Basic phylogenetics

23rd IPFA/PEI workshop on blood-borne pathogens
Lisbon, May 25th 2016

prof.dr. Hans L. Zaaijer, medical microbiologist
Sanquin - Blood-borne Infections & AMC - Clinical Virology
Amsterdam NL

Why phylogenetics?

- science (evolution and origin of organisms)
- outbreak management (eg. epidemiology of Ebola: recent recurrences are not imported but endogeneous)
- patient care (eg. hep C - which pills? genotyping of HCV)

in blood transfusion:
- epidemiology of infections in donors.
- did donor or product A infect recipient B?

phylogenetics = hierarchical, historical relationship
Greek: φυλή, φυλον = tribe, clan, race
γενετικος = origin, source, birth

case: patients infected by occult HBV donor?

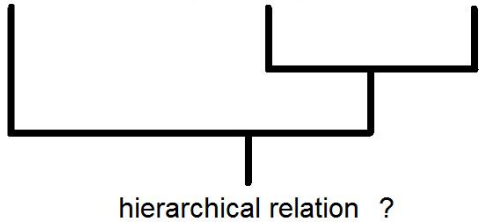
- 2008: HBV NAT screening introduced for all Dutch donors.
- 2009: female donor X is found to have occult HBV infection.
  25 years ago, while applying as a new donor, she reported having had acute hepatitis, just after her marriage some years before.
- Look back:
  - 13 recipients could no be traced
  - 27 recipients had died
  - 15 recipients tested: 4 HBsAg+ , ***infected by donor?***

Lieshout-Krikke ea; Transfusion 2016 March; 691-699.

comparing organisms

for example: using yes/no properties ('binary characters')

mammal	mammal	mammal
single toe	single toe	paired toe
no ruminant	no ruminant	ruminant
stripes	no stripes	no stripes
no horns	no horns	horns



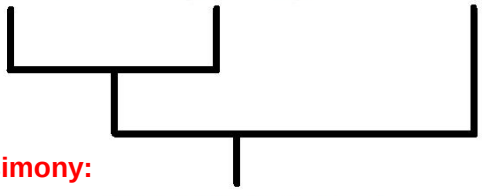
a 'cladogram'

hierarchical relation ?

comparing organisms

for example: using yes/no properties ('binary characters')

mammal	mammal	mammal
single toe	single toe	paired toe
no ruminant	no ruminant	ruminant
stripes	no stripes	no stripes
no horns	no horns	horns



principle of parsimony:
simplest = best

hierarchical relation !

adapted from Open University – 'Cladistics'

quantitative hierarchy of relationships

example of distances and visualisation

	Amsterdam	The Hague	Rotterdam	Utrecht
Amsterdam	0 km			
The Hague	60 km	0 km		
Rotterdam	80 km	20 km	0 km	
Utrecht	25 km	45 km	40 km	0 km

'quantitative distance' between zebra, horse and cow

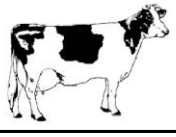
counting the differences:

	zebra	**horse**	**cow**
zebra	**0**		
horse	**1**	**0**	
cow	**4**	**3**	**0**

single toe
no ruminant
stripes
no horns

single toe
no ruminant
no stripes
no horns

paired toe
ruminant
no stripes
horns



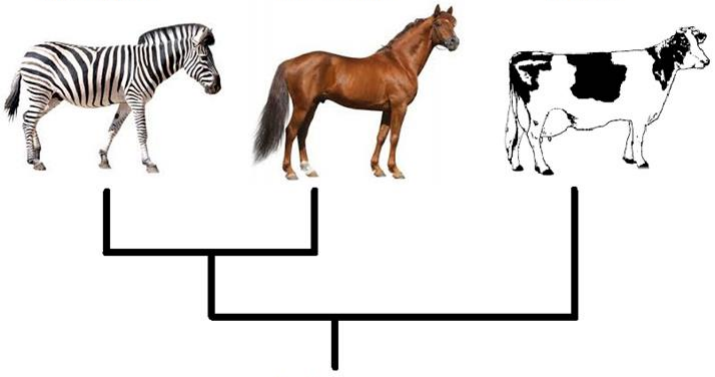
comparing organisms

genetic distance between organisms

single toe
no ruminant
stripes
no horns

single toe
no ruminant
no stripes
no horns

paired toe
ruminant
no stripes
horns



cladogram

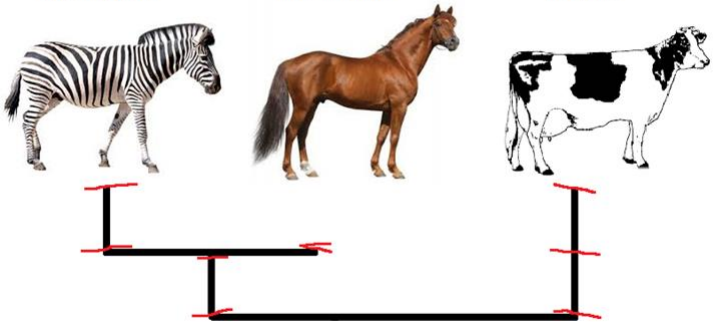
comparing organisms

genetic distance between organisms

single toe
no ruminant
stripes
no horns

single toe
no ruminant
no stripes
no horns

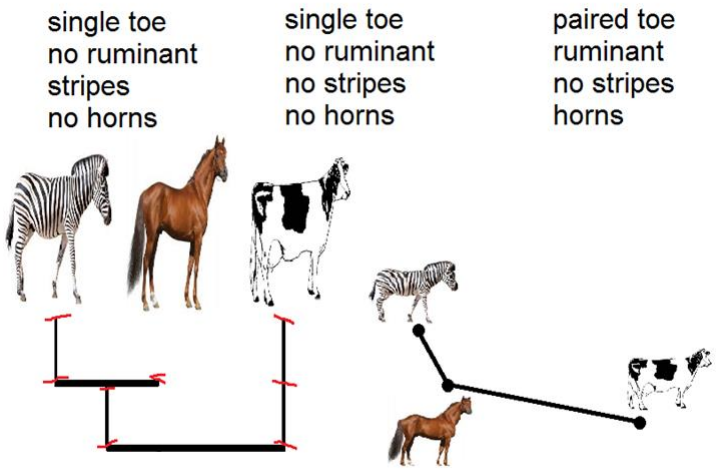
paired toe
ruminant
no stripes
horns



quantitative : phylogram

comparing organisms

genetic distance between organisms



quantitative : phylograms

Comparing isolates of HIV, HBV, HCV, S.aureus from donors, blood products and recipients

No horns, hoofs or stripes,
but we can compare DNA sequences :

1) determine the differences (distances) between sequences:
   a) align the DNA sequences,
   b) count number of different nucleotides (= '4-state characters'),

2) visualise and compare the differences to reference strains.

Requirements:
1) Sufficient variation (information) in the DNA sequences.
2) For comparison: isolates from same period, same region.

comparing HBV isolates

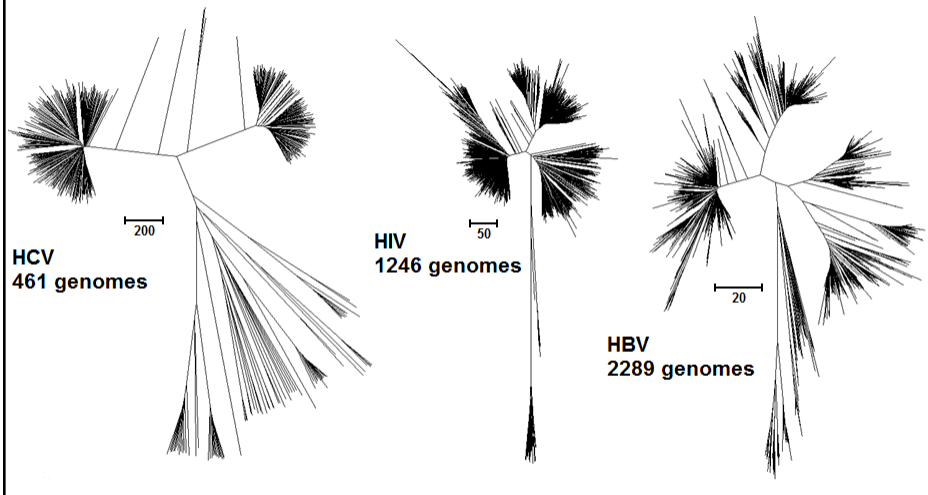
genetic distance between full HBV genomes from GenBank

