Molecular Epidemiology of HBV Genotypes Circulating In Acute Hepatitis B Patients In The Campania Region

Caterina Sagnelli [1], Massimo Ciccozzi [2,3], Mariantonietta Pisaturo [4], Gianguglielmo Zehender [5], Alessandra Lo Presti [2], Loredana Alessio [4], Mario Starace [4], Domenica Lovero [2], Evangelista Sagnelli [4,6], Nicola Coppola [4]

1 :Department of Clinical and Experimental Medicine and Surgery "F. Magrassi e A. Lanzara", Second University of Naples, Naples, Italy
2: Epidemiology Unit, Department of Infectious, Parasite and Immune-Mediated Diseases, Italian Institute of Health, Rome, Italy.
3: University of Biomedical Campus, Rome, Italy.
4: Department of Mental Health and Public Medicine, Section of Infectious Diseases, Second University of Naples, Naples, Italy.
5: Department of Biomedical and Clinical Research, Infectious Diseases Chair, University of Milan, Milan, Italy.
6: Infectious and Tropical Diseases Unit, A.O.R.N. Sant'Anna e San Sebastiano, Caserta, Italy.

Background

- **The incidence of acute hepatitis B (AVH-B): decline in Italy (12/100,000 inhabitants in 1985 to 1/100,000 in 2011).**
- **The incidence rate per 100,000 inhabitants in 2011 was 0.0% in the age group 0-14, 0.5% in 15-24, 1.2% in 25-34 and 2% in those over 35 years.**
- **The majority of Italian patients with AVH-B are adults over 35 and show symptomatic, seldom severe, clinical course, in some cases progresses to fulminant liver failure [Stroffolini 2008; Sagnelli 2009].**

- **Few data are available on epidemiological impact of HBV subgenotypes in Italy, D3 associated with percutaneous transmission and A2 with sexual exposure [Zehender 2008].**

- **Italy has become a land of immigration (Africa, eastern Europe and eastern Asia) and consequently new HBV genotypes have been introduced [Coppola 2010].**

Methods

125 AVH-B consecutive pts observed from September 1999 to July 2008 [Coppola 2010]

53* pts

Increase in AST and ALT of at least 10 times UNL, HBsAg/ HBc-Ag IgM= pos/pos; Anti-HDV/ anti-HIV= neg/neg. HBV polymerase region evaluated by direct sequencing in plasma samples obtained at first observation.

HBV data sets: Two different sequence data sets were analyzed.

The 1st data set: 128 HBV polymerase gene sequences, of which 53 were isolates and 75 were genotype and subgenotype specific reference sequences (National Centre for Biotechnology Information).

The 2nd data set: only HBV-D3 subgenotype isolates and used to estimate the evolutionary rate and time of the most recent common ancestor (tMRCA).

***included in the present study based on the availability of plasma sample stored at -40°C**

Phylogenetic analysis: data sets were aligned (CLUSTAL X software). ModelTest program v 3.7 was used to select simplest evolutionary model that adequately fitted sequence data. A maximum likelihood (ML) phylogenetic tree was inferred with PhyML program using GTR + I + G nucleotide substitution model for 1st data set.

Evolutionary rate estimate, time-scaled phylogeny and phylodynamics: To estimate evolutionary rate on 2nd data set both strict and relaxed clock with uncorrelated log normal rate distribution were compared. A Bayesian Markov chain Monte Carlo (MCMC) method, BEAST software 1.7.4 was used.

Selection pressure analysis: dN/dS rate (ω; non-synonymous to synonymous) by ML approach implemented in HyPhy using 2nd data set.

Viral gene flow analysis: MacClade vers. 4 program to test viral gene out/in flow among HBV-D3 (2nd data set) with different risk factors (modified Slatkin and Maddison test). The viral gene flow among different risk factors was traced with State changes and stasis tool (MacClade software), when multiple MPRs were present, the algorithm calculated the average migration count over all possible MPRs for each pair.

