



Changement des profils génétiques des souches de VIH-1 au Sénégal

Mame Salane Thiam BASSE
LBV_HALD/CHAN

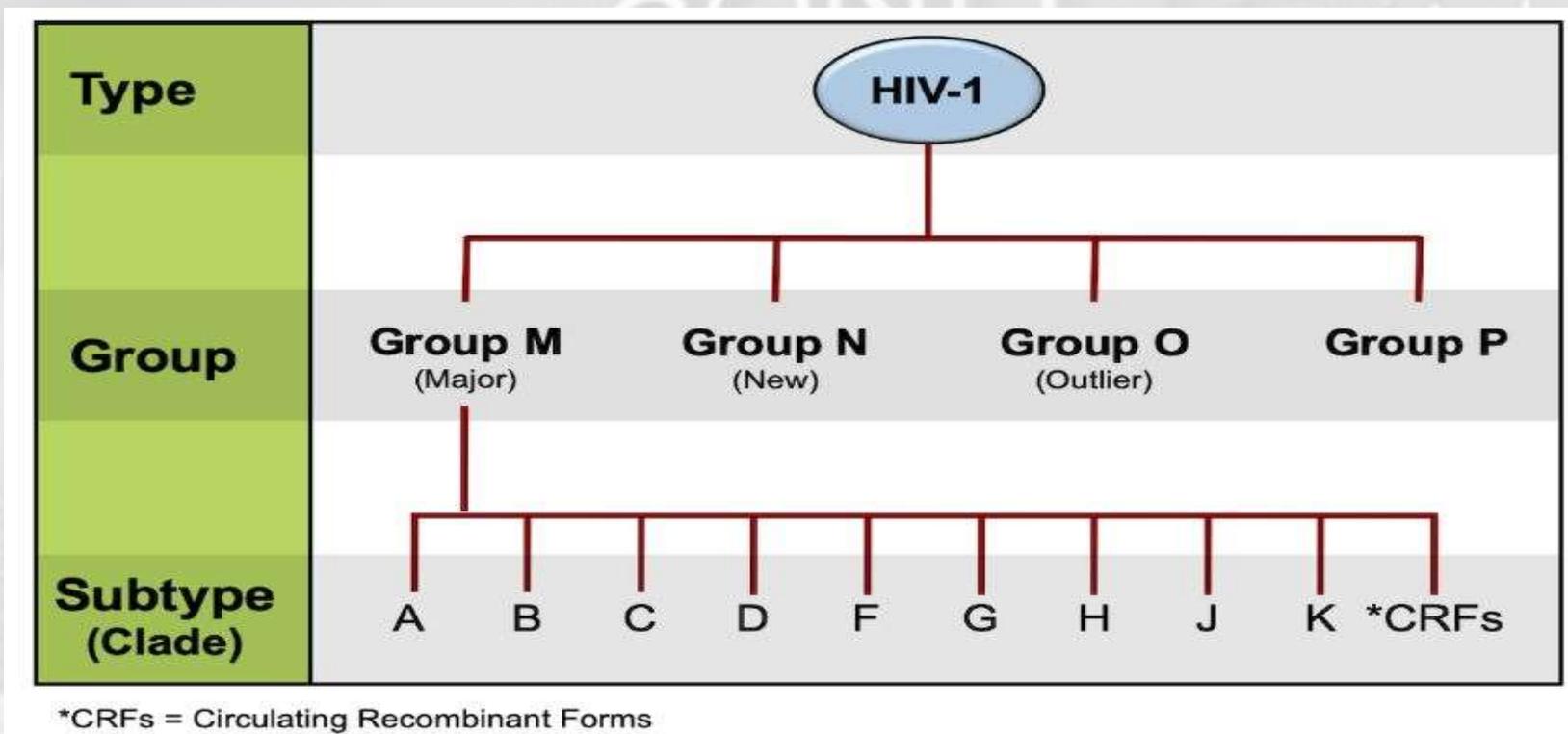
Plan

- Contexte
- Objectif
- Méthodologie
- Résultats
- Discussion
- Conclusion
- Remerciements

