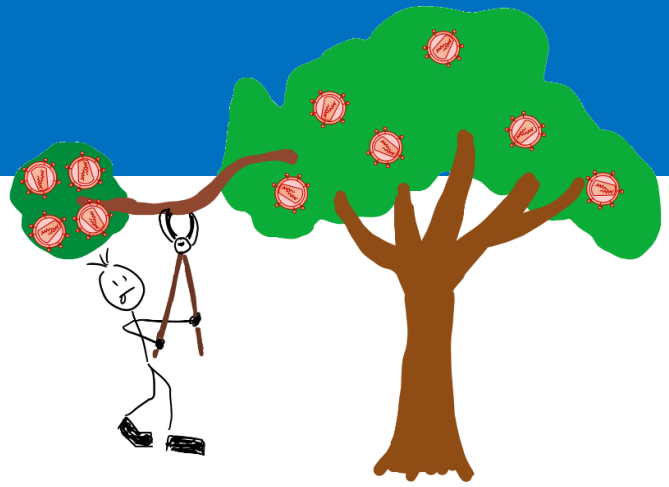
Phylogénie et clusters de transmission virale



Dr Benoit Visseaux

Maitre de conference des Université – Praticien Hospitalier

Laboratoire de Virologie

Hôpital Bichat Claude Bernard, Paris

HÔPITAUX UNIVERSITAIRES PARIS NORD VAL DE SEINE

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS

lame UMR 1137

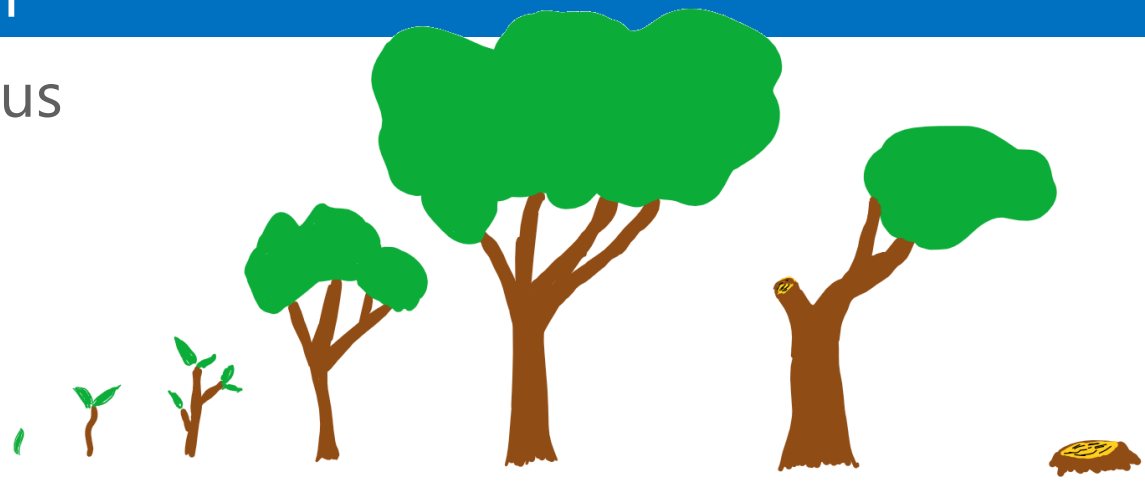
anrs France REcherche Nord & sud Sida-hiv Hépatites — Agence autonome de l'Inserm

Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

université PARIS DIDEROT

Qu’est ce qu’un arbre phylogénétique ?

Quelles informations nous donnent-ils…



L'évolution…

- Une idée (pas si) ancienne :

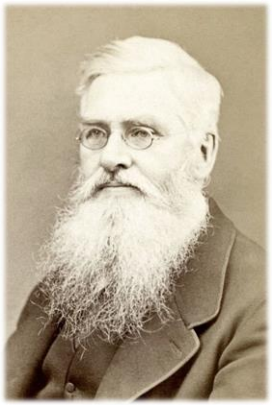
Jean-Baptiste de Lamarck
1744-1829

1809 : théorie de Lamarck

arehn.asso.fr

- Acquisition de caractères par adaptation à l'environnement
- Transmission des caractères à la descendance

Charles Darwin
1809-1882



Alfred Russel Wallace
1823-1913

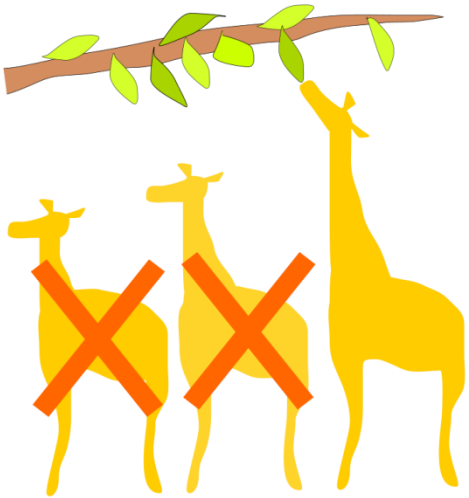
1859 : théorie de Darwin

arehn.asso.fr

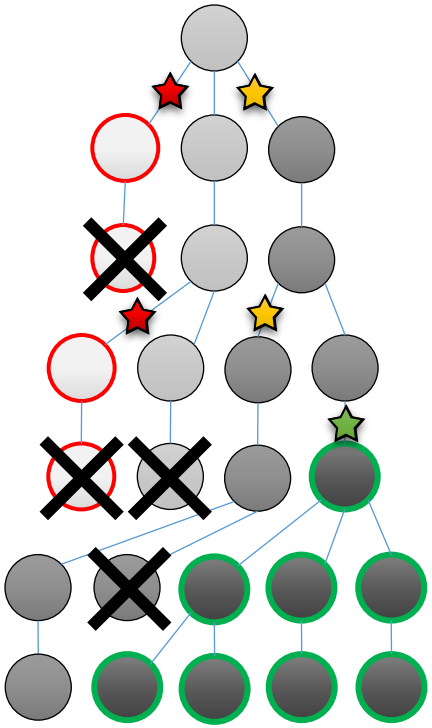
- Acquisition de caractères **au hasard**
- **Ajout de la sélection naturelle** (positive et négative)
- Transmission des caractères à la descendance

Evolution et phylogénétique

Changements phénotypiques



Changements génotypiques



Les mutations sont à l'origine des variations
(Premier moteur : taux de mutations)

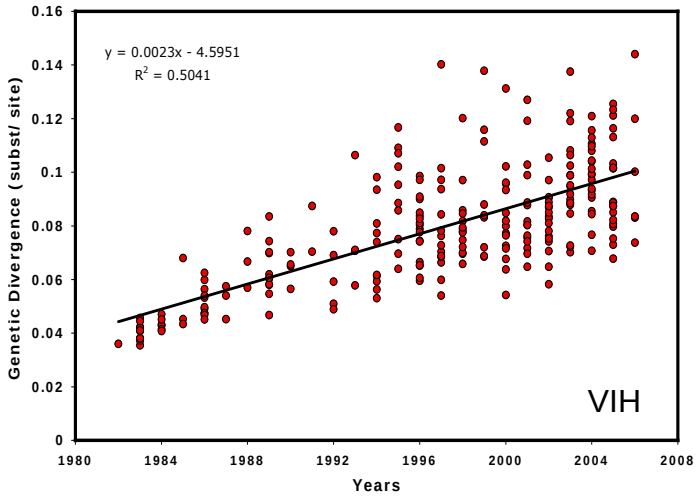
Elles peuvent ensuite être sélectionnées en fonction de leur impact
(Second moteur : la sélection)

- ★ **Négative** *(défavorable)*
- ★ **Neutre**
- ★ **Positive** *(favorable)*

Génome : archive de l'évolution !

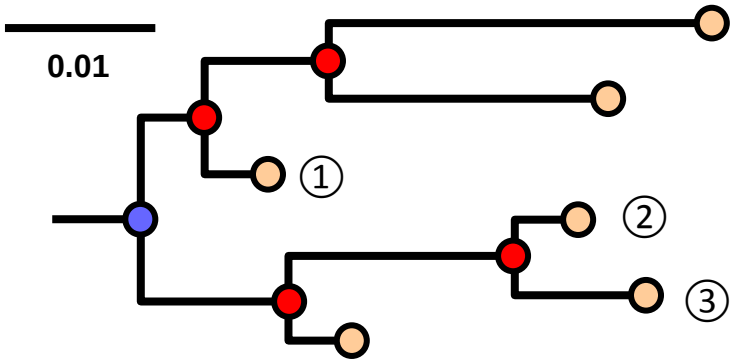
- **L'évolution d'un organisme laisse des empreintes (mutations) au sein du génome**
- Ces empreintes peuvent être, le plus souvent, mesu

- **Pour un organisme donné, l'évolution se déroule à**
- Le nombre de différences entre deux séquences gén t leur niveau de relation

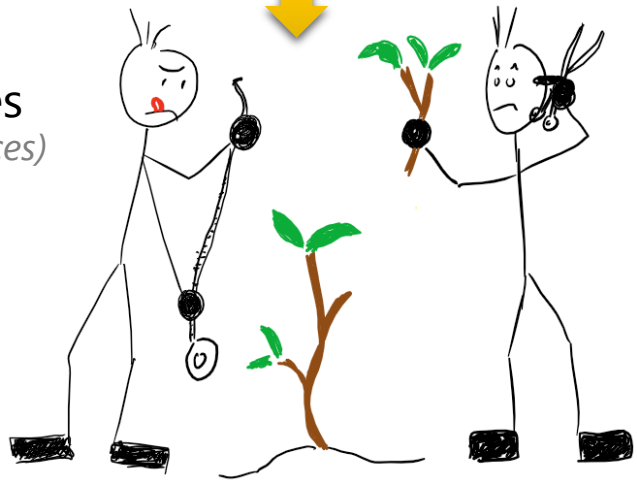


➔ Les approches phylogénétiques vont permettre de représenter et d'étudier les relations évolutives des différents organismes

