and AIDS-related complex. *Science* 1985; 229: 561-563.
195. **Gomez-Gonzalo M.**, Carretero M., Ruas J. et al. The hepatitis B virus X protein induces HIV-1 replication and transcription in synergy with T-cell activation signals: functional roles of NF-kappaB/NF-AT and SP1-binding sites in the HIV-1 long terminal repeat promoter. *J Biol Chem* 2001; 276(38): 35435-35443.
196. **Eskil A.**, Magnus P., Peterson G. et al. Hepatitis B antibodies in HIV-infected homosexual men are associated with more rapid progression to AIDS. *AIDS* 1992; 6: 571-574.
197. **Sinicco A.**, Raiteri R., Sciandra M. et al. Co-infected and superinfection of hepatitis B virus in patients infected with immunodeficiency virus: no evidence of faster progression to AIDS. *Scand J Infect Dis* 1997; 29: 111-115.
198. **Rockstroh J.K.** Management of hepatitis B and C in HIV co-infected patients. *J AIDS* 2003; 34: S59-65.
199. **Nunez M.**, Puoti M., Camino N. et al. Treatment of chronic hepatitis B in the human immunodeficiency virus-infected patient: present and future. *CID* 2003; 37: 1678-1685.
200. **Thio C.L.**, Seaberg E.C., Skolasky R., Phair J. et al. HIV-1, hepatitis B virus, and risk of liver-related mortality in the Multicenter Cohort Study (MACS). *Lancet* 2002; 360: 1921-1926.
201. **Carr A.**, Cooper D.A. Restoration of immunity to chronic hepatitis B infection in HIV-infected patients on protease inhibitor. *Lancet* 1997; 349: 995-996.
202. **Fasano M.**, Saracino A., Carosi G. et al. Hepatitis B and immigrants: a SIMIT multicenter cross-sectional study. *Infection* 2013; 41(1):53-9.
203. **Gaeta G.B.**, Cuomo G. Epidemiologia delle epatiti virali in Italia. *Ligand Assay* 2006. 11 (4): 314-318.

204. **Stroffolini T.** The changing pattern of hepatitis B virus infection over the past three decades in Italy. *Dig. Liver Dis* 2005. 37(8): 622-627.
205. **Simi** Home Page <http://www.simi.iss.it/seieva.htm>
206. **Stroffolini T.**, Mele A., Tosti ME. et al. The impact of hepatitis B mass immunisation campaign on the incidence and risk factors of acute hepatitis B in Italy. *J Hepatol* 2000; 33(6): 980-985.
207. **Shouval D.** Is universal vaccination against hepatitis B sufficient for control of HBV infection? Lessons from the immunization campaign in Italy. *J Hepatol* 2000; 33(6): 1009-1011.
208. **Zanetti A.R.**, Tanzi E., Romanò L., Grappassonni I. Vaccination against hepatitis B: the Italian strategy. *Vaccine* 1993; 11: 521-524.
209. **Struffolini T.**, Cialdea L., Tosti.E. et al. Vaccination campaign against hepatitis B for 12-year-old subject in Italy. *Vaccine* 1997; 5: 583-585.
210. **Mele A.**, Mariano A., Tosti M.E. et al. Acute hepatitis delta virus infection in Italy: incidence and risk factors after the introduction of the universal anti-hepatitis B vaccination campaign. *Clin Infect Dis* 2007; 44: 17-24.
211. **Raimondo G.**, Isgrò G., Caccamo G., Pollicino T., Squadrito G. Calabrian HBV Study Group. Is there a downgrading in the alert about the hepatitis B virus infection in Italy? *Dig Liver Dis* 2007; 39(3): 257-261.
212. **Scalia G.**, Russo R., Palermo I., Gussio M., Bellissimo F., Fatuzzo F., Montineri A., Patamia I., Costanzo M., Zappala D., Russo R., Cacopardo B. Clinical, virological and epidemiological aspects of acute B hepatitis in the Catania area (Italy). *Infez Med* 2010; 18(3): 169-174.
213. **Grenfell B.T.**, Pybus O.G., Gog J.R. et al. Unifying the epidemiological and evolutionary dynamics of pathogens. *Science* 2004; 303(5656): 327-332 21.
214. **Pybus O.G.**, Rambaut A. Evolutionary analysis of the dynamics of viral infectious disease. *Nat Rev Genet* 2009; 10(8): 540-550.
215. **Rutschmann F.** Molecular dating of phylogenetic trees: a brief review of current methods that calculate divergence times. *Divers Distrib* 2006; 12: 35-48.