Contexte

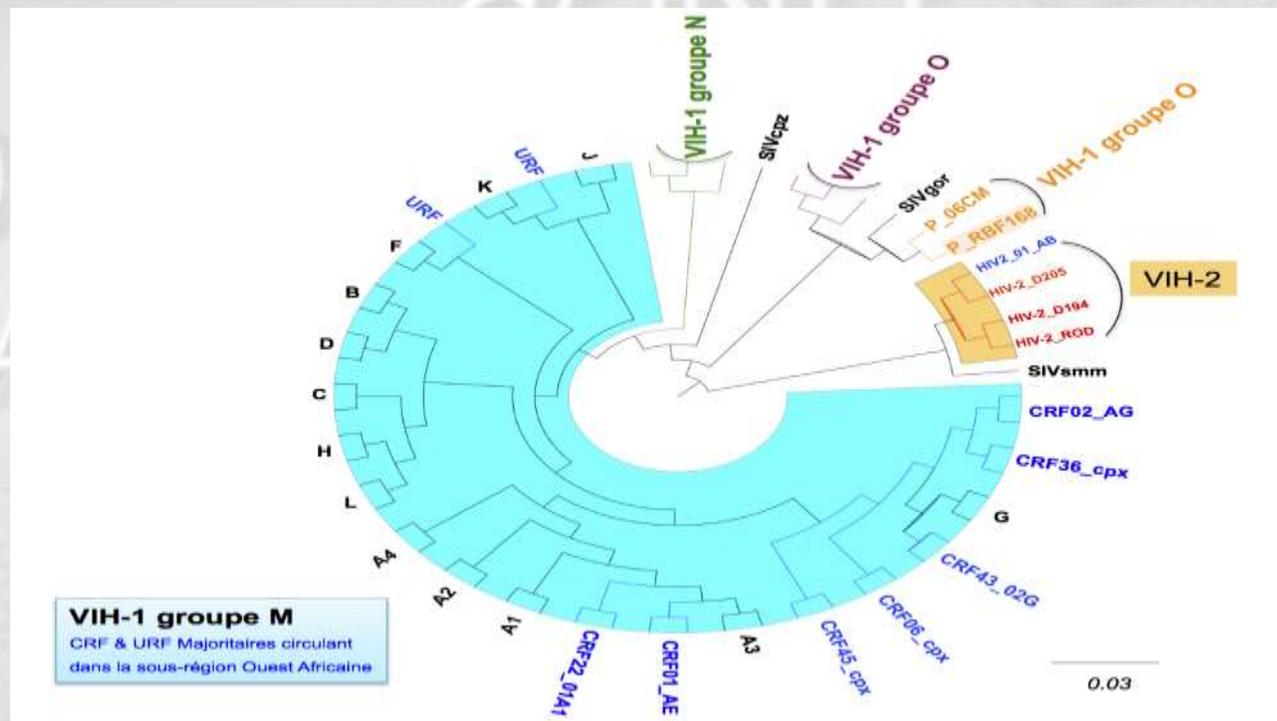
Diversité génétique du VIH

- A l'échelle mondiale



Contexte(I)

- En Afrique de l'Ouest



Contexte (II)

Au Sénégal

- ❑ Grande diversité génétique documentée depuis les années 2000
- ❑ Prédominance du CRF02_AG dans la Pop. Générale (**Harr F Njai et al., 2021**)
- ❑ Prédominance du sous-type C avec une diversité intra sous-types chez les HSH (**H.Diop-Ndiaye et al., 2013, N.Leye et al., 2013**)
- ❑ Complexité des souches prédominantes (CRF02_AG et C) avec co-circulation d'autres souches (**A.Bangue et al., 2019**)
- ❑ Scénarii de co et super-infections également documentées (**J. BOUGON et al., 2021**) favorisant la diversité génétique

Contexte (III)

Cette diversité génétique résulte

- Du taux élevé de la réplication couplée à une faible fidélité de la transcriptase inverse;
- Des phénomènes de recombinaison qui s'interfèrent en présence de deux virus,
- De l'hétérogénéité observée de la prévalence du VIH-1,

Contexte (IV)

Intérêt et nécessité d'étudier la diversité génétique

- ❑ Précision et amélioration outils de diagnostic biologique
- ❑ Choix des molécules dans le TARV (ex: Profil VIH-2 et Groupe O)
- ❑ Documentation de l'évolution et la distribution des souches
- ❑ Etude des réseaux de transmission pour améliorer les stratégies de prévention ciblées (Pop HSH par ex.)
- ❑ Cependant, il existe peu de données sur l'évolution et la répartition des sous-types, sous-sous-types du VIH-1 et les proportions des CRFs au fil des ans, en particulier au Sénégal, où la diversité génétique du VIH-1 est remarquable

Objectif

Déterminer l'évolution de la diversité génétique des séquences des gènes *gag* et *env* du **VIH-1** au Sénégal de 1990 à 2020.

Méthodologie (I)

Collecte des séquences

- Recueillir les données à partir de la base de données publique de Los Alamos (<http://www.hiv.lanl.gov>) ;
- Renseigner le code alpha-2 du Sénégal (SN) pour le pays d'échantillonnage ;
- Trouver toutes les séquences pour un gène ou une région spécifique (VIH-1, SIVcpz & SIVgor) ;
- Les régions génomiques d'intérêt sont : *gag* [790-1921], et *env* [6225-8250] relativement à la séquence de référence (HXB2).