Les méthodes phylogénétiques



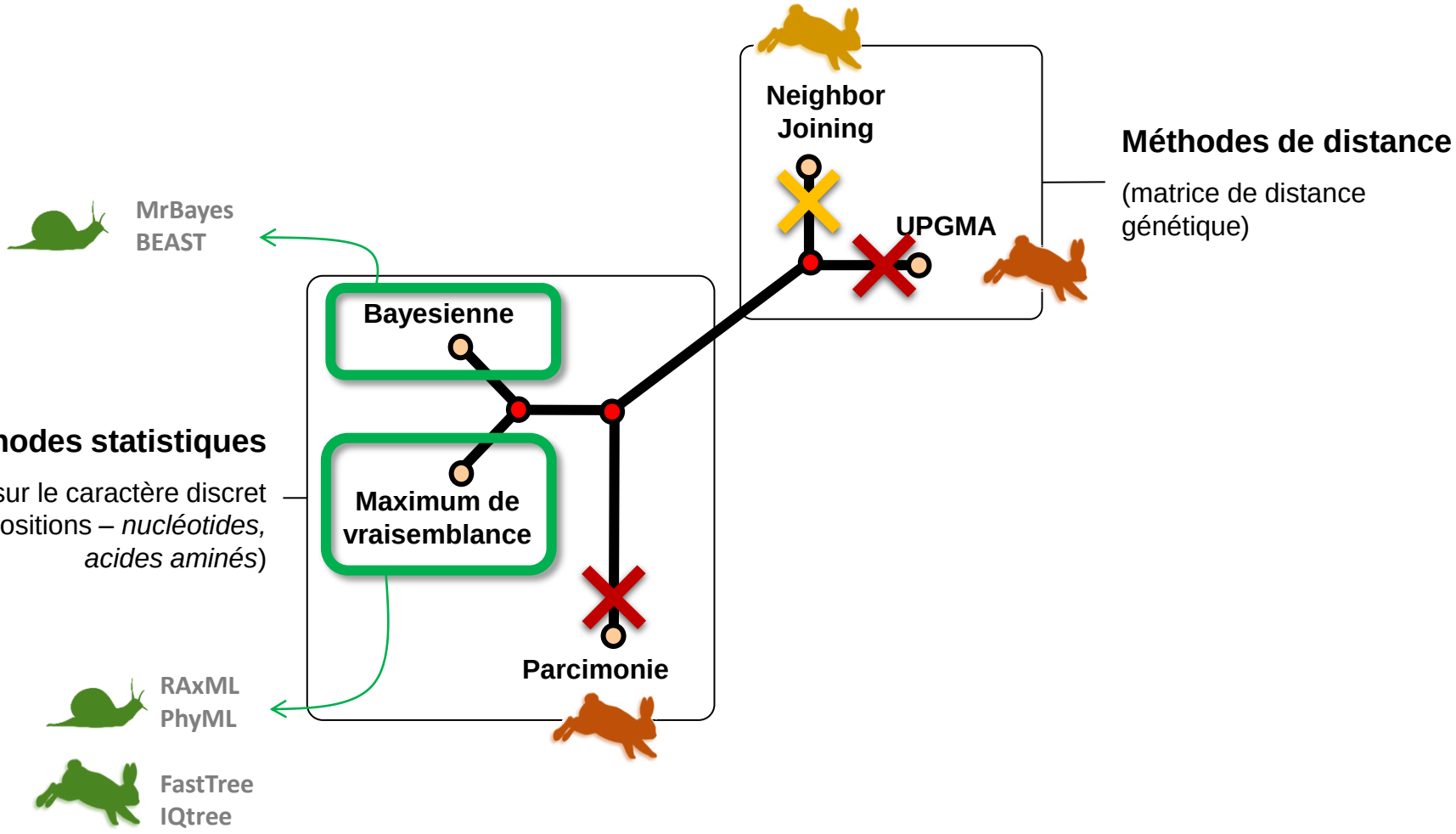
Longueur des branches :
distance entre les séquences
(N différences / longueur des séquences)



Topologie :
homologie entre séquences
(distance + position des mutations)

Reconstruction phylogénétique

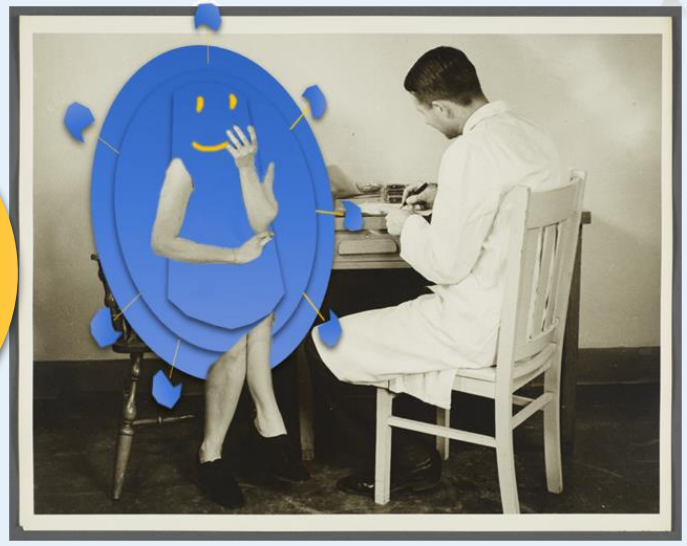
- Les deux grandes familles de méthodes de reconstruction phylogénétiques :



Epidémiologie vs. phylogénie…

L'enquête épidémiologique

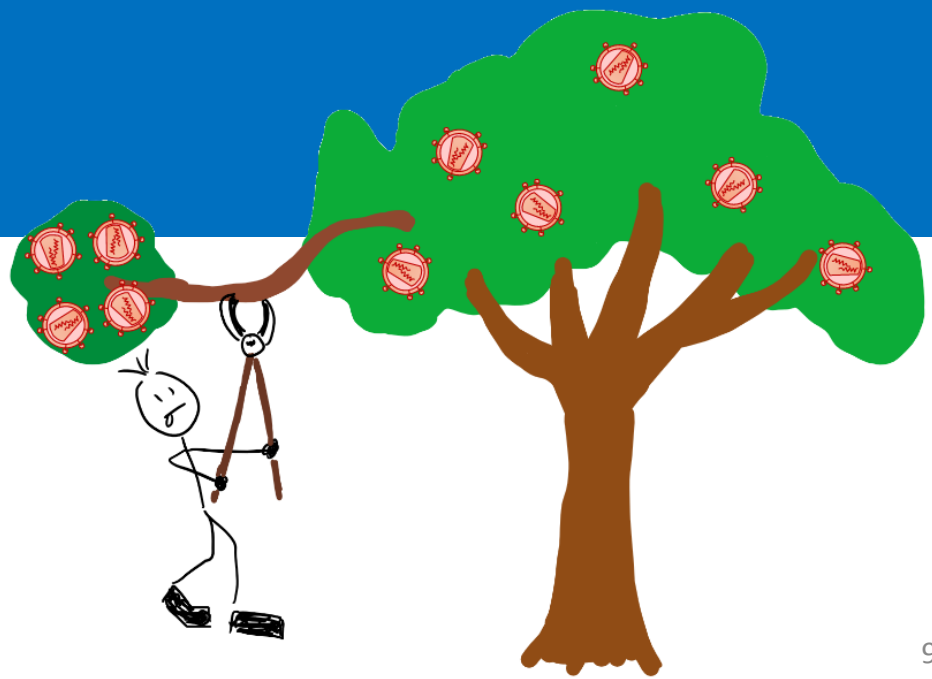
- Recueil des informations épidémiologiques (temps réel)
- Description des populations et de leurs évolutions
- Identification des facteurs de risques (associés ou pronostiques)
- Identification de la source d'une infection (source, datation, dispersion)
- (…)



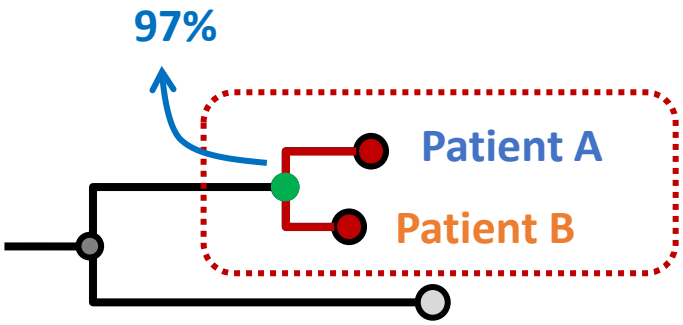
L'enquête phylogénétique

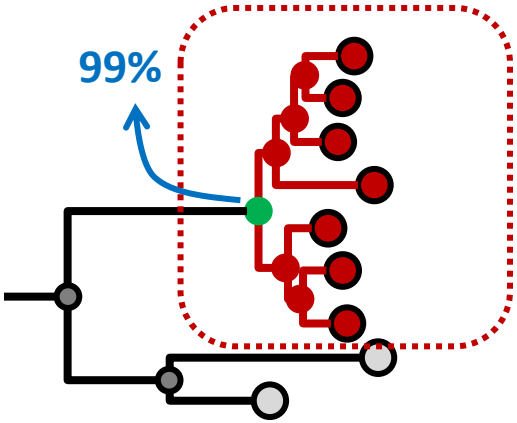
- Reconstruction des événements épidémiologiques (passés)
- Etude de la dispersion de marqueurs génétiques (virulence, résistance)
- Identification des chaines de transmission, des facteurs associés
- Identification de la source d'une infection (source, datation, dispersion)
- (…)

Chaînes de transmission récentes



« Chaînes de transmissions récentes »