216. **Drummond A.J.**, Rambaut A., Shapiro B., Pybus O.G. Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol Biol Evol* 2005; 22(5): 1185-1192.
217. **Holmes E.C.** Evolutionary history and phylogeography of human viruses. *Annu Rev Microbiol* 2008; 62: 307-328.
218. **Lemey P.**, Rambaut A., Drummond A.J., Suchard M.A. Bayesian phylogeography finds its roots. *PLoS Comput Biol* 2009; 5(9): e1000520.
219. **Murhekar M.V.**, Murhekar K.M., Sehgal S.C. Epidemiology of hepatitis B virus infection among the tribes of Andaman and Nicobar Islands, India. *Trans R Soc Trop Med Hyg.* 2008; 102:729-734.
220. **Bozdayi G.**, Türkyilmaz A.R., Idilman R., Karatayli E., Rota S., Yurdaydin C., Bozdayi A.M. Complete genome sequence and phylogenetic analysis of hepatitis B virus isolated from Turkish patients with chronic HBV infection. *J Med Virol* 2005; 76(4): 476-481.
221. **Garmiri P.**, Rezvan H., Abolghasemi H., Allain J.P. Full genome characterization of hepatitis B virus strains from blood donors in Iran. *J Med Virol* 2011; 83:948-952.
222. **Tallo T.**, Tefanova V., Priimägi L., et al. D2: major subgenotype of hepatitis B virus in Russia and the Baltic region. *J Gen Virol* 2008; 89:1829-1839.
223. **Zehender G.**, Shkjezi R., Ebranati E., Gabanelli E., et al. Reconstruction of the epidemic history of hepatitis B virus genotype D in Albania. *Infect Genet Evol* 2012; 12(2):291-8.
224. **De Maddalena C.**, Giambelli C., Tanzi E., Colzani D., Schiavini M., Milazzo L., Bernini F., Ebranati E., Cargnel A., Bruno R., Galli M., Zehender G. High level of genetic heterogeneity in S and P genes of genotype D hepatitis B virus. *Virology* 2007; 365(1): 113-124.
225. **Kitab B.**, El Feydi A.E., Afifi R., et al. Hepatitis B genotypes/subgenotypes and MHR variants among Moroccan chronic carriers. *J Infect* 2011; 63:66-75.
226. **Michitaka K.**, Tanaka Y., Horiike N., et al. Tracing the history of hepatitis B virus genotype D in western Giappone. *J Med Virol* 2006; 78:44-52.
227. **Sugauchi F.**, Kumada H., Acharya S.A. et al. Epidemiological and sequence differences between two subtypes (Ae and Aa) of hepatitis B virus genotype A. *The Journal of general virology* 2004; 85(Pt 4): 811-820.

228. **Makuwa M.**, Souquiere S., Telfer P. et al. Identification of hepatitis B virus subgenotype A3 in rural Gabon. *J Med Virol* 2006; 78(9): 1175-1184.
229. **Pourkarim M.R.**, Amini-Bavil-Olyae S., Lemey P., Maes P., Van Ranst M. HBV subgenotype misclassification expands quasi-subgenotype A3. *Clin Microbiol Infect* 2011; 17(6): 947-949.
230. **Van Steenbergen J.**, Niesters H., Op de Coul E. et al. Molecular epidemiology of hepatitis B virus in Amsterdam 1992–1997. *J Med Virol* 2002; 66(2): 159-165.
231. **van Houdt R.**, Bruisten S.M., Geskus R.B. et al. Ongoing transmission of a single hepatitis B virus strain among men having sex with men in Amsterdam. *J Viral Hepat* 2010; 17(2): 108-114.
232. **van Ballegooijen W.M.**, van Houdt R., Bruisten S.M. et al. Molecular sequence data of hepatitis B virus and genetic diversity after vaccination. *Am J Epidemiol* 2009; 170(12): 1455-1463.
233. **Hahne S.**, van Houdt R., Koedijk F. et al. Selective hepatitis B virus vaccination has reduced hepatitis B virus transmission in the Netherlands. *PloS one* 2013; 8(7): e67866.
234. **Fujisaki S.**, Yokomaku Y., Shiino T. et al. Outbreak of infections by hepatitis B virus genotype A and transmission of genetic drug resistance in patients coinfecting with HIV-1 in Giappone. *J Clin Microbiol* 2011; 49(3): 1017-1024.
235. **Kramvis A.**, Paraskevis D. Subgenotype A1 of HBV--tracing human migrations in and out of Africa. *Antivir Ther* 2013; 18(3 Pt B): 513-521.
236. **Thompson J.D.** et al. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994; 22: 4673-4680.
237. **Felsenstein J.** Phylogeny inference package (version 3.2). *Cladistics* 1989; 5: 164-166.
238. **Tamura K.**, Dudley J., Nei M., Kumar S. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 2007; 24(8): 1596-1599.
239. **Posada D.**, Crandall K.A. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 1998; 14: 817-818.