Méthodologie(II)

Filtrage des séquences

- ❑ Un scripte **Python** a été utilisé pour transformer les fichiers **fasta** téléchargés en fichiers **CSV** ;
- ❑ Ensuite ces derniers sont importés dans une base de données **MySql** grâce à l'outil **Laragon**, avec les colonnes suivantes :
 - ✓ sous-types (type)
 - ✓ pays (country)
 - ✓ année de collecte (annee)
 - ✓ nom (id_patient)
 - ✓ accession et les séquences en acide nucléique (an).

Méthodologie(III)

Stockage des séquences

Scripte Python

```
def fastaToCsv(inputFileName, outputFileName):  
    testR = open(inputFileName, "r")  
    testW = open(outputFileName, "w")  
    testW.write("type;country;year;id;accession;an")  
    l = testR.readline()  
    while(l) :  
        if(l[0] == ">"):  
            l = l.replace("\n", ";")  
            l = l.replace(">", "\n>")  
            l = l.replace(".", ";")  
        else:  
            l = l.replace("\n", "")  
        testW.write(l)  
        l = testR.readline()  
  
    testR.close();  
    testW.close();  
  
fastaToCsv("newgag/GAGNONALIGNE1.fasta", "gag.csv")
```

Base de données avec l'outil Laragon

📌 id	type	annee	📌 id_patient	📌 accession	an
64	>D	2001	01SN_160HALD_12_3_01	AJ583742	aaagccggaacaggaaggacaa.
65	>A1	2000	00SN_69HPD_4_9_00	AJ583752	aaaggaaggacaccaaataaaa.
66	>B	2000	00SN_101hald_3_7_00	AJ583740	aaaggaaggacaccaaataaaa.
67	>U	2000	00SN_58HALD_25_9_00	AJ583736	aaaggaaggacaccaaataaaa.
68	>02_AG	2001	258HALD	FN599699	aacagactggaacccacagcccc.
69	>02_AG	2010	SN_5745HALD	HG424419	aacagccccaccagcagagagct.
70	>02_AG	2000	208HPD	FN599693	aacagccccccagcagagagcc.
71	>02_AG	2010	SN_10023HALD	HG424423	aagccagggatggatggcccaaa.
72	>02_AG	2000	00SN_46hald_6_6_00	AJ583747	aaggaaggacaccaaataaaa.
73	>02_AG	2008	SL160	HG315064	aataaaagcattgacagaaattg.
74	>DG	2011	S-512-JADY-J0	KU685097	aatttccattagtcctattgaaact.
75	>B	2010	SN_SL84	HG424408	aatttccaattagtcctattgaaac.
76	>A3	2010	S-024-GOSE-J0	KU685071	aatttccattagtcctataaaact.
77	>02_AG	2010	S-003-GAGY-J0	KU685057	aatttccattagtcctattgaaact.
78	>02_AG	2007	2881HA_RT	GQ351918	acaatggccattgacagaaagaaa.

Méthodologie(IV)

Stockage des séquences

- SELECT CONCAT(`type`,`;`,SUBSTR(annee, 3),`,`;`,accession`,`;`, an) FROM env WHERE **doublons != 1** and **type = "** and **annee != '-'** ORDER BY annee