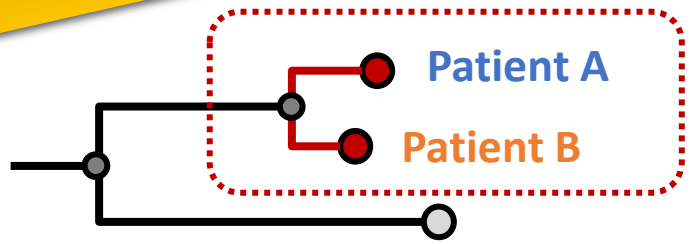


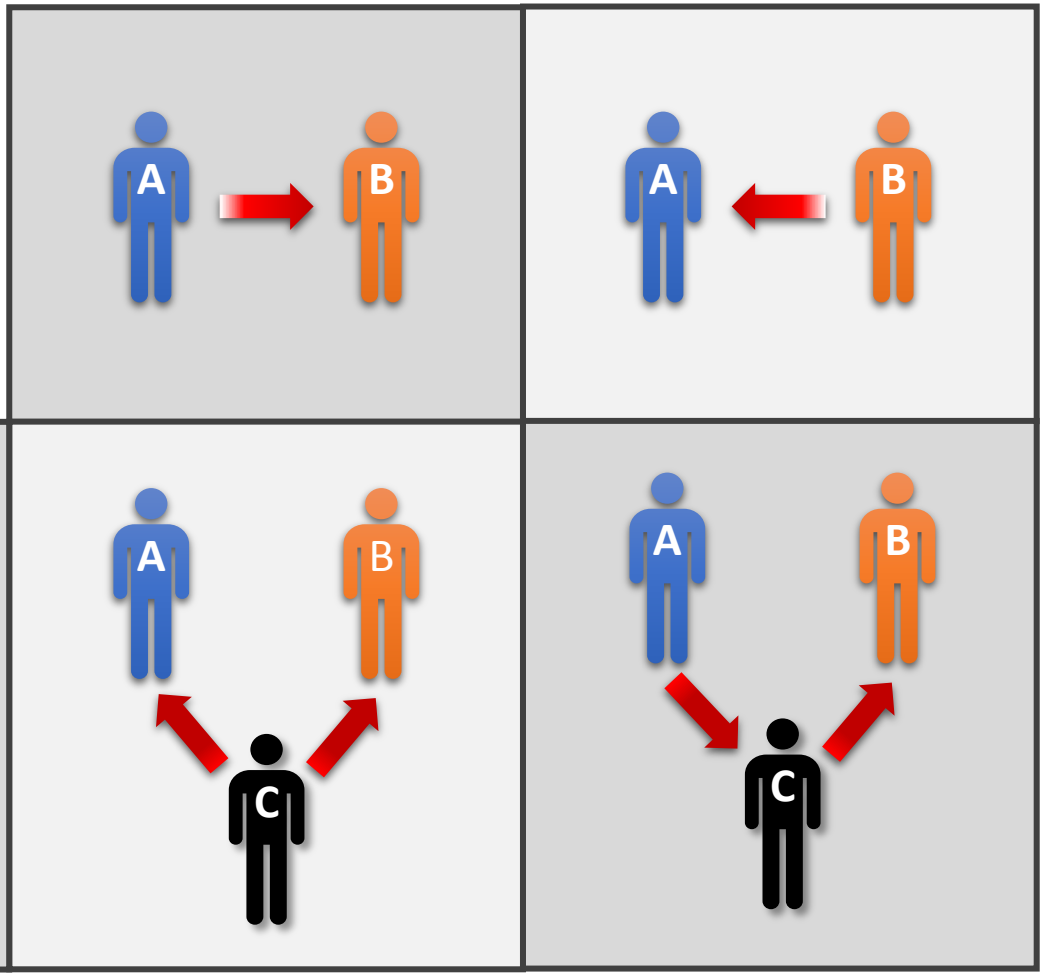


- **Patients réunis sur un arbre phylogénétique par :**
  - Un **même nœud**
  - Une **distance génétique** *(nombre de mutations)* maximale de 4,5%
  - Un **support de branche** *(confiance statistique)* >95%

Quelles informations dans le VIH ?

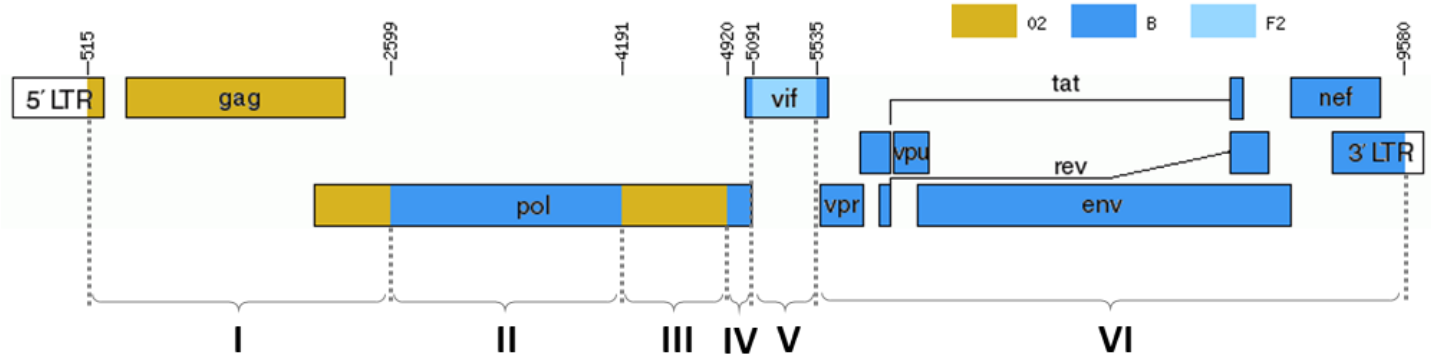
Quelles sont les personnes appartenant à un même réseau de transmission ?





Un exemple : CRF94_cpx

- Profils de typage inhabituels pour plusieurs souches virales :
  - **Protéase** : CRF02_AG
  - **Reverse transcriptase** : B
  - **Recherche des séquences proches** : ramène une dizaine de patients sur la base de la Pitié-Salpétrière ?!?!



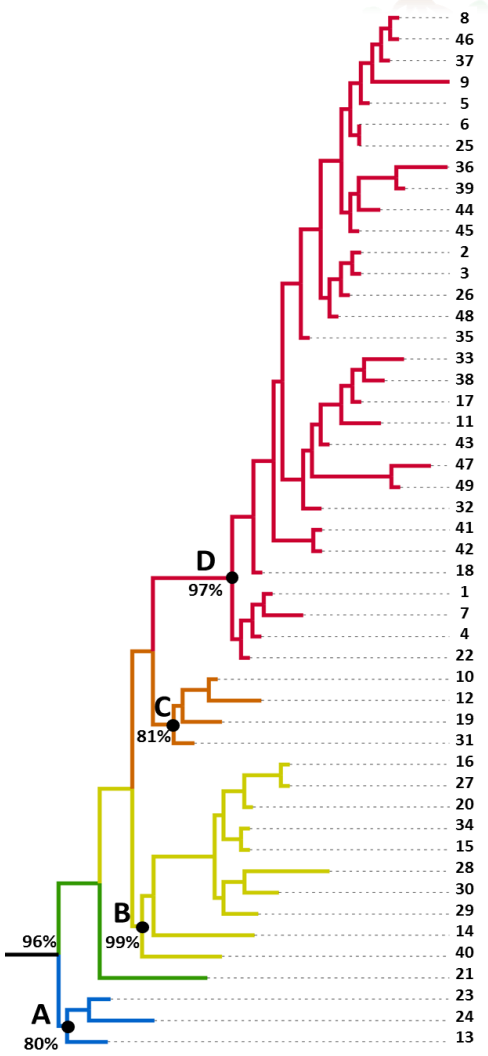
Analyse plus fine du cluster…

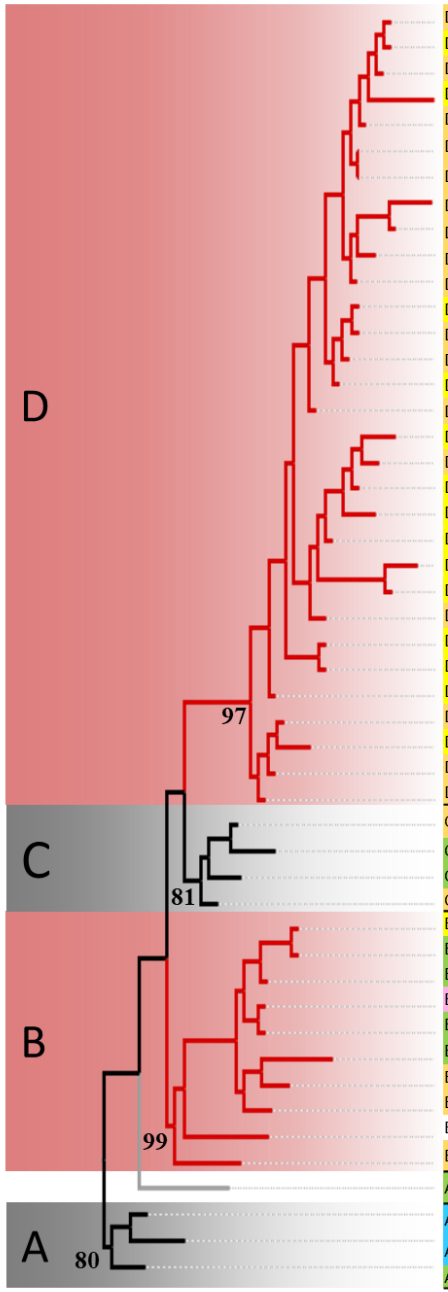
- **Enquête systématique auprès des centres de l'AC43 :**
  - 50 patients identifiés au 31 décembre 2017
  - Date du premier diagnostic : janvier 2013

- **Analyse phylogénétique**
  - Distance génétique maximale 3,4%
  - Un support de branche > 95%

  ➔ C'est un cluster de transmission récent !