The demographics and clinical, biochemical and virological characteristics at baseline

Number of patients	53
Males, N. (%)	40 (75.6)
Age, years, M ± SD	
Risk factor, No. (%):	
Unsafe sexual intercourse	
Intravenous drug use (IDU)	
Surgery, endoscopy, dental care	
Tattoo/Acupuncture	
Unknown	
Missing	
HBV genotype, No. (%):	
D (D1; D2; D3; D4)	34 (64.1) (7; 2; 24; 1)
A (A2)	14 (26.4) (14)
F (F1)	2 (3.8) (2)
E	3 (5.7)
Born in Italy, No. (%):	44 (83.02)

17% in other countries:
1=Ghana,
1=Dom. Rep.,
1= Algeria,
1=Tunisia,
1= Nigeria,
2 =Albania ,
2=Russia

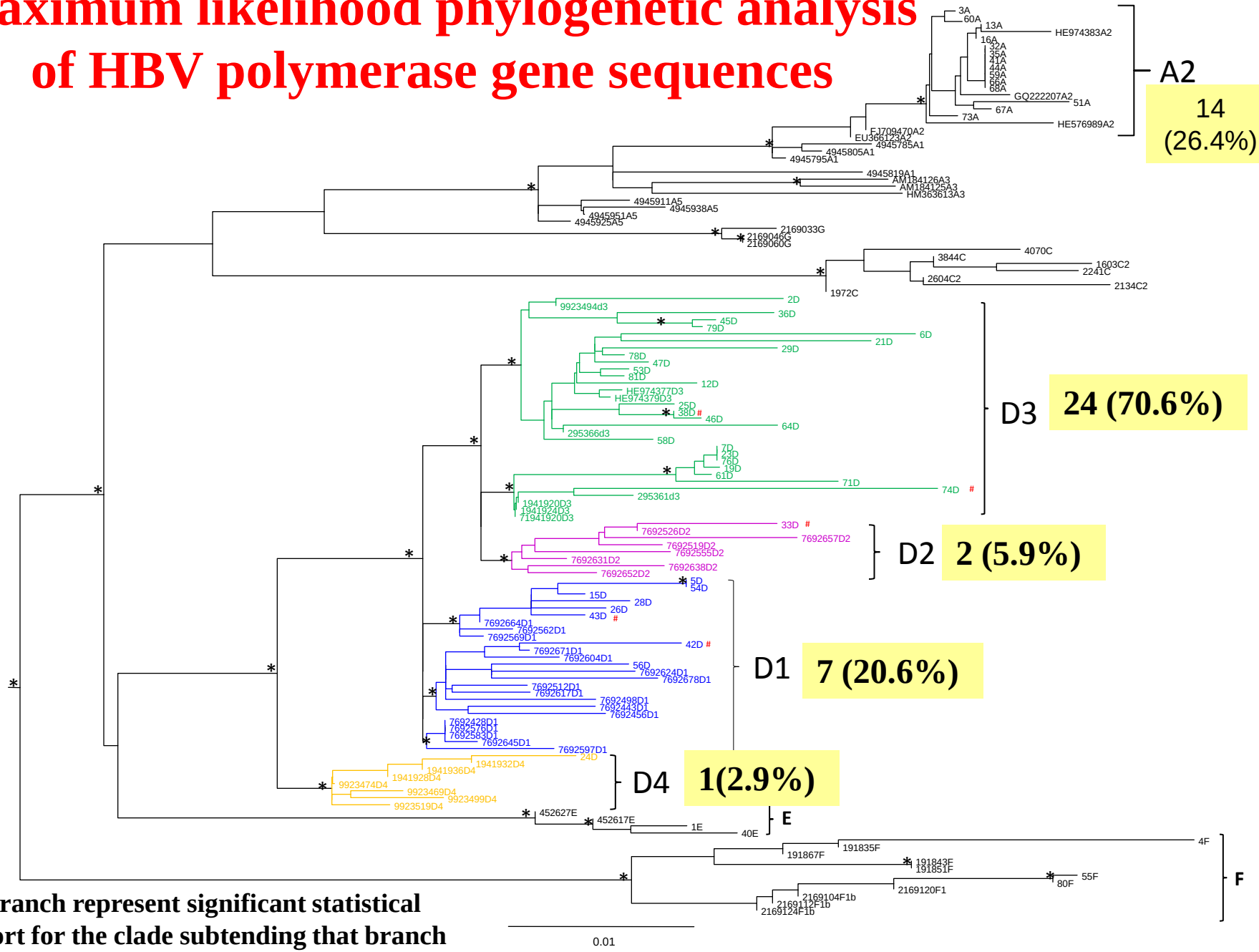
The demographics and clinical, biochemical and virological characteristics at baseline