Données filtrées avec séparateur point-virgule

id	type	annee	id_patient	accession	an
64	>D	2001	01SN_160HALD_12_3_01	AJ583742	aaagccggaacaggaaggacaa.
65	>A1	2000	00SN_69HPD_4_9_00	AJ583752	aaaggaaggacaccaaataaaa.
66	>B	2000	00SN_101hald_3_7_00	AJ583740	aaaggaaggacaccaaataaaa.
67	>U	2000	00SN_58HALD_25_9_00	AJ583736	aaaggaaggacaccaaataaaa.
68	>02_AG	2001	258HALD	FN599699	aacagactggaacccacagcccc.
69	>02_AG	2010	SN_5745HALD	HG424419	aacagccccaccagcagagagct.
70	>02_AG	2000	208HPD	FN599693	aacagccccccagcagagagcc.
71	>02_AG	2010	SN_10023HALD	HG424423	aagccagggatggatggcccaaa.
72	>02_AG	2000	00SN_46hald_6_6_00	AJ583747	aaggaaggacaccaaataaaa.
73	>02_AG	2008	SL160	HG315064	aataaaagcattgacagaaatttg.
74	>DG	2011	S-512-JADY-J0	KU685097	aatttccattagtctattgaaact.
75	>B	2010	SN_SL84	HG424408	aatttccaattagtctattgaaac.
76	>A3	2010	S-024-GOSE-J0	KU685071	aatttccattagtctattgaaact.
77	>02_AG	2010	S-003-GAGY-J0	KU685057	aatttccattagtctattgaaact.

```
CONCAT(`type`,`;`,SUBSTR(annee, 3),`,`;`,accession`,`;`, an)
>B;83;K03455;tggaagggctaattcactcccaacgaagacaagatatctctg
>D;90;AB485648;tggaagggctaatttggtccaagaaaagacaagagatc
>D;90;AB485649;tggatgggttaatttggtccaagaaaagacaagagatc
>C;90;AY713416;gagagagcgtcaatattaagaggggggaaattagattc
>C;90;AY833059;atgagagtgacggggatggtgaggaatttcaaccatgg
>C;90;JQ715373;atgagagtgacggggatggtgaggaattgtcaaccatgg
>C;90;L22944;ggatcatcaaaatcctatatcaaacagtaagtactacatag
>D;90;L22945;tgtttattcattcagaattgggtgtcaacatagcagaatagg
>09_cpx;94;AY093602;tttgaatgctgataaattgtgggtcacgggtctact
>C;94;AY249756;agaagaagaaataataattaggtctgaaaatagacaa
>02_AG;94;AY249757;gaatagtgattagatctgaaaatcacaaaca
>02_AG;94;AY249758;agaaggaagatagtgattagatctgaaaatag
>02_AG;94;AY249760;gattagatctgaaaatcacaaacaatgccaac
>02_AG;94;AY249761;agaaaaaaataatacaattatctgaaaatag
```

Méthodologie(V)

Stratification des séquences

- ✓ 1990-1994
- ✓ 1995-1999
- ✓ 2000-2004
- ✓ 2005-2009
- ✓ 2010-2014

```
SELECT
CONCAT(`type`, '.', annee, '.', id_patient, '.', accessio
n, CHAR(13), an) FROM gag WHERE annee BETWEEN 1990
AND 1994 AND doublons != 1 ORDER BY annee
SELECT
CONCAT(`type`, '.', annee, '.', id_patient, '.', accessio
n, CHAR(13), an) FROM gag WHERE annee BETWEEN 1995
AND 1999 AND doublons != 1 ORDER BY annee
```

Résultats (I)