240. **Guindon S.**, Lethiec F., Duroux P., Gascuel O. PHYML Online--a web server for fast maximum likelihood-based phylogenetic inference.. *Nucleic Acids Res* 2005; 33(Web Server issue): W557-9.
241. **Strimmer K.**, Von Haeseler A. Likelihood-mapping: a simple method to visualize phylogenetic content of a sequence alignment. *Proc Natl Acad Sci USA* 1997; 94(13): 6815-6819.
242. **Drummond A.J.**, Rambaut A. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* 2007; 7: 214-222.
243. **Kass R.E.**, Raftery A.E. Bayes factor. *J Am Stat Assoc* 1995; 90: 773-795.
244. **Lanave C.**, Preparata G., Saccone C., Serio G. A new method for calculating evolutionary substitution rates. *J Mol Evol* 1984; 20(1): 86-93.
245. **MrBayes** Home Page <http://mrbayes.csit.fsu.edu/>
246. **Huelsenbeck J.P.**, Ronquist F. et al. Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* 2001; 294: 2310-2314.
247. **Hastings W.K.** Monte Carlo sampling methods using Markov chains and their application. *Biometrika* 1970; 57: 97-109.
248. **Spread** Home Page <http://www.kuleuven.be/aidslab/phylogeography/SPREAD.html>
249. **Deterding K.**, Constantinescu I., Nedelcu F.D., et al. Prevalence of HBV genotypes in Central and Eastern Europe. *J Med Virol* 2008; 80(10): 1707-1711.
250. **Asimina F.**, Dimitrios P., Maria D., Mathaios M., Reiberger T., Nicolaos M. Molecular epidemiology of chronic hepatitis B virus infection in Greece. *J Med Virol* 2011; 83(2): 245-252.
251. **Echevarría J.M.**, Avellón A., Magnius L.O. Molecular epidemiology of hepatitis B virus in Spain: identification of viral genotypes and prediction of antigenic subtypes by limited sequencing. *J Med Virol* 2005; 76(2): 176-184.
252. **Servant-Delmas A.**, Mercier M., El Ghouzzi M.H. et al. National survey of hepatitis B virus (HBV) polymorphism in asymptomatic HBV blood donors from 1999 to 2007 in Francia. *Transfusion* 2010; 50(12): 2607-2618.

253. **Chandra P.K.**, Biswas A., Datta S., Banerjee A. et al. Subgenotypes of hepatitis B virus genotype D (D1, D2, D3 and D5) in India: differential pattern of mutations, liver injury and occult HBV infection. *J Viral Hepat* 2009; 16(10): 749-756.
254. **Paraskevis D.**, Magiorkinis G., Magiorkinis E. et al. Dating the origin and dispersal of hepatitis B virus infection in humans and primates. *Hepatology* 2013; 57, 908-916.
255. **Zehender G.**, Ebranati E., Gabanelli E. et al.. Enigmatic origin of hepatitis B virus: an ancient travelling companion or a recent encounter? *World J Gastroenterol* 2014; 20, 7622-7634.
256. **Godoy B.A.**, Alvarado-Mora M.V., Gomes-Gouvêa M.S. et al. Origin of HBV and its arrival in the Americas--the importance of natural selection on time estimates. *Antivir Ther* 2013; 18, 505-512.
257. **Araujo N.M.**, Waizbort R., Kay A. Hepatitis B virus infection from an evolutionary point of view: how viral, host, and environmental factors shape genotypes and subgenotypes. *Infect Genet Evol* 2011; 11, 1199-1207.
258. **Geggus D.** The French slave trade: an overview. *The William and Mary Quarterly* 2001; 119-138.
259. **Pakendorf B.**, Bostoen K., De Filippo C. Molecular perspectives on the Bantu expansion: a synthesis. *Language Dynamics and Change* 2011; 1, 50-88.
260. **Oliver R.A.**, Atmore A. Africa since 1800. *Cambridge University Press*, 1994.
261. **Chami F.A.** A review of Swahili archaeology. *African Archaeological Review* 1998; 15, 199-218.
262. **Vernet T.** Slave trade and slavery on the Swahili coast (1500-1750). *Slavery, Islam and Diaspora* 2009; 37-76.
263. **Gauniyal M.**, Chahal S.M., Kshatriya G.K. Genetic affinities of the Siddis of South India: an emigrant population of East Africa. *Human biology* 2008; 251-270.
264. **Vernet T.** East Africa: slave migrations. *The Encyclopedia of Global Human Migration*, 2013.

265. **Devesa M.**, Pujol F.H. Hepatitis B virus genetic diversity in Latin America. *Virus Res* 2007; 127, 177-184.
266. **Lago B.V.**, Mello F.C., Kramvis A. et al. Hepatitis B virus subgenotype A1: evolutionary relationships between Brazilian, African and Asian isolates. *PLoS One* 2014; 9(8): e105317.
267. **Forbi J.C.**, Vaughan G., Purdy M.A. et al. Epidemic history and evolutionary dynamics of hepatitis B virus infection in two remote communities in rural Nigeria. *PLoS One* 2010; 5(7): e11615.
268. **Rantala M.**, van de Laar M.J. Surveillance and epidemiology of hepatitis B and C in Europe - a review. *Euro Surveill* 2008; 13(21).
269. **Tamada Y.**, Yatsunami H , Masaki N. et al. Hepatitis B virus strains of subgenotype A2 with an identical sequence spreading rapidly from the capital region to all over Japan in patients with acute hepatitis B. *Gut* 2012; 61, 765-773.