Number of patients	53
Days from onset of symptoms, M ± SD	32.26 69.14
with severe acute hepatitis B, No. (%):	5 (9.4)
AST (IU/ml), M ± SD	2000.32 1229.37
ALT (IU/ml), M ± SD	2768.40 1193.57
	18.58 10.11
	3.50 (20.1-113)
	225.35 82.70
	4 E 7 2.9 E 10
Anti-HCV, No. (%): Negative Positive	 47 (88.7) 6 (11.3)
Clinical outcome, No. (%): Chronic viral hepatitis B Resolution OLT Missing	 1 (1.9) 49 (96.1) 1 (1.9) 2
Seroconversion to anti-HBsAg in 49 pts with resolution, No. (%): Yes No	 31 (63.3) 18 (36.7)

6 (11.5%) anti-HCV/HCV RNA-positive for at least 12 months at time of HBV superinfection,
All 6 pts became HCV RNA-negative and remained so throughout the clinical course of AVH-B.

All 53 patients enrolled were followed up for at least 6 months.

Maximum likelihood phylogenetic analysis of HBV polymerase gene sequences



***: a branch represent significant statistical support for the clade subtending that branch (bootstrap support >70%). #: foreign patients.**

Demographic, epidemiological, clinical and virological characteristics of 14 patients with HBV-A and 34 with HBV-D

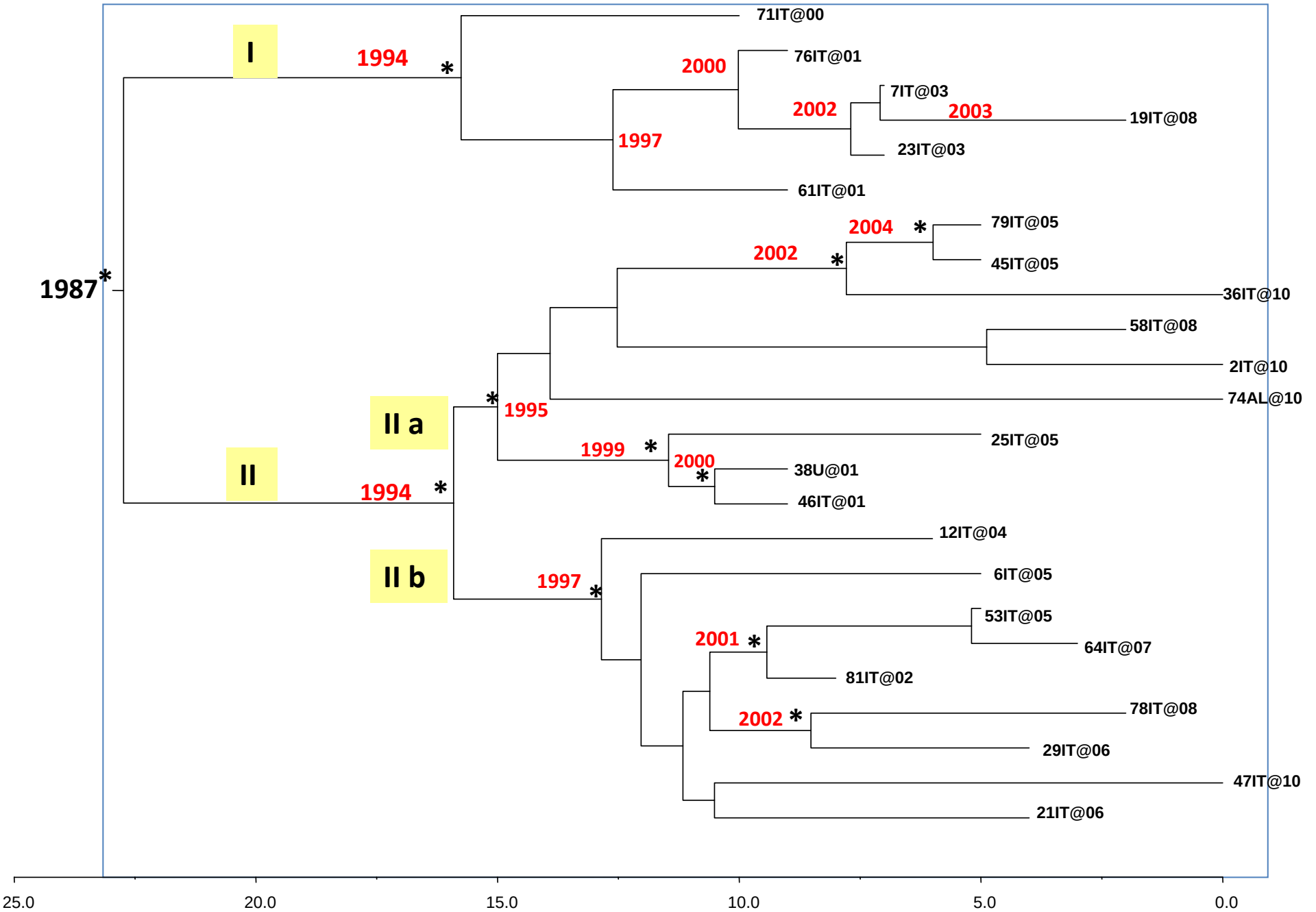
	HBV -A	HBV-D	*p*-value
Number of patients	14	34	
Males, No. (%):	13 (92.9)	23 (67.6)	0.06
Age, years, M ± SD	36.00 10.78	37.50 15.75	ns
Risk factor, No. (%):			
Unsafe sexual intercourse	4 (36.4)	10 (45.5)	ns
Intravenous drug use	1 (9.1)	9 (40.9)	0.064
Surgery, endoscopy, dental care	6 (54.5)	0	0.0001
Tattoo/Acupuncture	0	3 (13.6)	--
Unknown	3	12	--
Born in Italy, No.(%)	13 (92.9)	29 (85.3)	ns