❑ Collecte des séquences

- **286** séquences du gène *gag*
- **172** séquences du gène *env*

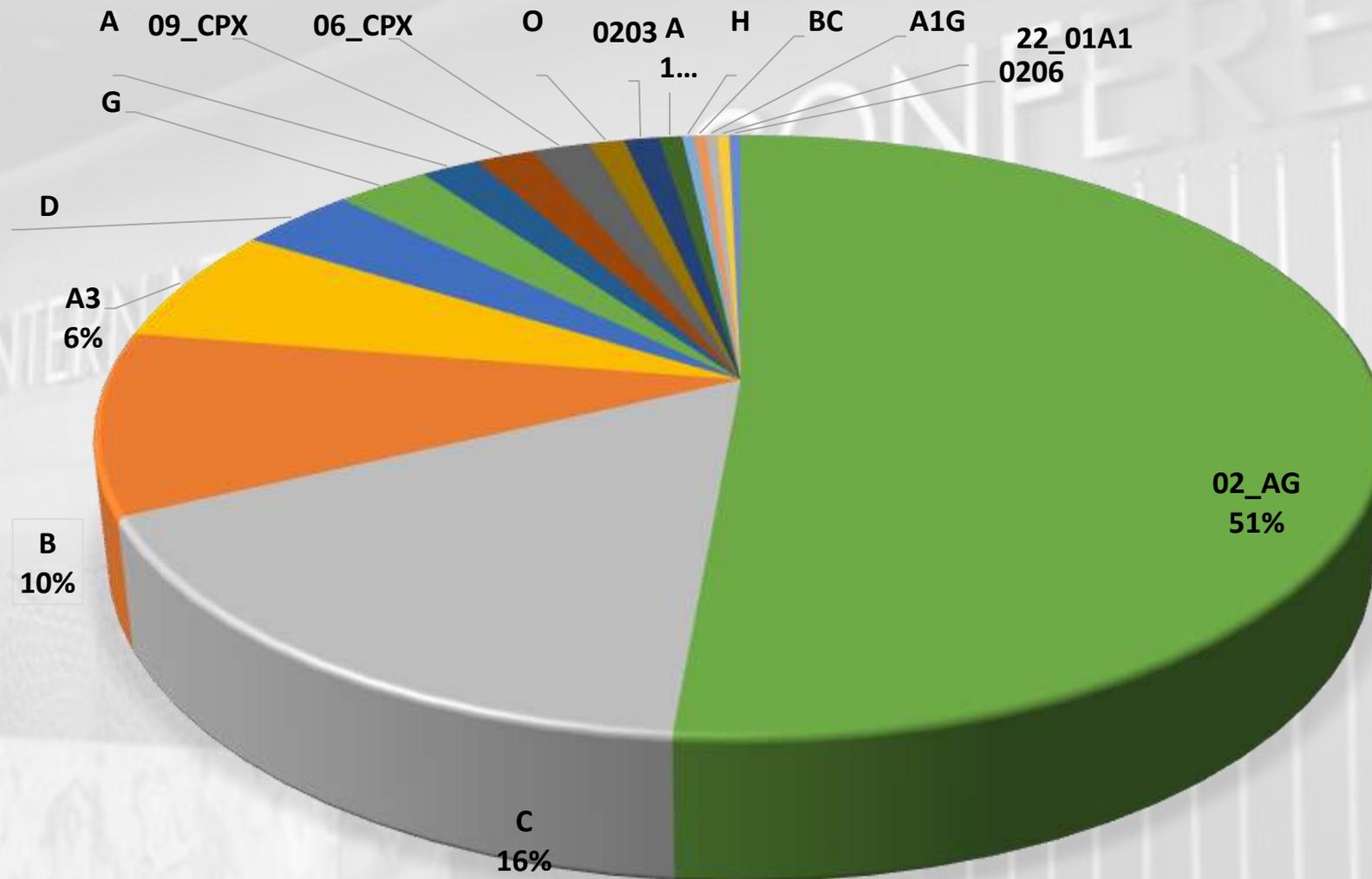
❑ Filtrage,

- **04** séquences de *gag*
- **03** séquences de *env*

Résultats (II)