- **Recueil de quelques informations** cliniques et épidémiologiques



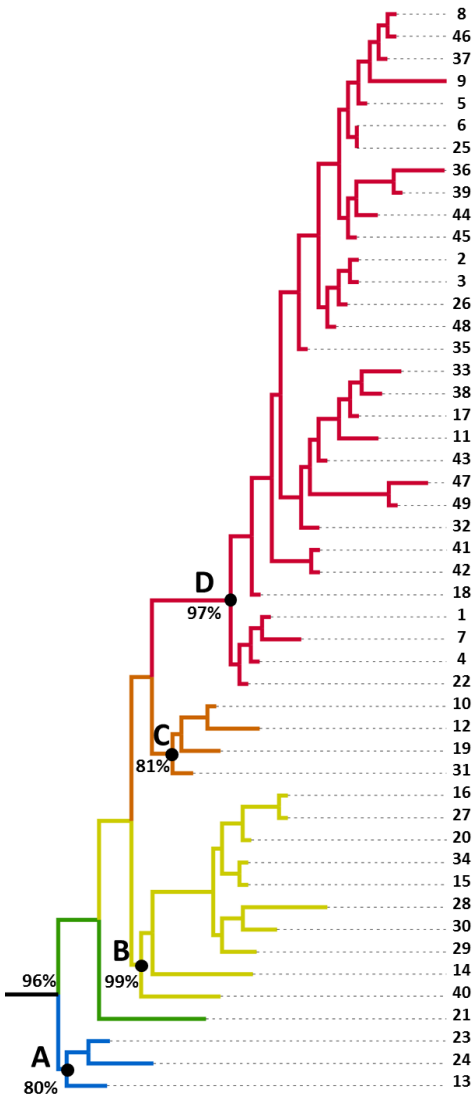


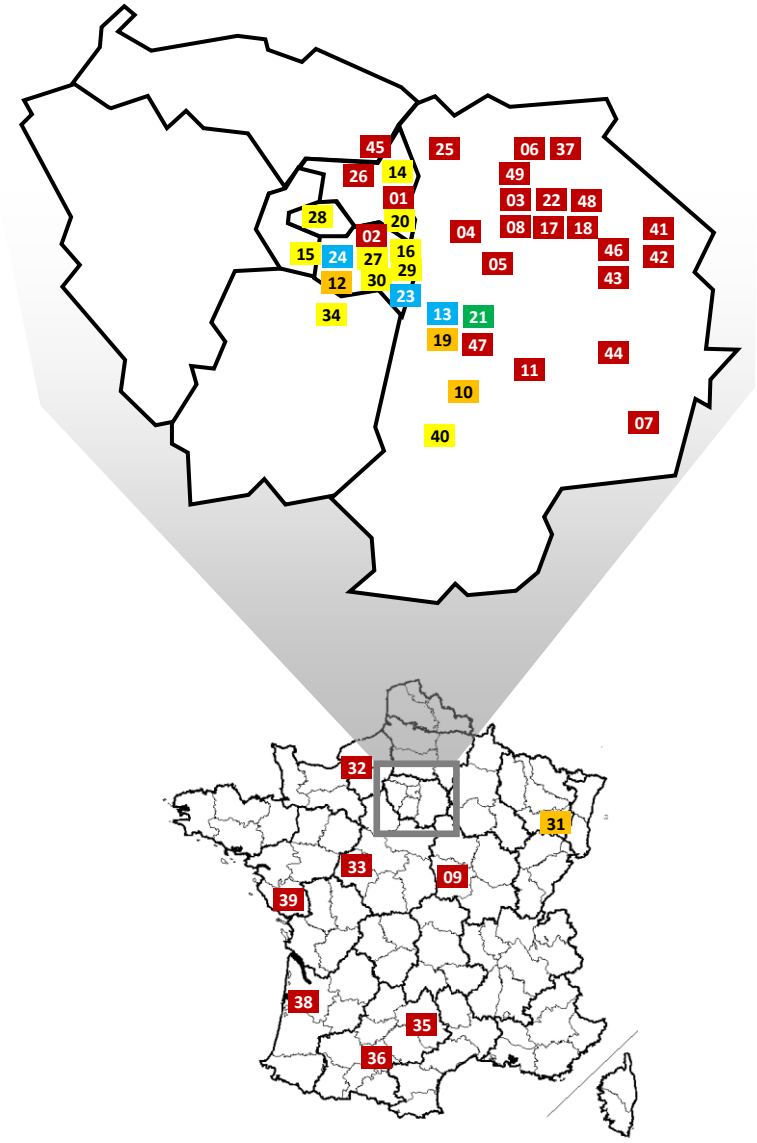
Patients	Year of diagnosis	Risk group	Stage of Infection	HIV Viral Load Log_{10} Copies/mL	CD4 cells count /mm3	Geno2pheno Tropism	
						V3 gene Pattern	False positive rate
D31	2016	MSM #	< 1 Year	4.23	658		
D30	2017	MSM	< 1 Year	6.41	742	2M/25QE	8.1 (X4)
D29	2016	MSM #	Primary	6.05	349	2M/25E	8.5 (X4)
D28	2017	MSM	Chronic	3.38	NA		
D27	2016	MSM #	Primary	> 7.00	298	2M/25E	8.5 (X4)
D26	2016	MSM #	Primary	6.31	320		
D25	2016	MSM #	Primary	5.18	740		
D24	2016	MSM	Chronic	5.42	538	2M/25Q	7.8 (X4)
D23	2016	MSM	Primary	>7.00	793	2M/25E	8.5 (X4)
D22	2016	MSM #	Primary	4.88	550	2M/25E	8.5 (X4)
D21	2016	MSM	< 1 Year	5.65	215	2M/25QE	8.1 (X4)
D20	2017	MSM	Advanced	6.44	23	2M/25KE	1.7 (X4)*
D19	2016	MSM #	Chronic	5.68	374		
D18	2016	MSM	Chronic	5.71	540		
D17	2017	MSM	Primary	6.18	NA		
D16	2016	MSM	Primary	5.78	252	2M/25E	8.5 (X4)
D15	2017	MSM #	Primary	6.90	561	2M/25E	8.5 (X4)
D14	2016	MSM #	Primary	5.42	535	2M/25E	8.5 (X4)
D13	2017	MSM #	Chronic	4.77	358		
D12	2017	MSM	Primary	5.81	402		
D11	2017	MSM	Primary	4.04	647	2M/25Q	7.8 (X4)
D10	2017	MSM	< 1 Year	5.97	190		
D9	2017	MSM	< 1 Year	5.60	607	2M/25E	8.5 (X4)
D8	2016	MSM #	Primary	5.92	201	2M/25E	8.5 (X4)
D7	2017	MSM	Primary	5.72	551	2M/25K	2.8 (X4)*
D6	2017	MSM	Primary	4.89	333	2M/25K	2.8 (X4)*
D5	2017	MSM #	< 1 Year	5.86	110		
D4	2016	MSM #	< 1 Year	3.90	236		
D3	2017	MSM #	Chronic	5.39	538		
D2	2016	MSM	Chronic	5.51	384		
D1	2016	MSM	Chronic	4.99	533		
C4	2016	MSM	< 1 Year	6.53	52		
C3	2015	MSM	Primary	5.20	1081	2M/25K	2.8 (X4)*
C2 *woman*	2015	Hetero.	Chronic	5.07	533		
C1	2016	MSM	Advanced	6.00	46	2M/25KE	1.7 (X4)*
B10	2017	MSM	Chronic	5.32	113	2I/25K	10.5 (R5)*
B 9	2015	MSM	Chronic	5.20	406		
B 8	2015	MSM	Primary	6.85	282	2I/25E	60 (R5)
B 7	2014	NA	NA	NA	NA		
B 6	2015	MSM	< 1 Year	5.16	371	2I/25K	10.5 (R5)*
B 5	2015	MSM	Primary	4.78	543	2I/25K	10.5 (R5)*
B 4	2016	MSM	< 1 Year	5.23	190		
B 3	2016	MSM	Chronic	5.06	259		
B 2	NA	Hetero.	Chronic	NA	NA		
B 1	2016	MSM	Chronic	4.52	370	2M/25Q	9.0 (X4)
A B	2015	MSM	Primary	5.64	407		
A 3	2013	MSM	Chronic	4.73	231	2I/25E	60 (R5)
A 2	2013	Hetero.	Chronic	2.94	578	2I/25E	56 (R5)
A 1	2015	MSM	Chronic	4.83	550	2I/25E	83 (R5)

- **32 patients**
- **Tous diagnostiqués en 2016-17**
- **15** (46%) **primo-infections**
- **7** (22%) **avec une sérologie négative de moins de 1 an**
- **14** (44%) **ayant un lien épidémiologique commun**

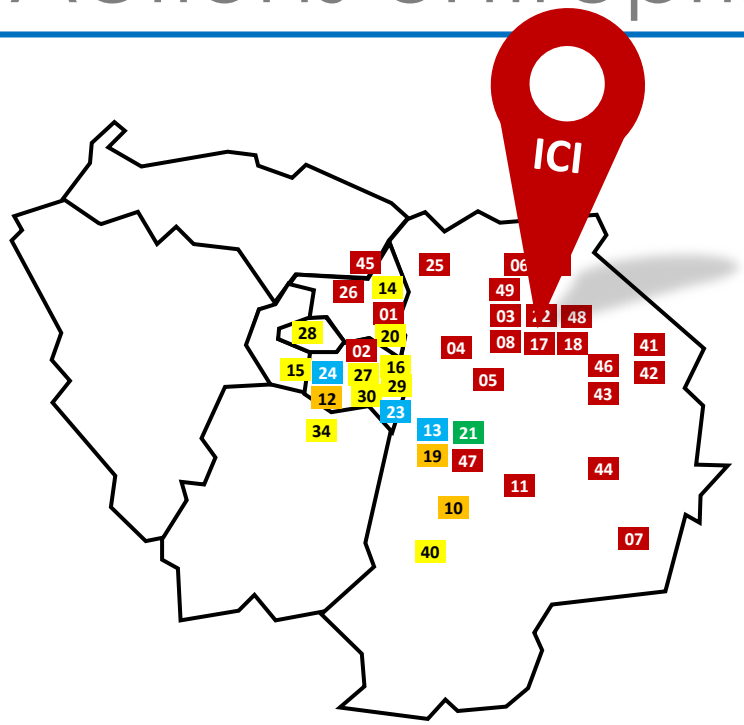
- **18 patients**
- **11** (61%) **diagnostiqués avant 2016**
- **4** (22%) **primo-infections**
- **3** (17%) **avec une sérologie négative de moins de 1 an**

Distribution géographique





Actions entreprises…



- Contact de l'ARS, du CEGGID local, d'association (AIDES), de la société concernée et sa médecine du travail

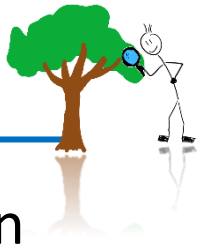
- **Reprise d'actions de prévention**
  - Affichage
  - Stand AIDES au sein de l'entreprise
  - Promotion des moyens de prévention et du dépistage

2e exemple : le « cluster breton »

Virologie - Rennes

Dr Charlotte Pronier
Dr Anne Maillard

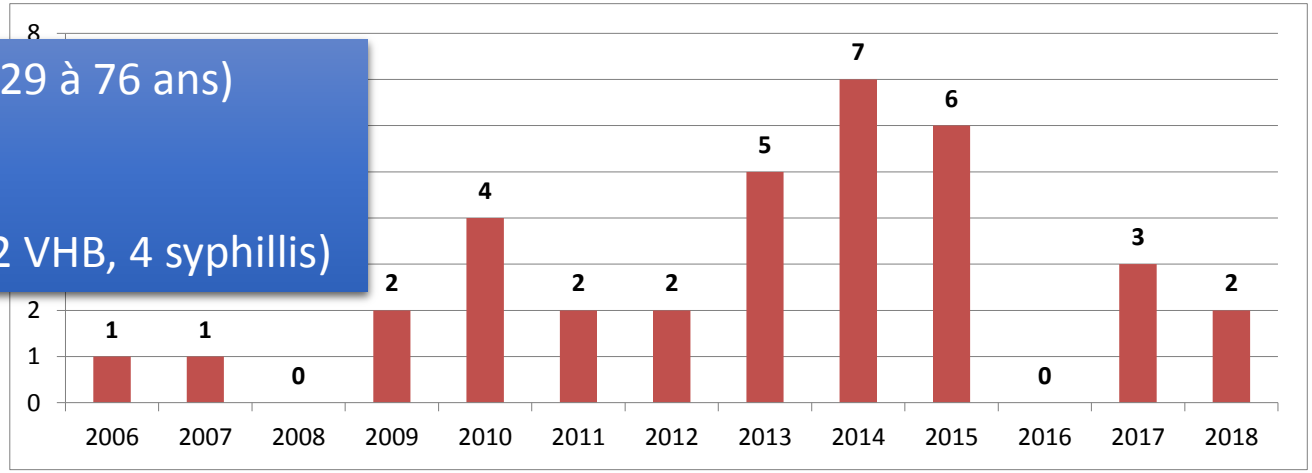
2e exemple : le « cluster breton »