Demographic, epidemiological, clinical and virological characteristics of 14 patients with HBV-A and 34 with HBV-D

	HBV -A	HBV-D	*p*-value
Number of patients	14	34	
Days from onset of symptoms, M ± SD	43.57 101.3	19.6 8.4	ns
with severe acute hepatitis B , No. (%)	0	4 (11.8)	ns
AST (IU/ml), M ± SD	1694 2492	2144 1389	ns
ALT (IU/ml), M ± SD	830.5 946.4	285.1 1207.1	ns
Bilirubin (mg/dl), M ± SD	17.15 6.67	18.58 10.1	ns
Prothrombin time (%), median (range)	68 (61.0-113)	60.5 (20.1-77.0)	ns
PTL x10^3/µL		9 90.9	ns
HBV DNA (IU/ml), M		7 1.8E8	ns
Anti-HCV, No. (%): Negative Positive	13 (92.9) 1 (7.1)	29 (85.29) 5 (14.70)	ns
Clinical outcome, No. (%): Chronic viral hepatitis B Resolution OLT Missing	0 14 (100) 0 0	1 (3.03) 31 (93.9) 1 (3.03) 1	ns
Seroconversion to anti-HBsAg in 45 pts with resolution, No. (%): Yes No	8 (57.1) 6 (42.9)	19 (61.3) 12 (38.7)	ns