DIVERSITE GENETIQUE

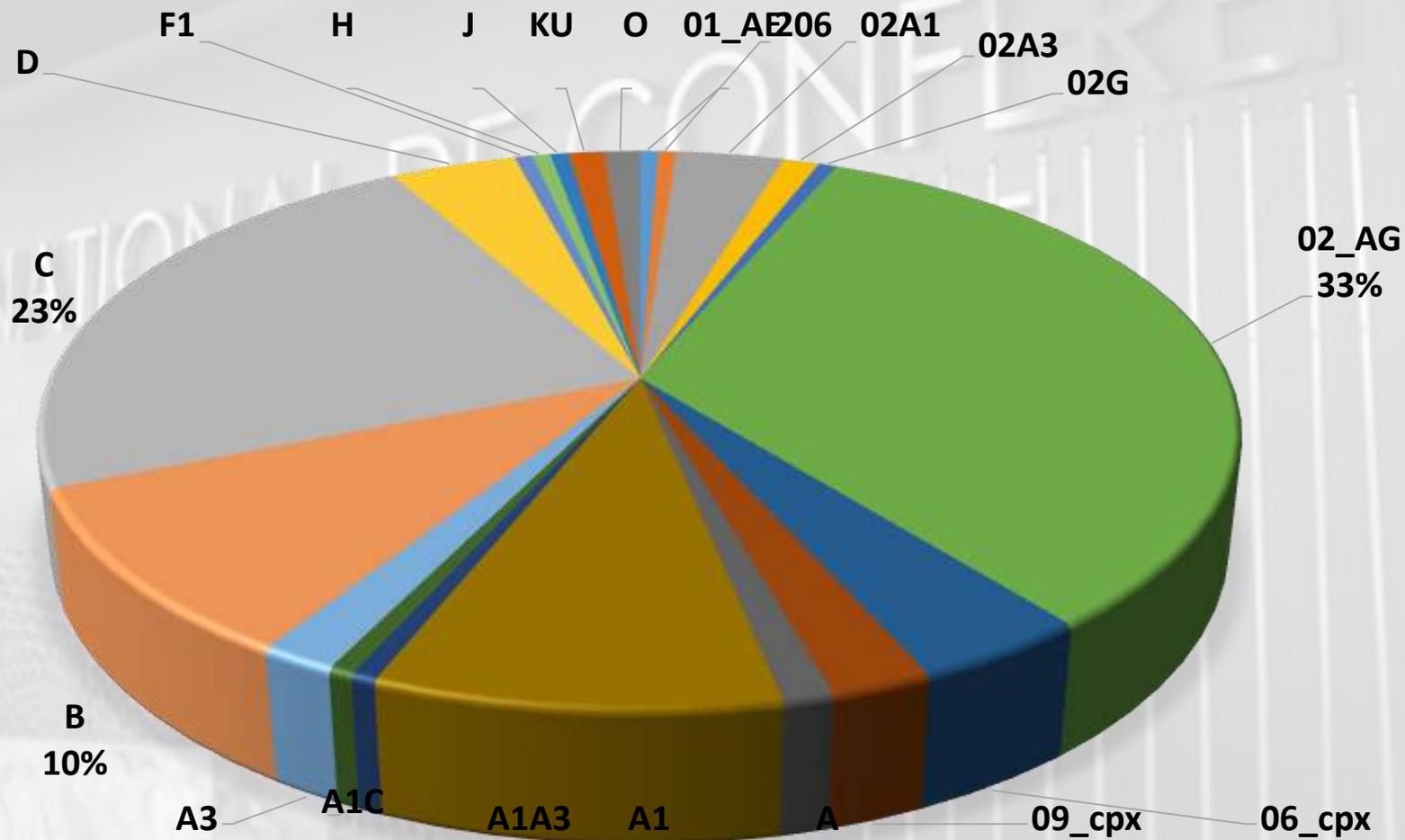
17 VARIANTS GAG



Résultats (III)