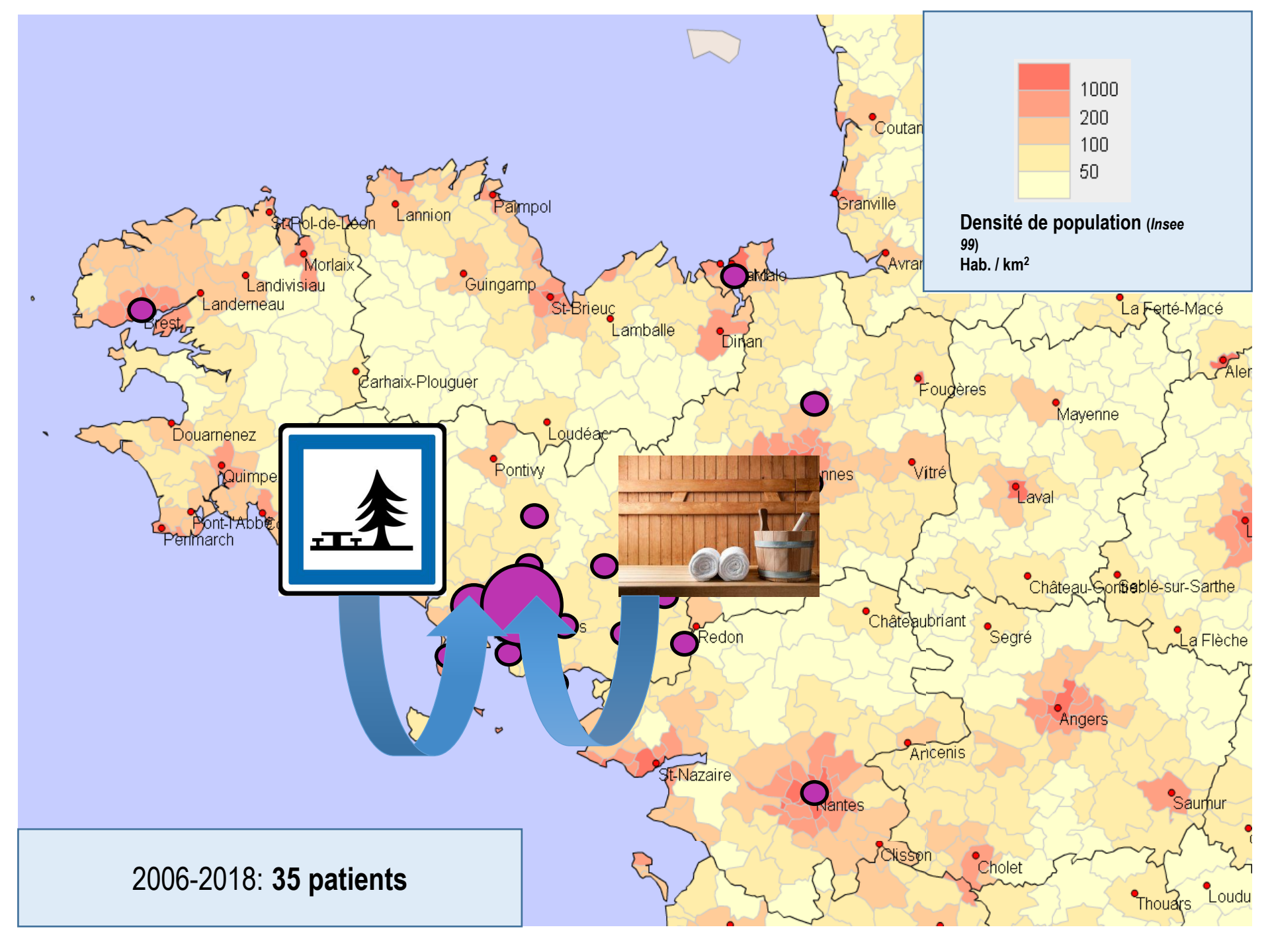


- A Rennes en 2010 : découverte de plusieurs cas d'infection avec un profil de résistance primaire identique
  - **M41L, E44D, T69D, L210E, T215S et K219R**
- Surveillance de nouveaux cas…
- En septembre 2018 : 35 patients avec ce profil

Répondent à la définition de cluster de transmission récente

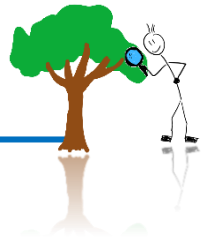
Age médian : 51 ans (29 à 76 ans)
34 hommes
2 primo-infections
Coinfections (7 VHC, 2 VHB, 4 syphillis)





2006-2018: **35 patients**

Combien de clusters en France ?



- Absence d'une base de contenant toutes les séquences effectuées en France
  (pas un cas isolé, existe seulement dans trois pays : UK, Pays-Bas, Suisse)

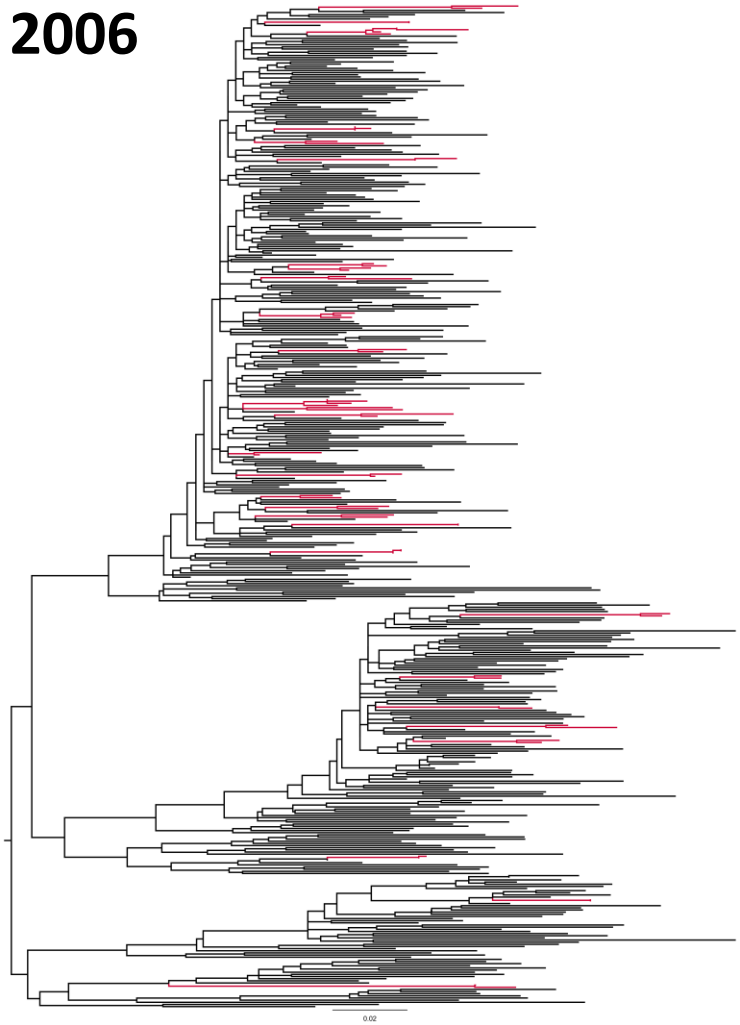
- **Evaluation du nombre de clusters de transmissions récentes dans deux jeux de données nationaux :**

  - **ODYSSEE :** coupes transversales des séquences virales obtenus chez des patients naïfs d'antirétroviraux en 2006, 2011 et 2016

  - **PRIMO :** surveillance continue des primo-infections détectés dans les laboratoires de Virologie français

ODYSSEE

- **2006**

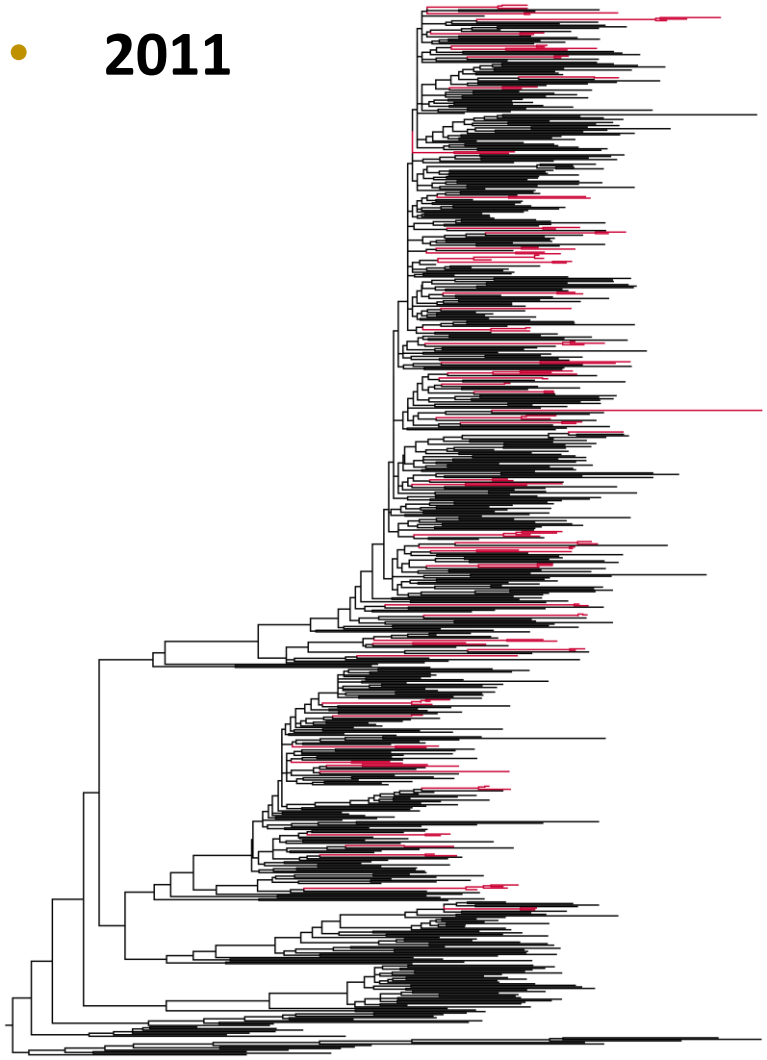


N = 467 séquences

29 clusters (61 séquences – 13%)

Taille du cluster	*nbre cluster*
2	24 (83%)
3	3 (10%)
4	1 (3%)