All 5 patients with severe acute hepatitis showed subgenotype D3

Bayesian time-scaled tree of 24 HBV-D3 polymerase sequences

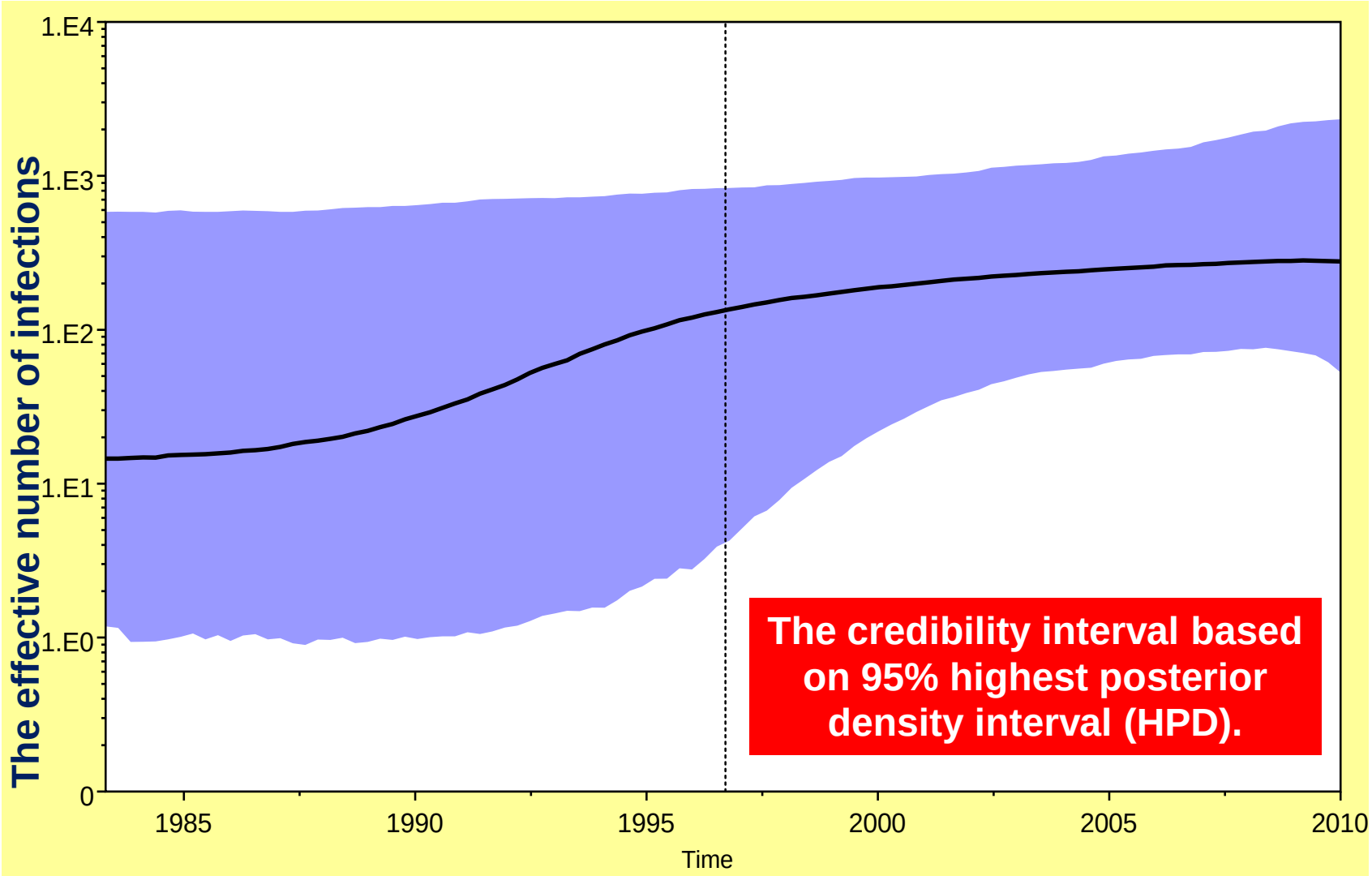


The reconstruction of the time-scaled Bayesian tree of the 24 HBV-D3 isolates

The estimation of the time of the tree's root gave a mean value of 23 years ago, corresponding to 1987 (95% HPD: 1870-1997).

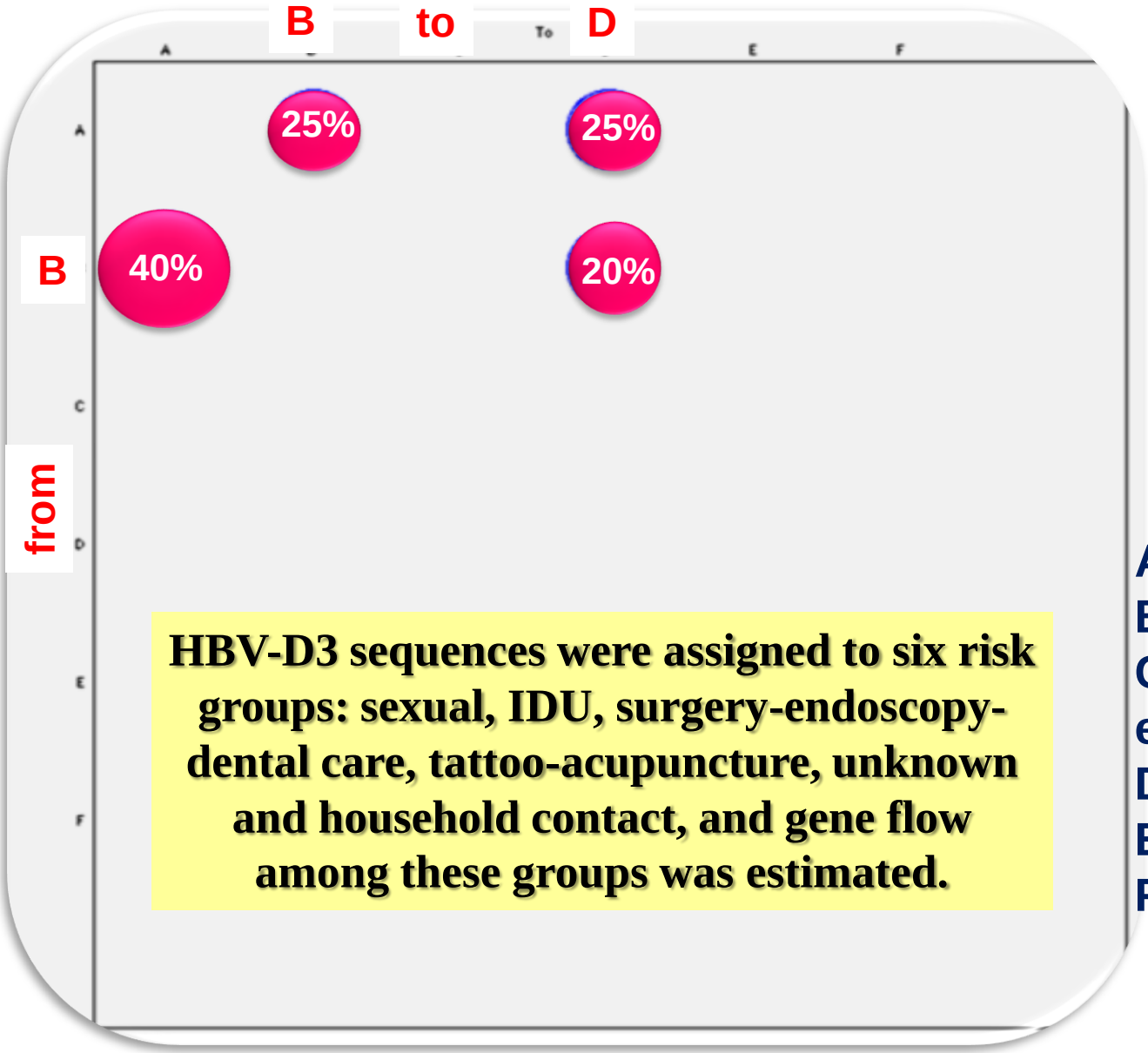
The Bayesian tree showed 2 main clades. Clade I included total of 6 isolates had tMRCA estimation of 16 years, corresponding to 1994 (95% HPD: 1941- 1999). Clade II dated back to 1994 (95% HPD 1941-1999) and is divided into subclade IIa included 2 statistically supported clusters. 1st including 3 isolates, dated back to 2002 (95% HPD: 1948-2005) and 2nd including 3 isolates (2 female IDU, 1 sexual risk), dated back 1999 (95% HPD: 1939-2001), subclade IIb dated back to 1997 and included 9 sequences, of which three were included in cluster dating back to 2001 (95% HPD: 1947-2002) and 2 included in another cluster dating back to 2002 (95% HPD: 1952-2006), both of them statistically supported.

Bayesian skyline plot (BSP) of the HBV–D3 sequences



The BSP showed that effective number of infections grew exponentially between 1987 and 1995 and then curve reached plateau approximately around year 2000

Maximum parsimony migration patterns of HBV-D3 polymerase sequences to/from different risk groups



HBV-D3 sequences were assigned to six risk groups: sexual, IDU, surgery-endoscopy-dental care, tattoo-acupuncture, unknown and household contact, and gene flow among these groups was estimated.

A= sexual intercourse
B= IDU
C= surgery, endoscopy, dental care
D= tatoo, acupuncture
E= Unknown
F= Familiarity

The different amino acids present at each site under positive selection are given in brackets

Positively selected sites (ω for sites > 1) HYPHY software	**97 (Y,D,H); 165 (L,H,F); 171 (T,A,N,I)**
Negatively selected sites (ω for sites < 1) HYPHY software	**H 176; H 240; F 244; H 259; G 275; N 279; K 282; K 284; Y 288; L 290; F 292; G 294; G 301; E 306; K 311**

The selection pressure analysis for polymerase region of HBV-D3 2nd data set, through computation of ratio of dN/dS (ω) with Hyphy programs, revealed limited positive selection and abundant negative selection for polymerase proteins, dN/dS rate (ω) being around 2.55 .

ω; non-synonymous to synonymous

- **FEL analysis revealed 15 codons under statistically significant negative pressure and REL analysis identified 3 sites under significant positive selection.**

- **The positively selected sites were not located in a specific domain of the polymerase protein, whereas some of the negatively selected site were located in reverse transcriptase RNA-dependent DNA polymerase domain and others were located in carboxy terminal domain of the viral DNA polymerase.**

Conclusion

- **The majority of pts were infected by HBV-D (D3); 26% with A2, not endemic in this region, prevailing in Africa and India.**

- **Time-scaled phylogeny reconstruction of HBV-D3 sequences: tree root had tMRCA estimate of 23 years and originated in 1987.**

- **tMRCA of tree root is recent and dates back to less than 30 years ago. HBV-D3 circulating today in Italy shares a common ancestor that existed in the late 1980s.**

- **HBV-D3 clades may be epidemiological networks originating from introduction of new strains after extinction of the majority of strains circulating before.**

- **In HBV-D3 sequences a viral gene flow was observed between IDV patients and who had acquired HBV by unsafe sexual practices, suggesting specific preventive strategies.**

- The level of endemicity of HBVinfection in Italy is today low and D infections predominant. new HBV strains have recently been introduced as a result of immigration from regions of high endemicity, as shown by two studies regarding the circulation of strains in new cases of HBV infection. The displacement of people from regions where HBV vaccination has not been implemented universally, in combination with poor socioeconomic conditions, can result in pockets of high endemicity in countries where the incidence rates are otherwise low This scenario results in an increased risk of infection and in the introduction of non-D genotypes, which over time could affect the expression pattern of HBV-related diseases and the response to antiviral therapy.
- The time-scaled phylogeny reconstruction of the acute HBV genotype D3 sequences showed that the tree root had a tMRCA estimate of 23 years and originated in 1987 (95% HPD: 1870-1997). We inferred the tMRCA of the internal nodes within Bayesian tree of D3 data set. Interestingly, our time-scaled reconstruction supported a relatively recent history of the currently circulating D3. The coalescent-based population dynamics analysis of D3 showed that the number of infections starting from 1987 slightly increased until 1995, remained steadily high in 2000 and is still the prevalent subgenotype in AVH-B in Italy.

- A recent phylogeographical reconstruction hypothesized that the HBV D originated in India and spread worldwide in the first decades of 20th century, probably favored by world wars and unsafe medical practices In the present study, where all the isolates were sampled from acute infections, the tMRCA of the tree root, corresponding also to the beginning of the increase in the number of primary infections, is recent and dates back to less than 30 years ago. Our results suggest that D3 circulating today in Italy shares a common ancestor that existed in the late 1980s.
- D3 clades observed in this study may be epidemiological networks originating from the introduction of new strains (still circulating in the population) after the extinction of the majority of strains circulating before. In D3 sequences, a significant viral gene flow was observed between IDU patients and subjects who had acquired HBV by unsafe sexual practices, suggesting specific preventive strategies.
- The selection pressure analysis in D3 data set showed 15 sites under negative selective pressure and 3 sites under positive selective pressure. The presence of abundant negative selective pressure most likely reflects a high degree of conservation of the polymerase protein, probably necessary forHBV biological function. The positive selection sites suggest evolutionary “hot spots” possibly involved in the severity of the associated illness. These positive selection sites should be more extensively studied for a better understanding of the correlation between the epidemiological and clinical characteristics of AVH-B.