HBV DNA sequences from GenBank :

```
HBV  1  A C T A T C C A G T A A A C A T A G T C A C T G ...
HBV  2  A C T G T C C A G T A A A C A T A C T C A C T G ...
HBV  3  A C T A T C C T G T A A A C A T A G T C A C T G ...
HBV  4  A C T A T C C T G T A A A C A T A G T C A C T G ...
HBV  5  A C T A T C C A G T A C A C A T A G T C A C T G ...
HBV  6  A C T A T C C A G T A C A C A T A G T C A C T G ...
HBV  7  A C T A T C C A G T A C A C A T A G T C A T T G ...
HBV  8  A C T A T C C A G T A A A C A T A G T C A T T G ...
HBV  9  A C T A T C C A G T A A A - - - A G T C A T T G ...
HBV 10  A C T A T C C A G T A A A C A T A G T C A T T G ...
 etc.
```

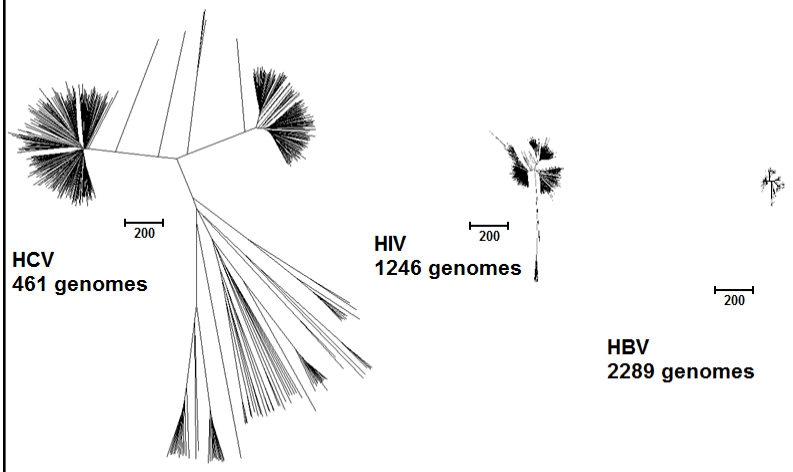
using public domain software:
1) align sequences, 2) calculate distance tabel, 3) construct tree

genetic diversity of HCV, HIV, HBV



genetic distance: # nt differences
phylogenetic tree: neighbour-joining

genetic diversity of HCV, HIV, HBV



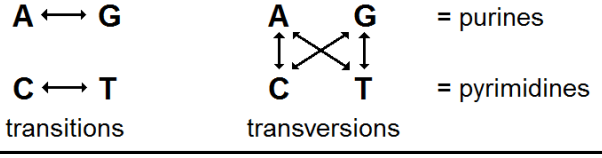
genetic distance: # nt differences
phylogenetic tree: neighbour-joining

comparison of HBVs in Dutch donors (2000)

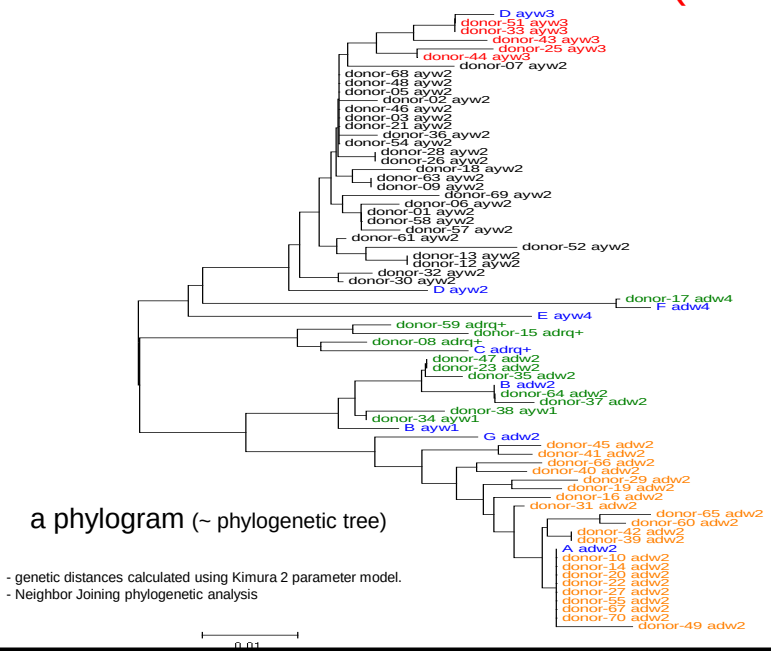
using more sophisticated genetic distances

HBV DNA sequences:

donor 1 A C T A T C C A G T A A A C A T A G T C A C T G ...
donor 2 A C T G T C C A G T A A A C A T A C T C A C T G ...
donor 3 A C T A T C C T G T A A A C A T A G T C A C T G ...
donor 4 A C T A T C C T G T A A A C A T A G T C A C T G ...
donor 5 A C T A T C C A G T A C A C A T A G T C A C T G ...
donor 6 A C T A T C C A G T A C A C A T A G T C A C T G ...
donor 7 A C T A T C C A G T A C A C A T A G T C A T T G ...
donor 8 A C T A T C C A G T A A A C A T A G T C A T T G ...
donor 9 A C T A T C C A G T A A A C A T A G T C A T T G ...
donor 10 A C T A T C C A G T A A A C A T A G T C A T T G ...
etc.



HBV in Dutch blood donors (2000)



a phylogram (~ phylogenetic tree)

- genetic distances calculated using Kimura 2 parameter model.
- Neighbor Joining phylogenetic analysis

how to construct a tree?
how reliable is my tree?

treebuilding - common computational methods:

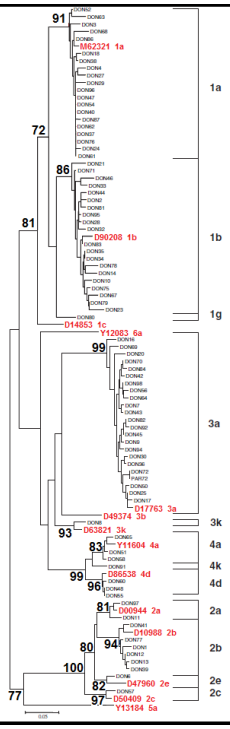
- maximum parsimony
- neighbor joining
- maximum likelyhood
- baysian

validation (confidence level) by 'bootstrapping':

- repeat the computation of tree 1000 times, each time on a slightly altered dataset; robust nodes in the tree remain mostly unaffected. (▸ percentage at each node)

HCV genotyping of 59 Dutch blood donors

van de Laar ea, Transfusion 2006; 1719.

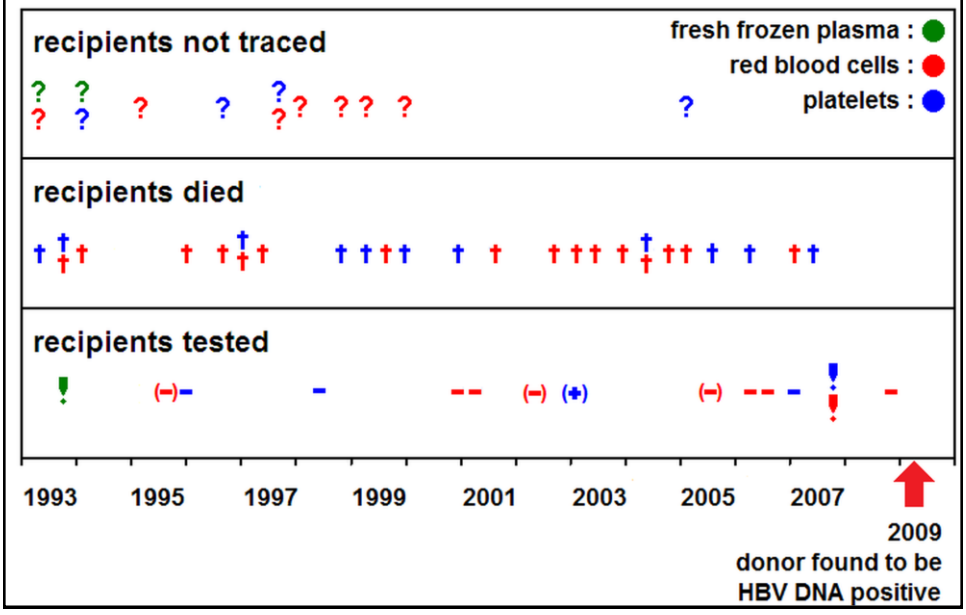


case: recipients of occult HBV donor

- In 2008 HBV NAT screening was introduced for all Dutch donors.
- Female donor X was found to have occult HBV infection.
  20 years ago, while applying as a new donor, she reported having had acute hepatitis, just after her marriage some years ago.
- Look back:
  - 13 recipients could no be traced
  - 27 recipients had died
  - 15 recipients tested: 4 HBsAg+ , ***infected by donor?***

Lieshout-Krikke ea; Transfusion 2016 March; 691-699.