- **2011**



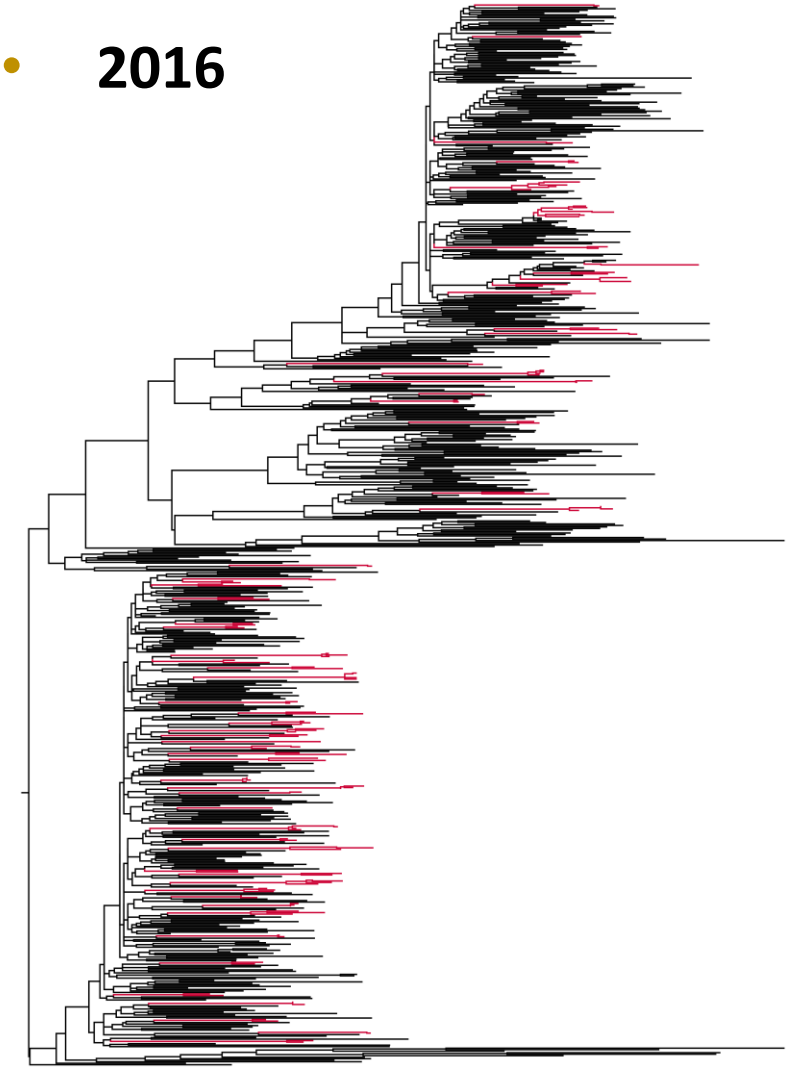
N = 681 séquences

56 clusters (129 séquences – 19%)

Taille du cluster	*nbre cluster*
2	43 (77%)
3	9 (16%)
4	4 (7%)

ODYSSEE

- **2016**



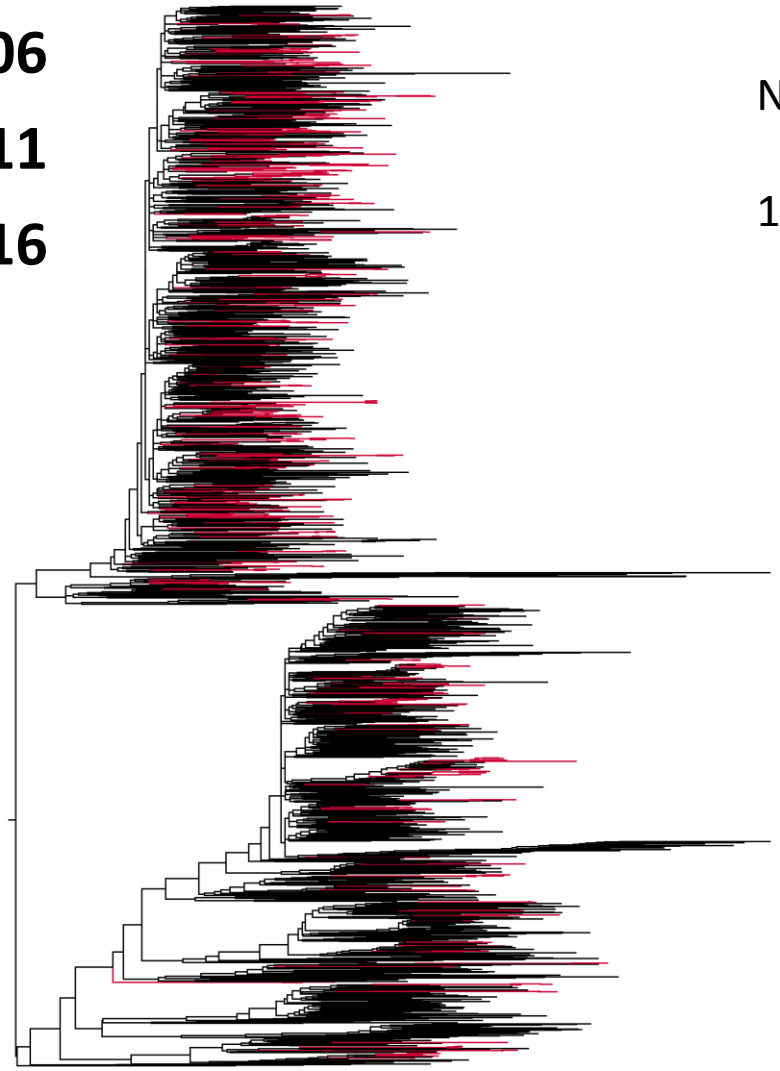
N = 705 séquences

65 clusters (163 séquences – 23%)

Taille du cluster	*nbre cluster*
2	43 (66%)
3	14 (22%)
4	6 (9%)
5	1 (2%)
6	1 (2%)

ODYSSEE – All in…

- **2006**
- **2011**
- **2016**

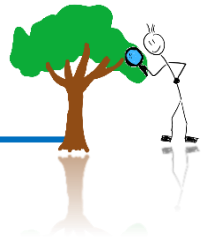


N = 1849 séquences

182 clusters (455 séquences – 25%)

Taille du cluster	*nbre cluster*
2	125 (69%)
3	37 (20%)
4	13 (7%)
5	3 (2%)
6	1 (1%)
7	3 (2%)

ODYSSEE – All in…



- Distribution temporelle des clusters larges :

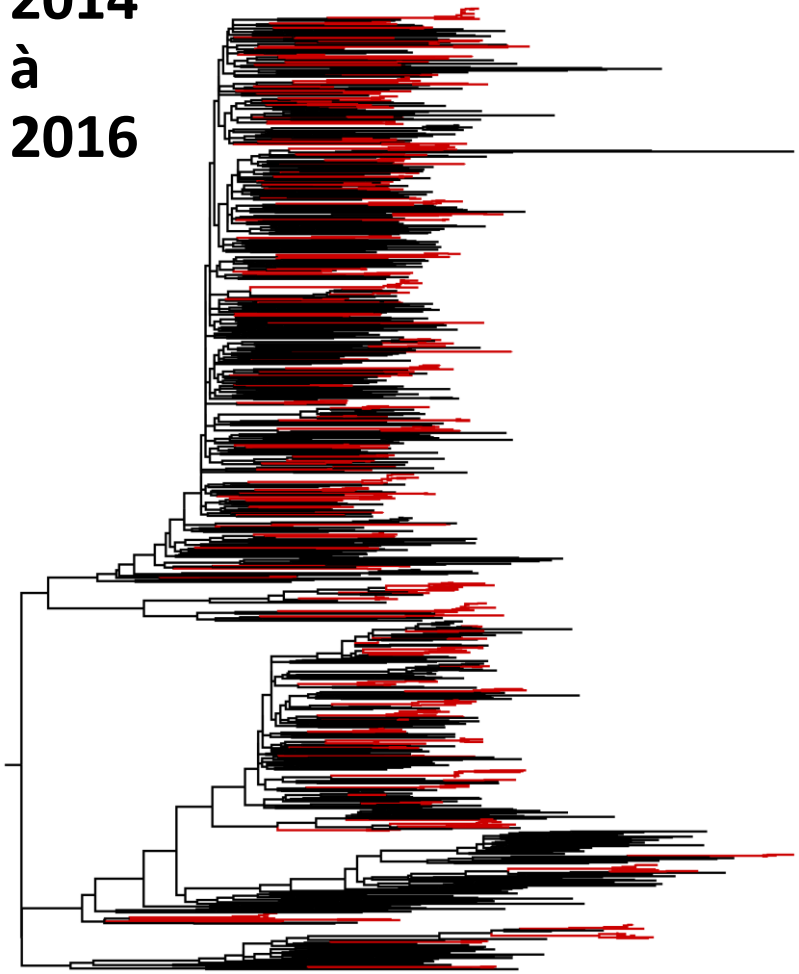
#	N séq.	2006	2011	2016
120	5	2	3	0
184	6	3	2	1
15	4	1	2	1
44	4	1	2	1
68	4	2	1	1
13	7	0	3	4
95	7	0	1	6
110	7	0	3	4
101	5	0	1	4
50	4	0	1	3
131	4	0	2	2
145	4	0	1	3
159	4	0	1	3
179	4	0	1	3
82	4	0	4	0
117	5	0	0	5
33	4	0	0	3
62	4	0	0	4
70	4	0	0	4
155	4	0	0	4

Clusters presque toujours étalés sur 5 à 10 ans

- Transmissions toujours actives ?
- Simple retour dans le circuit de soin d'infections anciennes ?