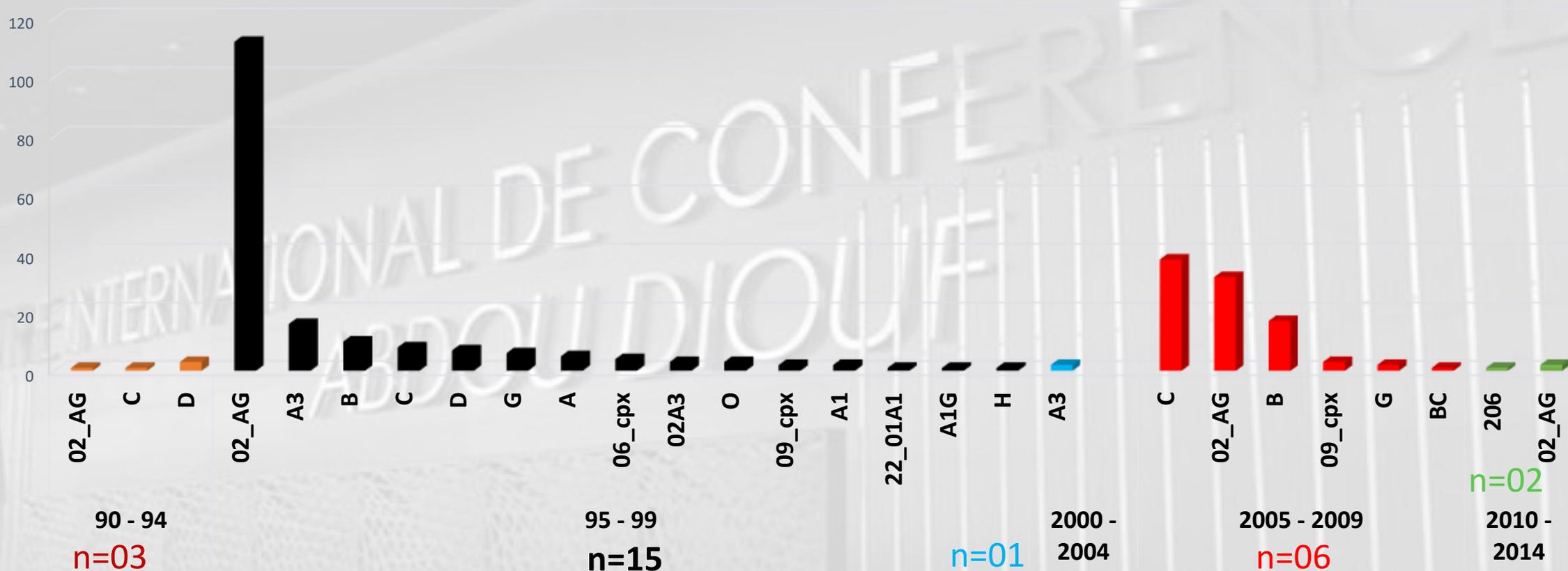
DIVERSITE GENETIQUE

21 VARIANTS ENV



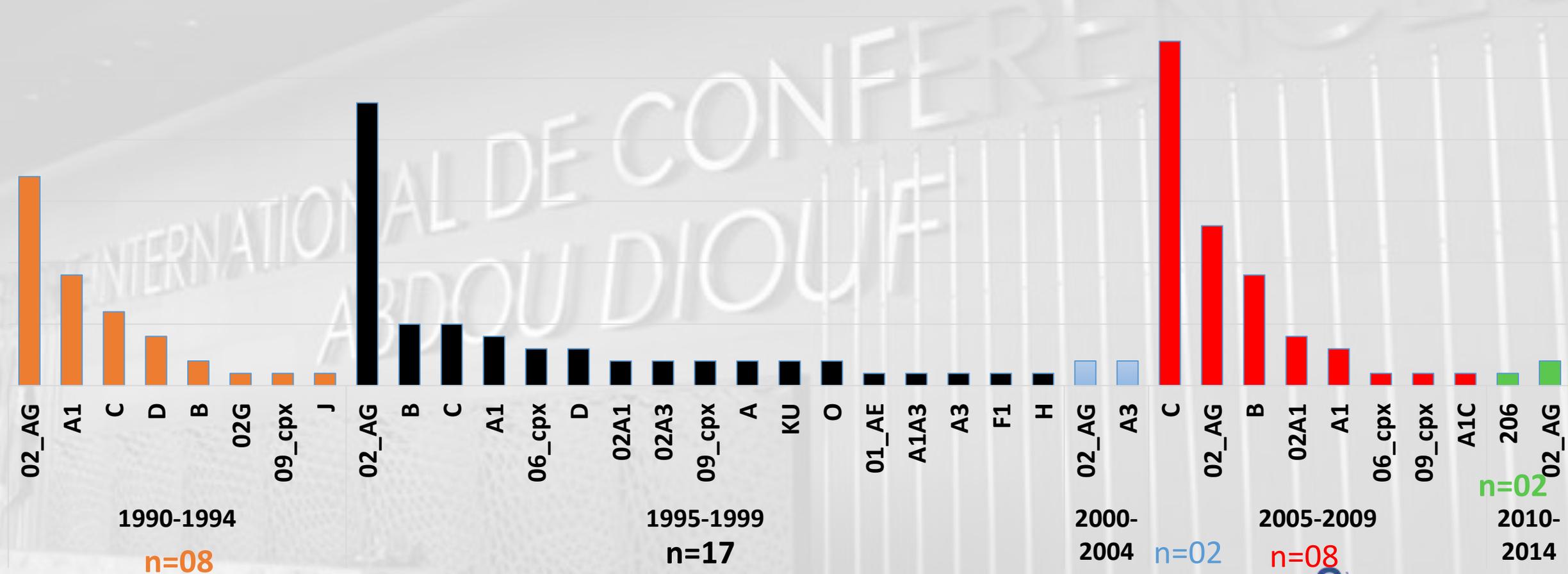
Résultats (IV)

Gène *Gag* du VIH-1 de 1990 - 2014



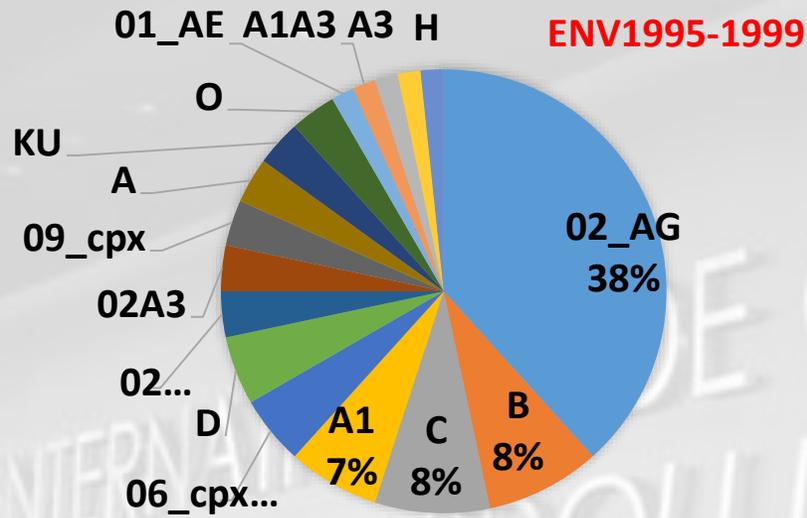
Résultats (V)

gène env du VIH-1 de 1990 à 2014

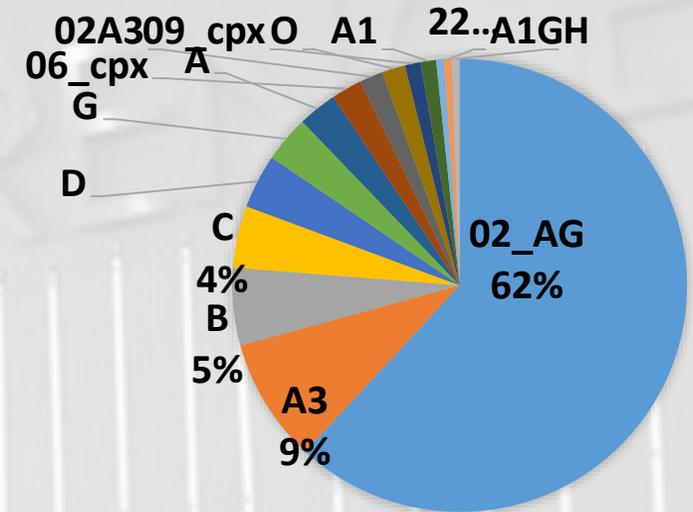


Résultats (VI)

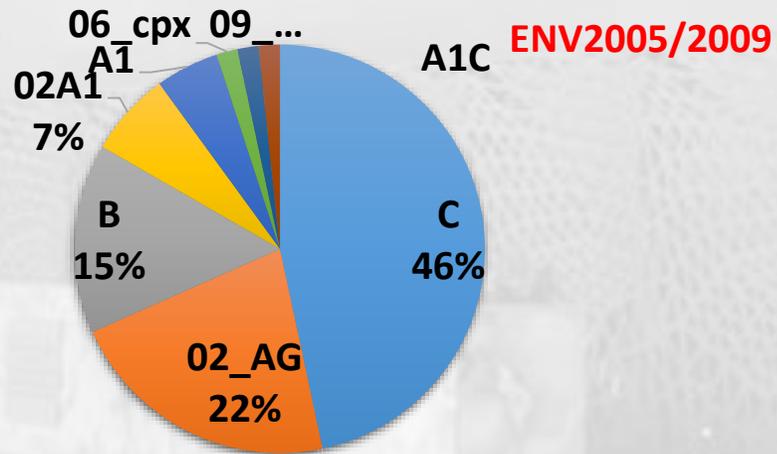
POPULATION EN GÉNÉRALE



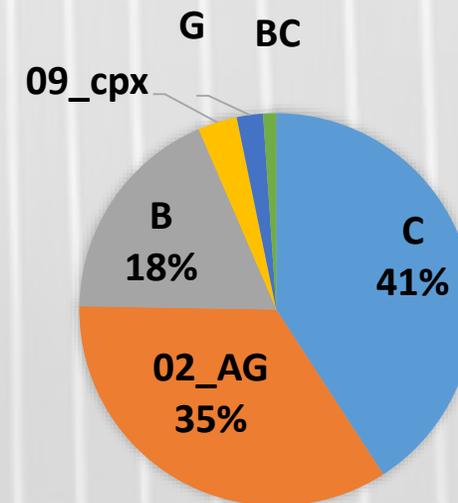
GAG1995-1999



POPULATION CHEZ LES HSH



GAG2005/2009



Discussion(I)

- ❑ De façon globale le CRF-02AG reste la forme recombinante circulante la plus répandue au Sénégal,
- ❑ Après stratification des données ,CRF-02AG dominer avec un bon pourcentage par rapport aux autres variants jusqu'en 2004,
- ❑ A partir de 2005 , l'analyse de la région génomique *env* et *gag* montre une baisse de la proportion de CRF02_AG au profit du sous-type C,

Discussion (II)

- ❑ C'est en 2005 que le Sénégal a enregistré plus de cas de VIH1 chez les Hommes ayant des rapports sexuels avec d'autres Hommes (HSH),
- ❑ D'après les données de la littérature, le sous type C est largement représenté dans le groupe des HSH au Sénégal (**H.Diop-Ndiaye et al .2013, N Leye et al., 2013**),

Conclusion

- ❑ Dans ce présent travail préliminaire, l'analyse des régions génomiques gag et env, a permis de confirmer la grande diversité génétique des gènes gag et *env* du VIH-1 isolés au Sénégal.
- ❑ Ces résultats doivent cependant être confirmés par l'analyse d'autres gènes en particulier, le gène *pol*.

Mame Salane Thiam^{1,2,3}, Abou Abdallah Malick Diouara², Mohammadou Lamine Dia¹, Aissatou Ahmet Niang¹, Fatoumata Diallo¹, Papa Adama Mboup¹, Mengué Fall³, Adja Aissatou Ba³, Ndeye Coumba Tourè¹, Abdoulaye Banire Diallo⁴, Halimatou Diop Ndiaye³

¹ Laboratoire de Bactériologie virologie du CHU Fann, – Université Cheikh Anta DIOP;

² Groupe de Recherche Biotechnologies Appliquées & Bioprocédés environnementaux (GRBA-BE), École Supérieure Polytechnique (ESP) – Université Cheikh Anta DIOP;

³ Laboratoire de Bactériologie Virologie CHU Aristide le Dantec – Université Cheikh Anta DIOP;

⁴ Laboratoire Bioinformatique, Université du Québec à Montréal - Montréal (Qc), Canada.

Merci de votre attention