3 HBV infected recipients (1993, 2007), associated with a donor with occult HBV infection (detected in 2009).

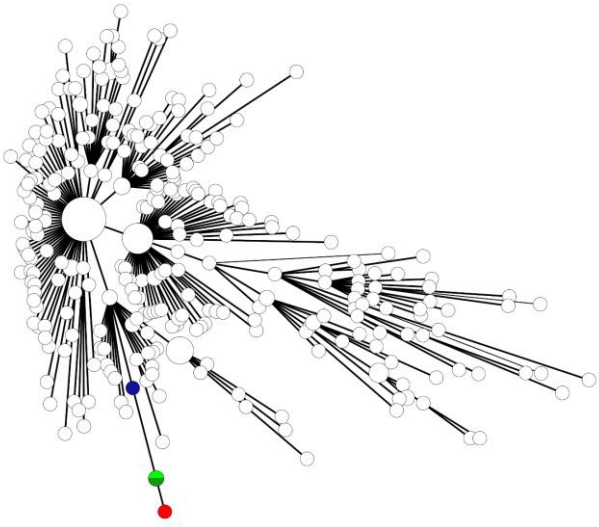


Comparing the HBVs from donor, recipients and reference strains

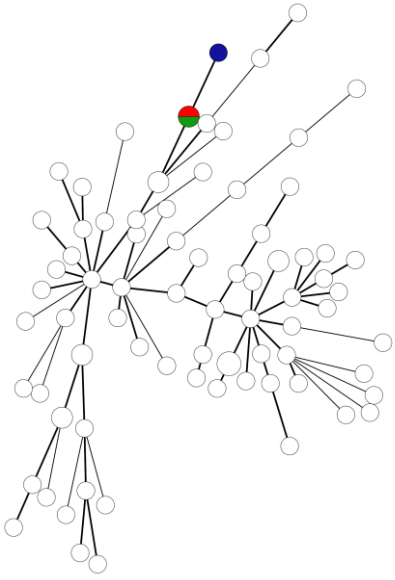
HBV DNA sequences from:

donor X	A C T A T C C A G T A A A C A T A G T C A C T G ...
recip.A	A C T G T C C A G T A A A C A T A C T C A C T G ...
recip.B	A C T A T C C T G T A A A C A T A G T C A C T G ...
recip.C	A C T A T C C T G T A A A C A T A G T C A C T G ...
ref. 1	A C T A T C C A G T A C A C A T A G T C A C T G ...
ref. 2	A C T A T C C A G T A C A C A T A G T C A C T G ...
ref. 3	A C T A T C C A G T A C A C A T A G T C A T T G ...
ref. 4	A C T A T C C A G T A A A C A T A G T C A T T G ...
ref. 5	A C T A T C C A G T A A A C A T A G T C A T T G ...
ref. 6	A C T A T C C A G T A A A C A T A G T C A T T G ...
etc.	

Minimum spanning tree, showing HBV surface gene sequences of donor X (red), patient A (light green), pat. B (dark green), pat. C (blue), and 375 contemporary Dutch cases of hepatitis B (RIVM).

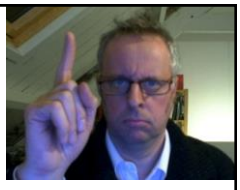


Minimum spanning tree, showing HBV core gene sequences of donor X (red), patient A (green), patient C (blue), and 86 contemporary Dutch cases of hepatitis B (RIVM).



Lieshout-Krikke,Transfusion 2016

Summary



- phylogenetic analysis may help to confirm or exclude donor or product A as source of infection for patient B.

- the chosen genetic sequence of the pathogen should carry sufficient variation in your population.

- sequences from recent local isolates must be available.

- phylogenetics are fun.