PRIMO

- **2014 à 2016**



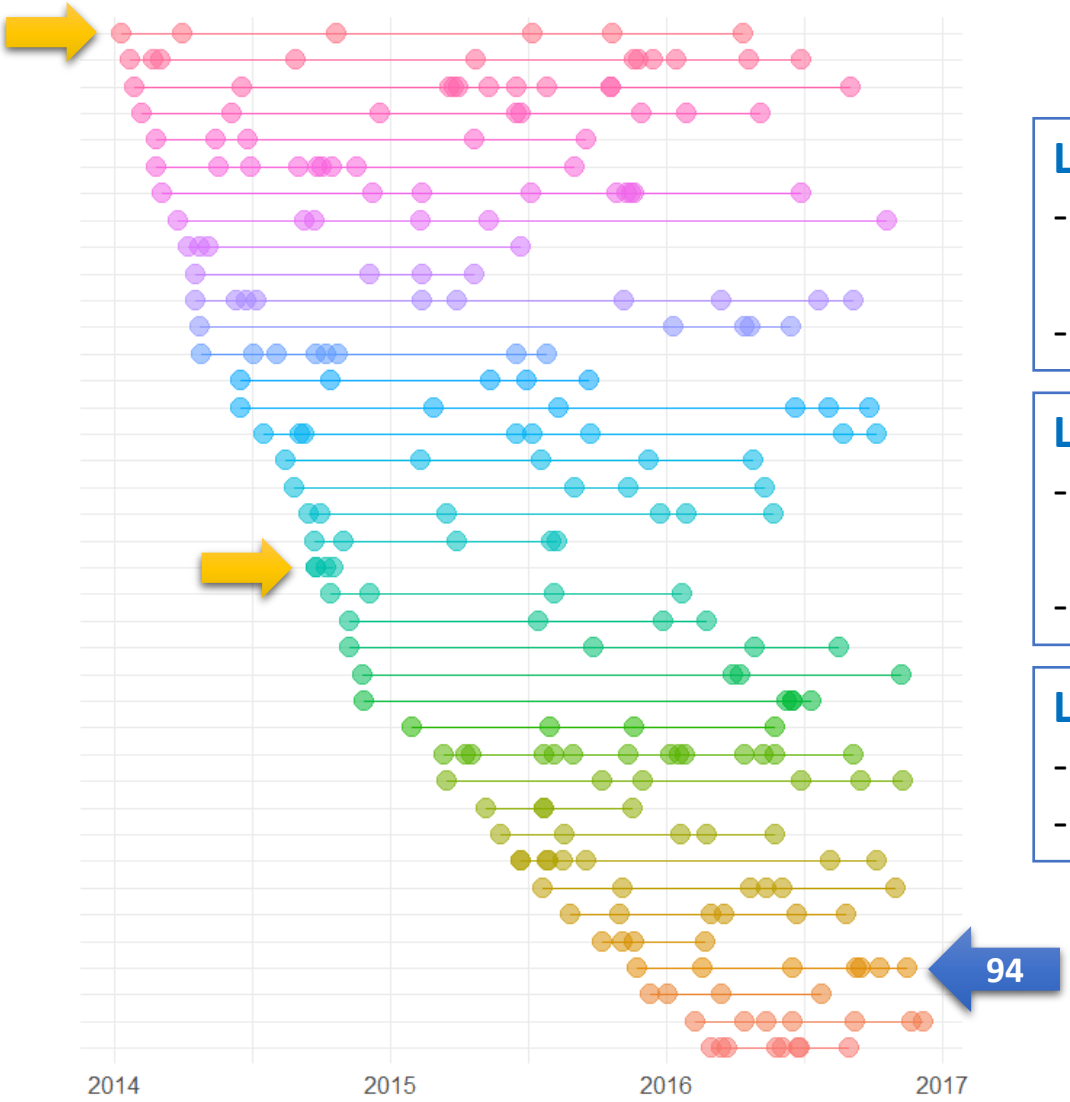
N = 1121 séquences

157 clusters (665 séquences – 59%)

Taille du cluster	*nbre cluster*
2	84 (55%)
3	22 (14%)
4	11 (7%)
5	8 (5%)
6	7 (4%)
7 - CRF94	2 (1%)
8	5 (3%)
9	1 (1%)
10	1 (1%)
11	2 (1%)
13	1 (1%)
14	1 (1%)

39 clusters (25% des clusters)
248 séquences (37% des séquences)

PRIMO – distribution temporelle

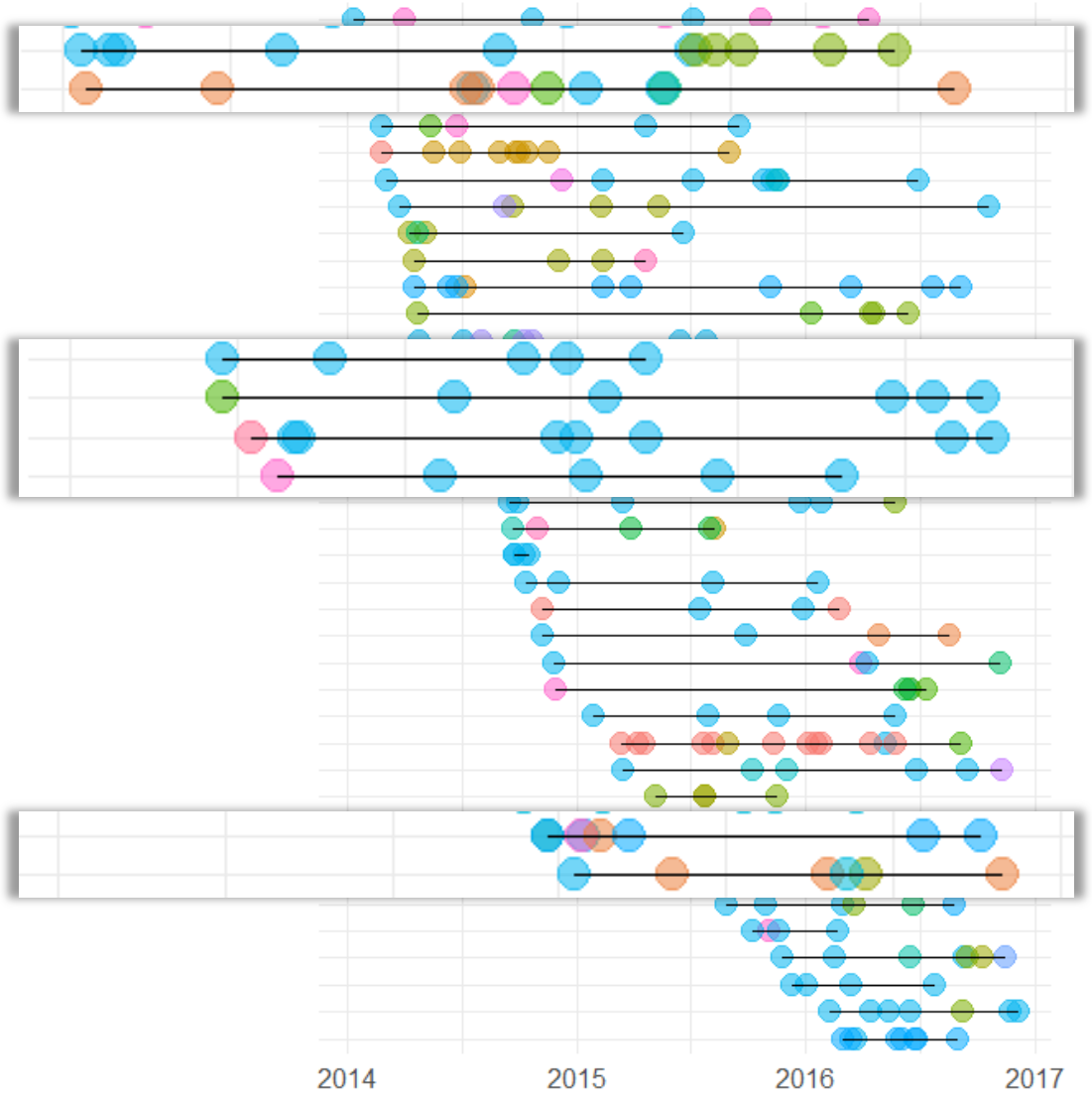


Le CRF94 :
- 34 patients fin 2016
- 7 présents ici

Le CRF98 :
- 8 patients à Bordeaux
- 2 dans cette base

Le cluster breton :
- 35 patients
- 0 ici

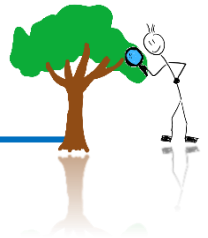
PRIMO – distribution géographique



Departement

- Bas-Rhin
- Bouches-du-Rh"ne
- Doubs
- FinistŠre
- Gard
- Gironde
- Haute-Garonne
- Herault
- Ille-et-Vilaine
- Indre-et-Loire
- IsŠre
- Loire-Atlantique
- Martinique
- Nord
- Paris
- Rh"ne
- Seine-et-Marne
- Seine-Maritime
- Seine-Saint-Denis
- Somme
- Val d'Oise
- Val de Marne
- Vienne
- Yvelines

Quelques leçons de ces arbres

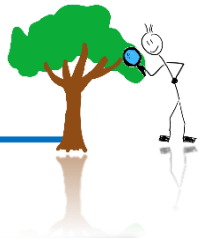


- **ODYSSEE :**
  - La part des clusters de transmissions récents augmente régulièrement
  - Les grands clusters se propagent sur de longues périodes temporelles

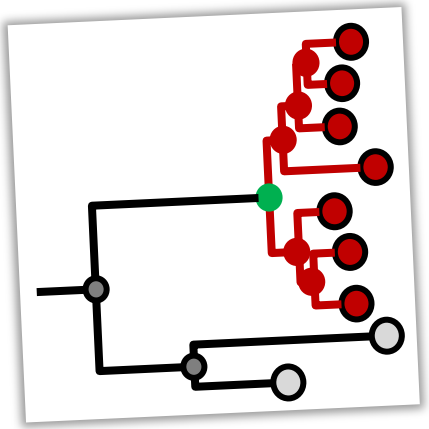
- **PRIMO :**
  - Part de clusters très importante dans cette population
  - Collecte très partielle et sous-estimation de l'importance réelle des clusters (CRF94, 98, cluster breton…)
  - Permet de détecter 39 « grands » clusters (4 à 14 patients)

Identifier ces clusters…
Pourquoi faire ?
Avec quels outils ?

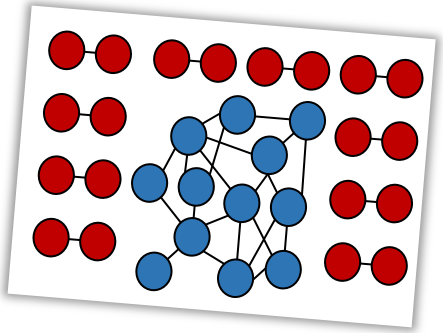
Des clusters ? Pour quoi faire ???



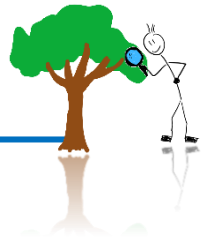
- Il est **de plus en plus simple** d'identifier ces clusters
  (bioinformatique de plus en plus simple et rapide)
- **Difficultés d'exhaustivité**… seulement possible à travers une bases nationales (ou plus)
  (enquêtes locales + études nationales, eDO, base nationale ?)



- **Action(s) possible(s) :**
  - **Etudes descriptives ➔ décisions de santé publique**
    (lieux, populations, caractéristiques virologiques des souches…)
  - **Identification de lieux** avec pratiques à risques
    ➔ actions locales
    - Action de prévention
      (promotion du dépistage, des préservatifs, de la PreP…)
    - Action de rattrapage
      (identification des populations, déclaration aux partenaires ?)
  - **Ne rien faire ?**



Et que font nos voisins ?



- **Pays avec une base nationale de séquences :**
  - Angleterre
  - Pays-Bas
  - Suisse
  - qq états aux USA

Epidémiologie	Résistance	Intervention
Epidémiologie	Résistance	Intervention
Epidémiologie	Résistance	Intervention
Epidémiologie	Résistance	Intervention

- **Pays avec projets affichés en cours de construction :**
  - USA (une base pour tous)
  - Mexique
  - (…)

➔ Projets à visée interventionnelle !!!

Une base nationale Française…

Constitution d'un groupe de travail pour réfléchir à la mise en place d'une base nationale sous l'égide de l'ANRS/Inserm et Santé Publique France

Investigateurs : L. Assoumou, F. Cazein, A. Chaillon, **M.L. Chaix,** J. Ghosn, F. Lot, **B. Visseaux,** M. Wirden

Objectif principal : réunir la première séquence effectuée pour chaque patient nouvellement diagnostiquée et agrégée avec quelques données épidémiologiques, pour…

- **Identifier rapidement les nouveaux clusters** de transmissions récentes et lancer des alertes
- **Renforcer la description épidémiologique des évolutions l'épidémie** par les données moléculaires

Une base nationale…

Laboratoires de Virologie (AC 43)

ANRS - Base transmission

- Identifiant
- Date prélèvement
- Séquences
- Données épidémiologiques
- Données sérologiques/virologiques

Santé publique France

COREVIH

Santé publique France

Santé publique France

e-do maladies à déclaration obligatoire

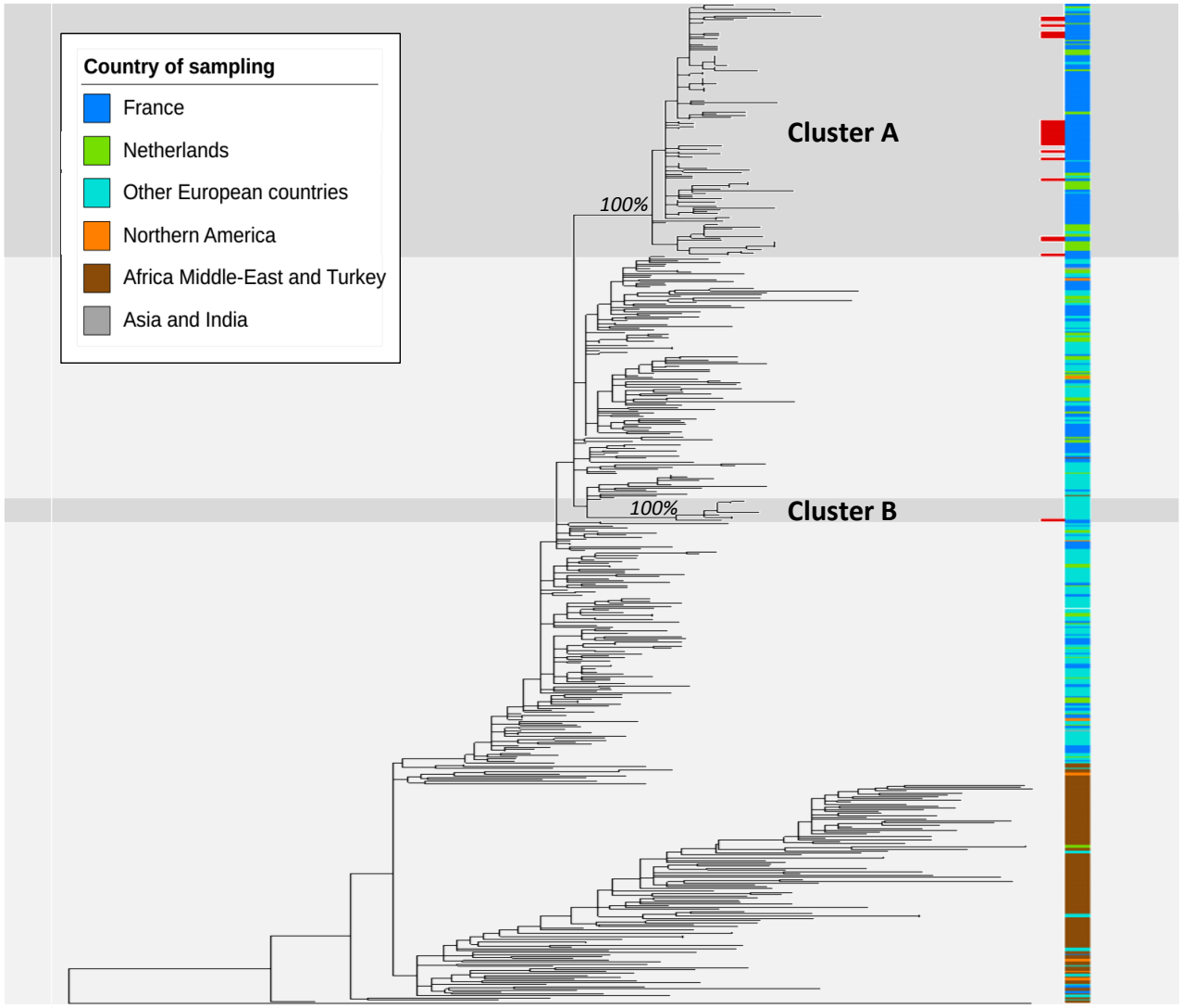
anrs France REcherche Nord & sud Sida-hiv Hépatites
Agence autonome de l'Inserm

Identifier les clusters significatifs « en temps réel »

Décrire la dynamique de l'épidémie sur des données consolidées

Une base nationale…

- 2003-2006 : Epidémie de VHC 4d parmi les MSM VIH+ à Paris



/ POUR LA VIE /
/ POUR UN WEEK-END /
/ POUR UNE FOIS /
Les situations varient.
Les modes de protection aussi.
Sexe entre hommes : trouvez vos modes de protection sur sexosafe.fr
PRÉSERVATIFS / DÉPISTAGES / PREP / TPE / TASP
Santé publique France
/ COUP DE FOUDRE /
/ COUP D'ESSAI /
/ COUP D'UN SOIR /
Les situations varient.
Les modes de protection aussi.
Sexe entre hommes : trouvez vos modes de protection sur sexosafe.fr
PRÉSERVATIFS / DÉPISTAGES / PREP / TPE / TASP
Santé publique France
/ S'AIMER /
/ S'ÉCLATER /
/ S'OUBLIER /
Les situations varient.
Les modes de protection aussi.
Sexe entre hommes : trouvez vos modes de protection sur sexosafe.fr
PRÉSERVATIFS / DÉPISTAGES / PREP / TPE / TASP
Santé publique France
/ AVEC UN AMANT /
/ AVEC UN AMI /
/ AVEC UN INCONNU /
Les situations varient.
Les modes de protection aussi.
Sexe entre hommes : trouvez vos modes de protection sur sexosafe.fr
PRÉSERVATIFS / DÉPISTAGES / PREP / TPE / TASP
Santé publique France

Bichat Claude Bernard

Virologie

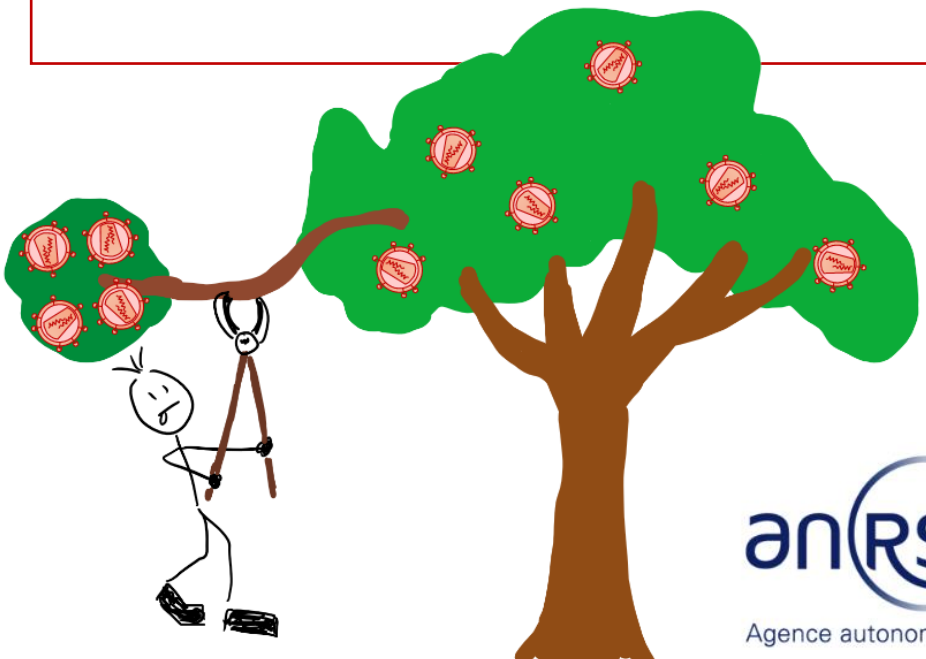
Mélanie Bertine
Gilles Collin
Pr Diane Descamps
Pr Charlotte Charpentier
Dr Florence Damond
Dr Nadhira Houhou-Fidouh
Dr Houria Ichou
Dr Lucile Larrouy
Dr Quentin Le Hingrat
Dr Vincent Mackiewicz
Dr Marine Perrier
Alexandre Storto

Maladies Infectieuses

Pr Yazdan Yazdanpanah
Dr Jade Ghosn
Dr Véronique Joly
Dr Roland Landman
Dr Sylvie Lariven
Dr Xavier Lescure
Pr Sophie Matheron
Dr Christophe Rioux

Pharmacologie

Dr Gilles Peytavin
Dr Minh Le



Pitié-salpêtrière

Virologie

Pr Vincent Calvez
Pr Anne-Geneviève Marcelin
Dr Marc Wirden

Méthodologie

Lambert Assoumou

Virologie - Rouen

F De Olivera
Pr JC Plantier

Virologie - Rennes

Dr Charlotte Pronier
Dr Anne Maillard

Virologie – Saint Louis

Dr Marie-Laure Chaix
Pr Constance Delaugerre

anrs France REcherche Nord & sud Sida-hiv Hépatites
Agence autonome de l'Inserm

iame UMR 1137

université PARIS DIDEROT PARIS 7

Instituts thématiques Inserm Institut national de la santé et de la recherche médicale

ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS