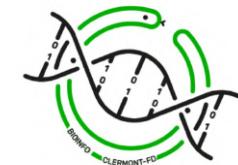


Stage de Master 1 Bio-Informatique 2024/2025

CONSTRUCTION DE PANGÉNOMES EN GRAPHE DU VIRUS DE L'HÉPATITE D

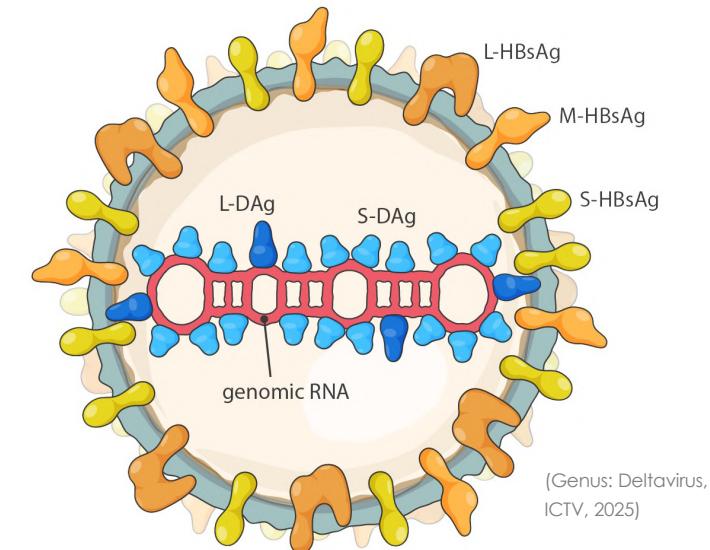
LAURA PLAYA PARIENTE



Encadrant : Xavier GRAND
UMR 1350 PaThLiv
équipe HeLip

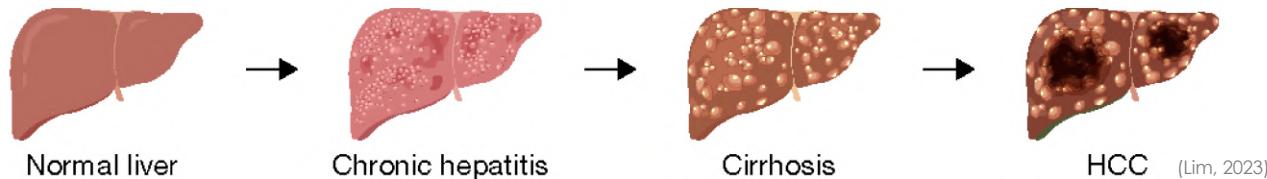
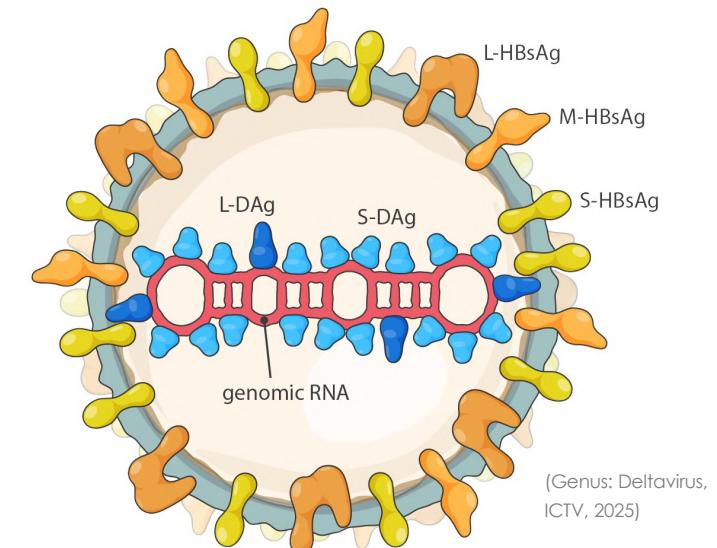
LE VIRUS DE L'HEPATITE D - VHD

- Virus satellite du virus de l'hépatite B
 - L'assemblage des virions nécessite les antigènes de surface du VHB



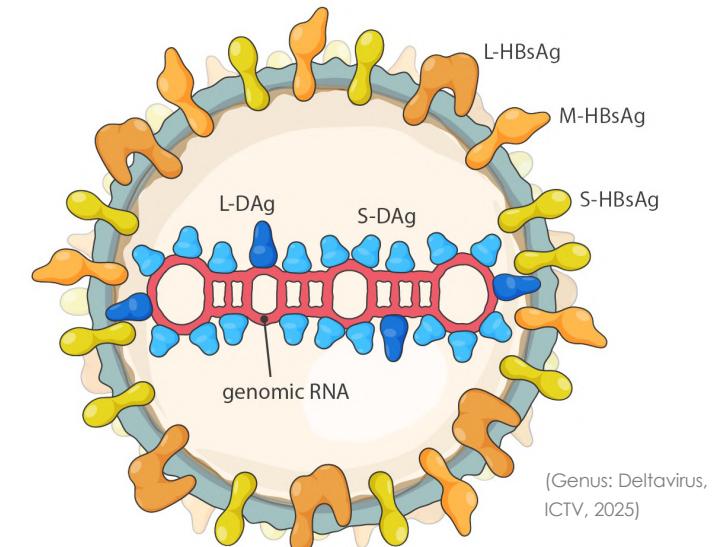
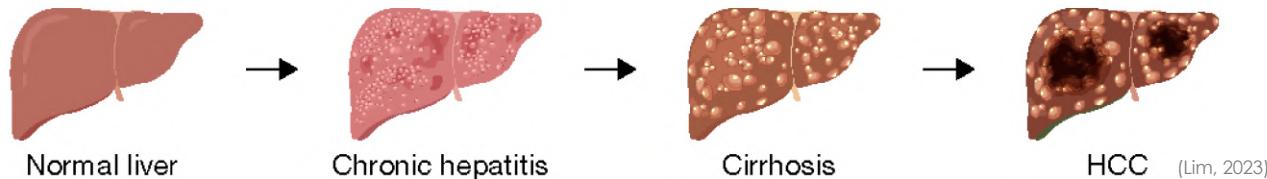
LE VIRUS DE L'HEPATITE D - VHD

- Virus satellite du virus de l'hépatite B
 - L'assemblage des virions nécessite les antigènes de surface du VHB
- Infection combinée VHB – VHD (OMS, 2025)
 - Plus grave des hépatites virales
 - Phase aiguë : risque accrue d'hépatite fulminante
 - **Phase chronique : progression accélérée vers cirrhose et carcinome hépatocellulaire**



LE VIRUS DE L'HEPATITE D - VHD

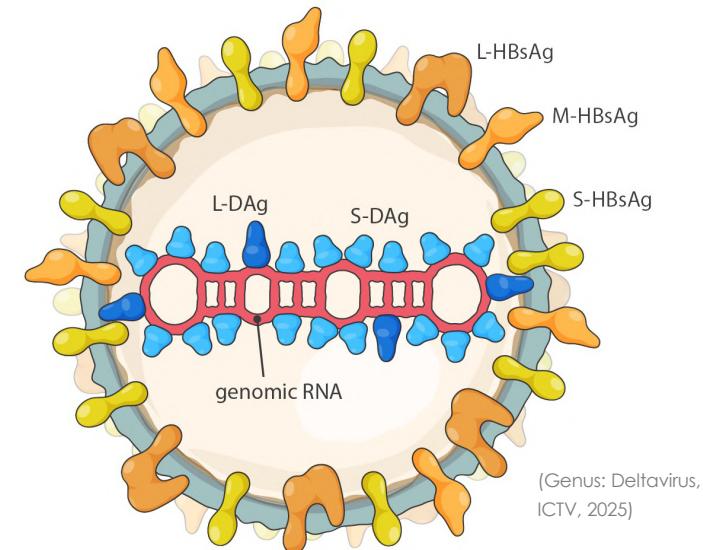
- Virus satellite du virus de l'hépatite B
 - L'assemblage des virions nécessite les antigènes de surface du VHB
- Infection combinée VHB – VHD (OMS, 2025)
 - Plus grave des hépatites virales
 - Phase aiguë : risque accrue d'hépatite fulminante
 - **Phase chronique : progression accélérée vers cirrhose et carcinome hépatocellulaire**



- 12 millions de personnes infectées dans le monde
- Faible efficacité des traitements disponibles, prévention possible par le vaccin VHB

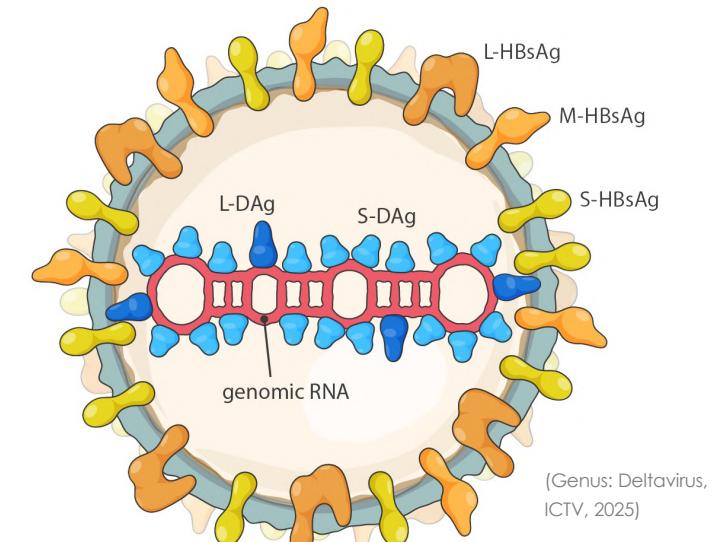
LE VIRUS DE L'HEPATITE D - VHD

- Virus satellite du virus de l'hépatite B
 - L'assemblage des virions nécessite les antigènes de surface du VHB
- Génome circulaire ARN simple brin polarité négative 1.7 kb
- 70 % de bases complémentaires => forme de bâtonnet



LE VIRUS DE L'HEPATITE D - VHD

- Virus satellite du virus de l'hépatite B
 - L'assemblage des virions nécessite les antigènes de surface du VHB
- Génome circulaire ARN simple brin polarité négative 1.7 kb
- 70 % de bases complémentaires => forme de bâtonnet
- Une seule ORF code 2 isoformes de l'antigène

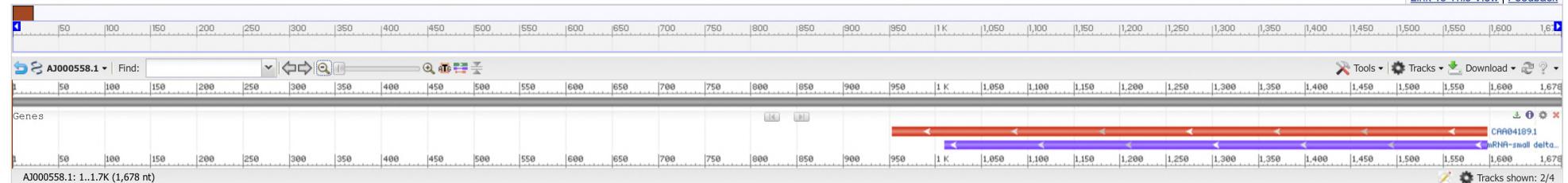


Hepatitis D Virus complete genome

GenBank: AJ000558.1

[GenBank](#) [FASTA](#)

[Link To This View](#) | [Feedback](#)

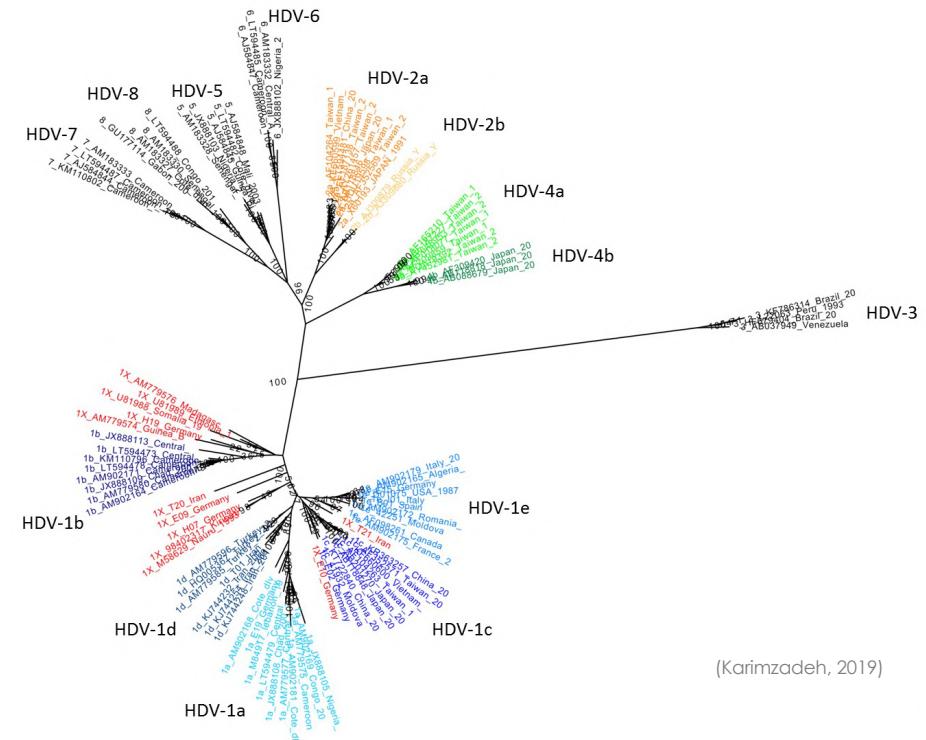


<https://www.ncbi.nlm.nih.gov/nuccore/AJ000558.1?report=graph>



LE VIRUS DE L'HEPATITE D - VHD

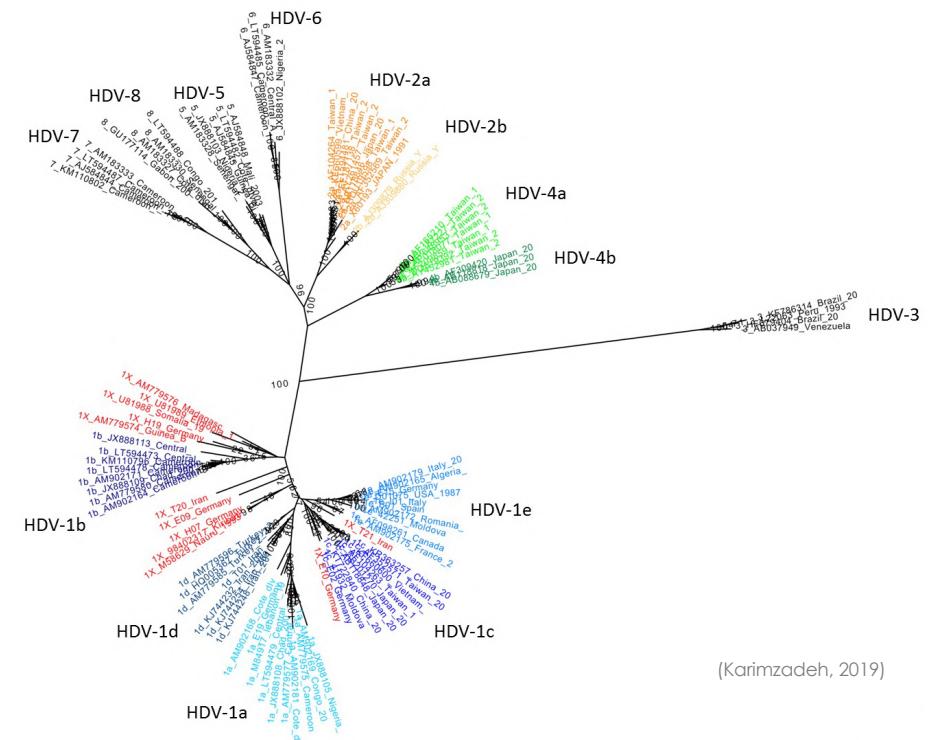
- 8 génotypes ($\geq 80\%$ d'identité sur le génome complet)
 - Subdivisés en sous-types ($\geq 90\%$ d'identité)



(Karimzadeh, 2019)

LE VIRUS DE L'HEPATITE D - VHD

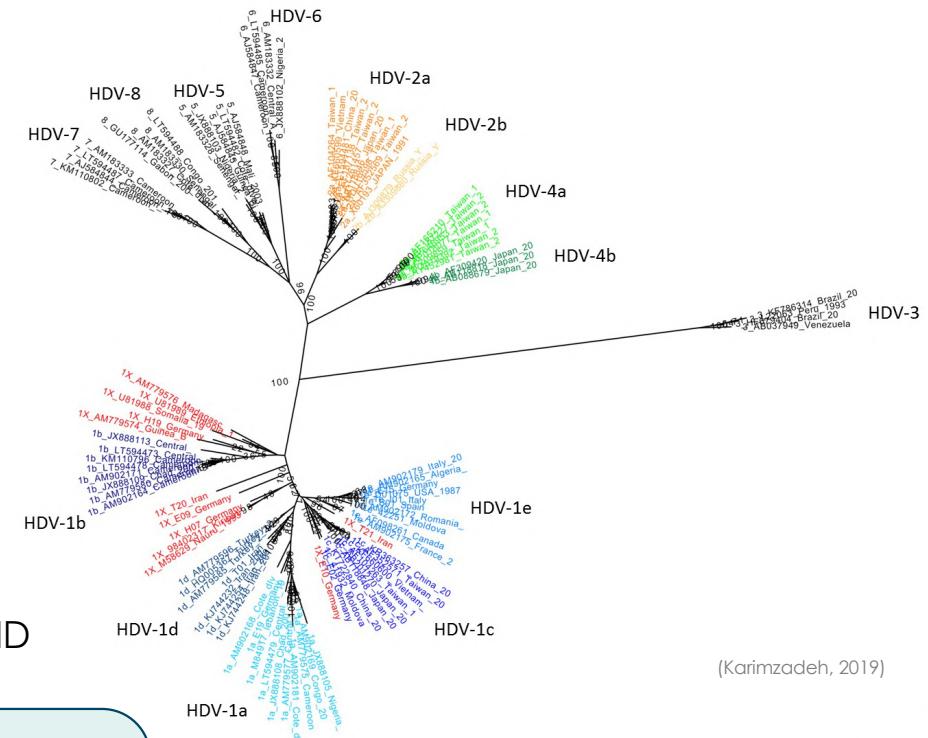
- 8 génotypes ($\geq 80\%$ d'identité sur le génome complet)
 - Subdivisés en sous-types ($\geq 90\%$ d'identité)
- Différences remarquables entre génotypes
 - Réplication in vitro
 - Évolution clinique
 - Réponse aux traitements



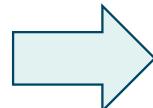
(Karimzadeh, 2019)

LE VIRUS DE L'HEPATITE D - VHD

- 8 génotypes ($\geq 80\%$ d'identité sur le génome complet)
 - Subdivisés en sous-types ($\geq 90\%$ d'identité)
- Différences remarquables entre génotypes
 - Réplication in vitro
 - Évolution clinique
 - Réponse aux traitements
- Absence de référence universelle et de ressource centralisée pour l'étude de la diversité génomique du VHD



(Karimzadeh, 2019)



Projet de création d'un centre de ressources de référence pour les génomes des virus hépatiques regroupant tous les génomes complets sous forme de **pangénomes en graphe**



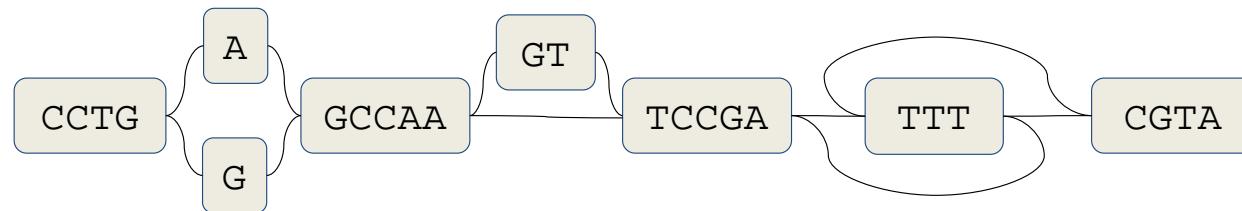


PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCCA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCAA--TCCGA**TTT**CGTA

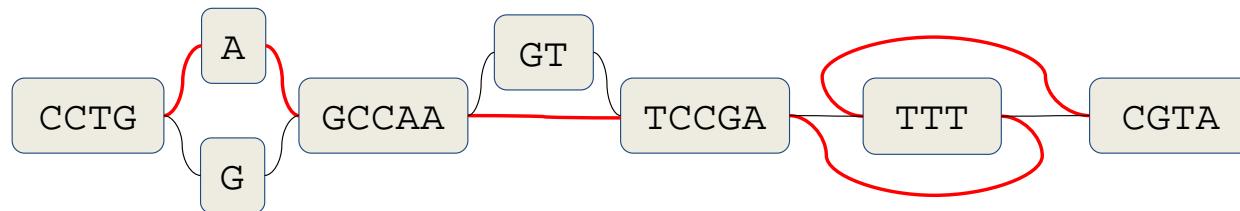
PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCAA--TCCGA**TTT**CGTA



PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCAA--TCCGA**TTT**CGTA

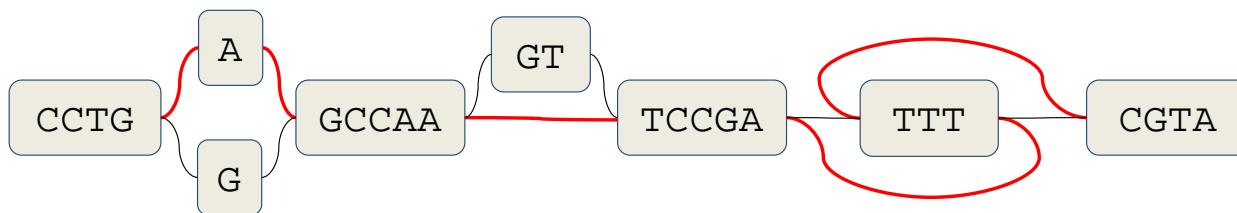


Chaque génome est un chemin sur le graphe

PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCAA--TCCGA**TTT**CGTA

- Représentation compacte
 - Parties conservées collapses dans un seul nœud
- Contient toute la variabilité génomique
 - Permet d'aligner davantage de lectures dans les régions variables sans biais de référence



Chaque génome est un chemin sur le graphe

Introduction

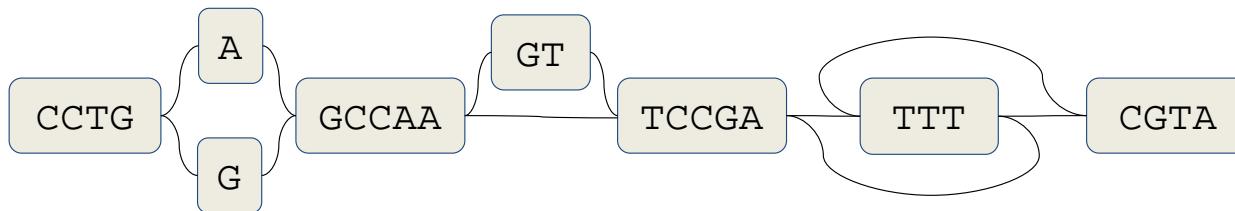
Matériel et méthodes

Résultats

Conclusion

PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCAA--TCCGA**TTT**CGTA

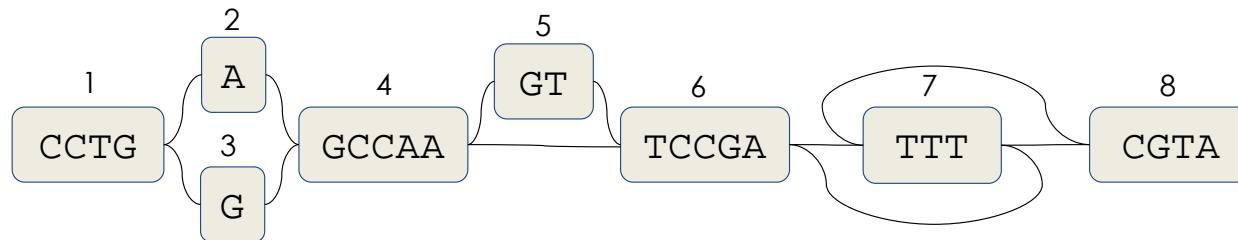


graphe.gfa			
H	VN:Z:1.0	CCTG	
S	1	A	
S	2	G	
S	3	GCCAA	
S	4	GT	
S	5	TCCGA	
S	6	TTT	
S	7	CGTA	
L	1	+	2
L	1	+	3
L	2	+	4
L	3	+	4
L	4	+	5
L	4	+	6
L	5	+	6
L	6	+	7
L	6	+	7
L	7	+	8
L	7	-	8
P	(A)	1+,2+,4+,6+,7-,8+	
P	(B)	1+,2+,4+,5+,6+,7+,8+	
P	(C)	1+,3+,4+,5+,6+,7-,8+	
P	(D)	1+,3+,4+,6+,7+,8+	

PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCA--TCCGA**TTT**CGTA

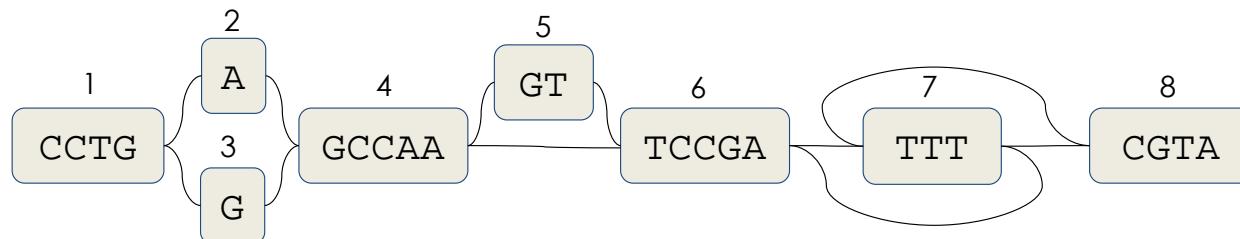
Nœuds (S = Segment)
fragments de séquence



graphe.gfa				
H	VN:Z:1.0	CCTG		
S	1	A		
S	2	G		
S	3	GCAA		
S	4	GT		
S	5	TCCGA		
S	6	TTT		
S	7	CGTA		
L	1	+	2	+
L	1	+	3	+
L	2	+	4	+
L	3	+	4	+
L	4	+	5	+
L	4	+	6	+
L	5	+	6	+
L	6	+	7	+
L	6	+	7	-
L	7	+	8	+
L	7	-	8	+
P	(A)	1+,2+,4+,6+,7-,8+		
P	(B)	1+,2+,4+,5+,6+,7+,8+		
P	(C)	1+,3+,4+,5+,6+,7-,8+		
P	(D)	1+,3+,4+,6+,7+,8+		

PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCA--TCCGA**TTT**CGTA



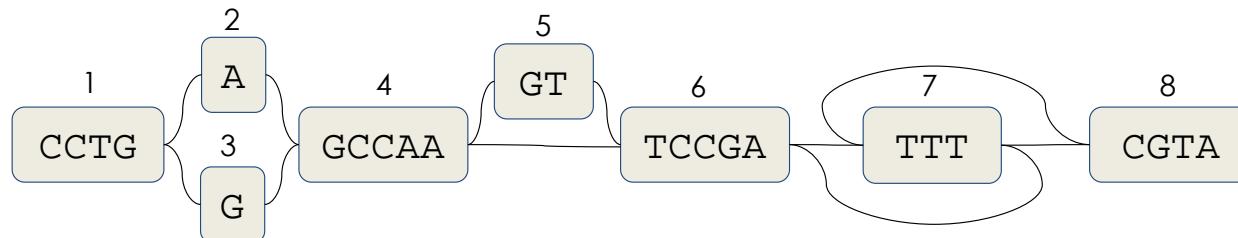
Nœuds (S = Segment)
fragments de séquence

Arêtes (L = Link)
liens entre deux nœuds

graphe.gfa				
H	VN:Z:1.0	CCTG	A	G
S	S	GCAA	GT	TCCGA
S	S	GT	TCCGA	CGTA
L	1	+	2	+
L	1	+	3	+
L	2	+	4	+
L	3	+	4	+
L	4	+	5	+
L	4	+	6	+
L	5	+	6	+
L	6	+	7	+
L	6	+	7	-
L	7	+	8	+
L	7	-	8	+
P	(A)	1+,2+,4+,6+,7-,8+		
P	(B)	1+,2+,4+,5+,6+,7+,8+		
P	(C)	1+,3+,4+,5+,6+,7-,8+		
P	(D)	1+,3+,4+,6+,7+,8+		

PANGÉNOME EN GRAPHE

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
- (B) CCTG**A**GCCAA**AGT**TCCGA**TTT**CGTA
- (C) CCTGGGCCA**AGT**TCCGA**AAA**CGTA
- (D) CCTGGGCCAA--TCCGA**TTT**CGTA



Nœuds (S = Segment)
fragments de séquence

Arêtes (L = Link)
liens entre deux nœuds

graphe.gfa

H	VN:Z:1.0	CCTG
S	1	A
S	2	G
S	3	GCCAA
S	4	GT
S	5	TCCGA
S	6	TTT
S	7	CGTA
L	1	+
L	1	+
L	2	+
L	3	+
L	4	+
L	4	+
L	5	+
L	6	+
L	6	+
L	7	+
L	7	-
L	7	+
L	7	+
P	(A)	1+, 2+, 4+, 6+, 7-, 8+
P	(B)	1+, 2+, 4+, 5+, 6+, 7+, 8+
P	(C)	1+, 3+, 4+, 5+, 6+, 7-, 8+
P	(D)	1+, 3+, 4+, 6+, 7+, 8+

Chemins (P = Path)
enchaînements de nœuds qui
reconstituent chaque génome



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)



Variation Graph Toolkit



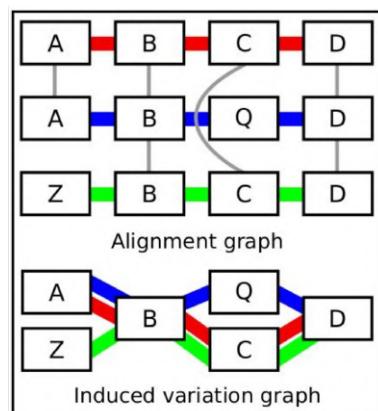
APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)



Pangenome Graph Builder (PGGB)

- Alignement tous génomes
- Construction du graphe



(Guarracino, 2021)

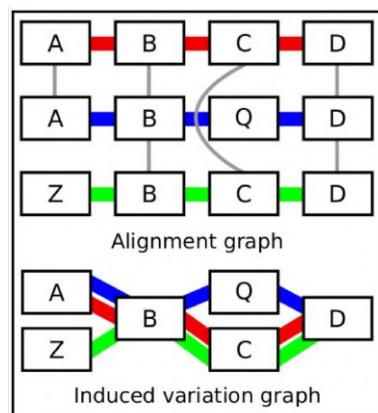
APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)



Pangenome Graph Builder (PGGB)

- Alignement tous génomes
- Construction du graphe

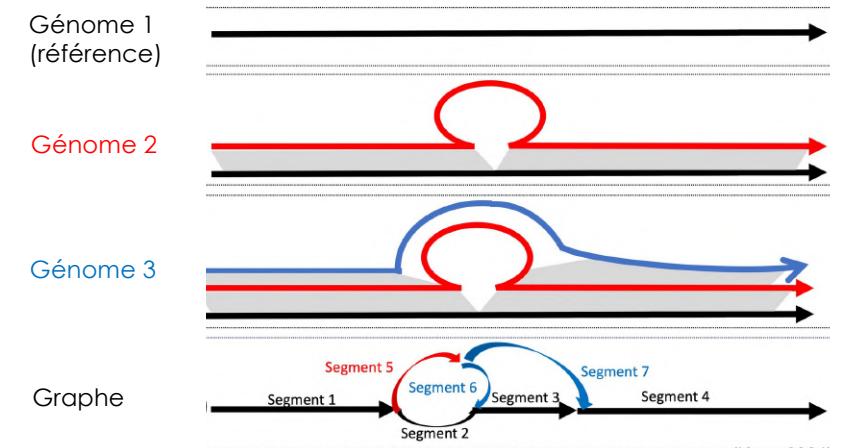


(Guarracino, 2021)



MiniGraph-Cactus

- Itérations d'alignement et modification du graphe
- Sur un premier génome de référence



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

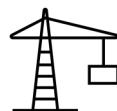
- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)
 - Génomes découpés en k-mers uniques de taille k avec recouvrement de k-1



Pantools, SplitMEM, Bifrost, TwoPaCo, Cuttlefish, ...



OBJECTIFS DU STAGE



Construction des premiers **pangénomes en graphe du VHD** avec les outils

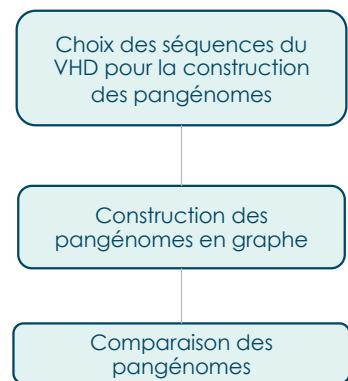
- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish



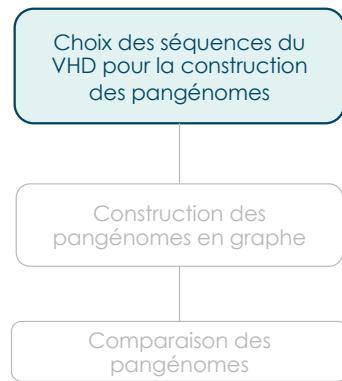
Comparaison des pangénomes produits



WORKFLOW



SEQUENCES VHD

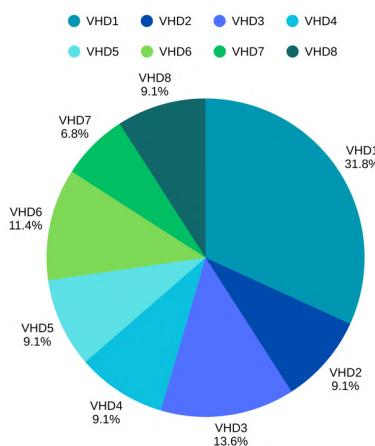
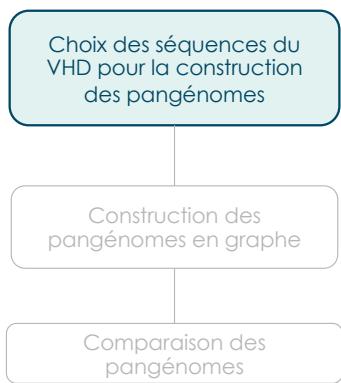


Séquences de génomes VHD dans GenBank

- Référence au laboratoire d'accueil (génotype VHD1)
- Références ICTV par génotype (1 par génotype)
- 3 articles de la littérature récente
 - Miao 2019   
 - Charre 2023  
 - Chowdhury 2025  
- 44 séquences au total
 - 22 pour la construction des graphes (Référence)
 - 22 pour validation



SEQUENCES VHD



Séquences de génomes VHD dans GenBank

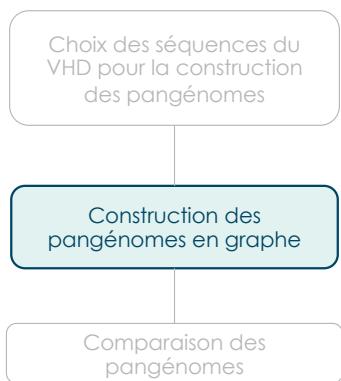
- Référence au laboratoire d'accueil (génotype VHD1)
- Références ICTV par génotype (1 par génotype)
- 3 articles de la littérature récente
 - Miao 2019
 - Charre 2023
 - Chowdhury 2025
- 44 séquences au total
 - 22 pour la construction des graphes (Référence)
 - 22 pour validation



Génotype	VHD1	VHD2	VHD3	VHD4	VHD5	VHD6	VHD7	VHD8	Total
Référence	9	2	3	2	1	2	2	1	22
Validation	5	2	3	2	3	3	1	3	22
Total	14	4	6	4	4	5	3	4	44



OUTILS



Linux Ubuntu 22.04.3

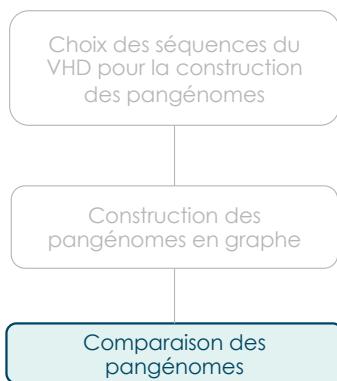
PGGB (Pangenome Graph Builder)
Minigraph-Cactus
Cuttlefish



.GFA

graph.gfa					
H	VN:Z:1.0				
S	1	CCTG			
S	2	A			
S	3	G			
S	4	GCCAA			
S	5	GT			
S	6	TCCGA			
S	7	TTT			
S	8	CGTA			
L	1	+	2	+	
L	1	+	3	+	
L	2	+	4	+	
L	3	+	4	+	
L	4	+	5	+	
L	4	+	6	+	
L	5	+	6	+	
L	6	+	7	+	
L	6	+	7	-	
L	7	+	8	+	
L	7	-	8	+	
P	(A)	1+,2+,4+,6+,7-,8+			
P	(B)	1+,2+,4+,5+,6+,7+,8+			
P	(C)	1+,3+,4+,5+,6+,7-,8+			
P	(D)	1+,3+,4+,6+,7+,8+			

OUTILS



Linux Ubuntu 22.04.3

PGGB (Pangenome Graph Builder)
Minigraph-Cactus
Cuttlefish

Bandage



gfatools

PGGE (Pangenome Graph Evaluator)



.GFA



.CSV



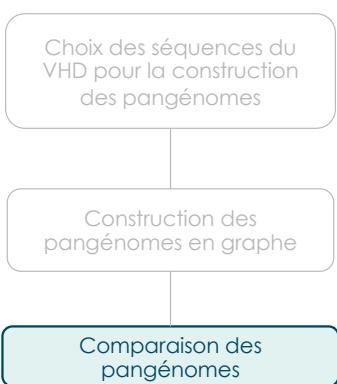
Introduction

Matériel et méthodes

Résultats

Conclusion

OUTILS



Linux Ubuntu 22.04.3

PGGB (Pangenome Graph Builder)

Minigraph-Cactus

Cuttlefish

Bandage

gfatools

PGGE (Pangenome Graph Evaluator)

Images Docker dans containers Singularity 3.11

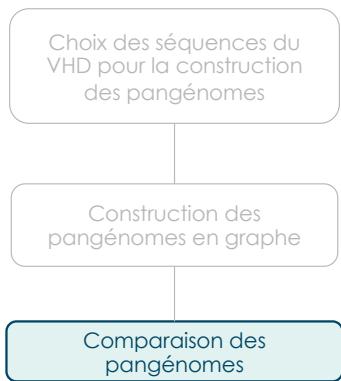


.GFA

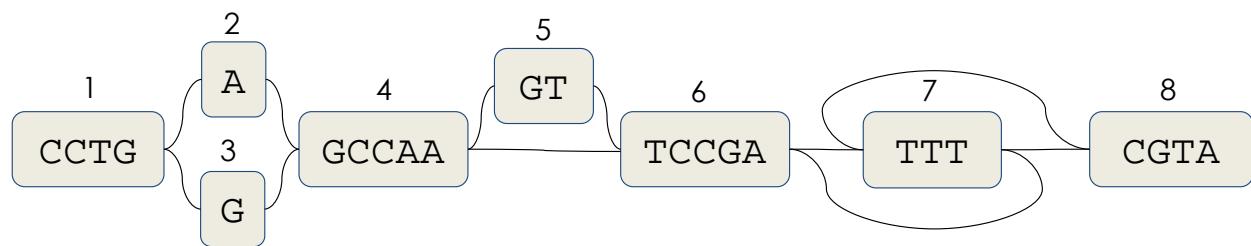


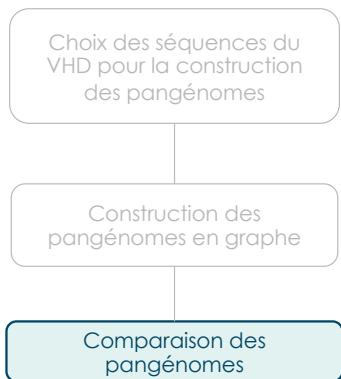
.CSV



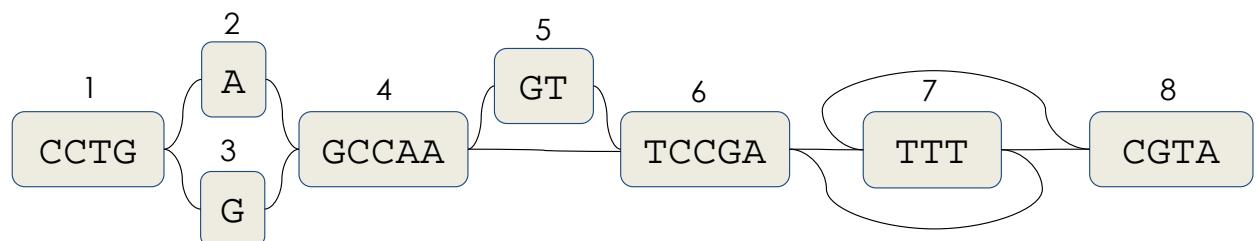


VISUALISATION AVEC BANDAGE

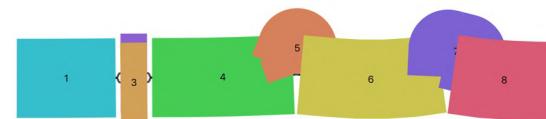


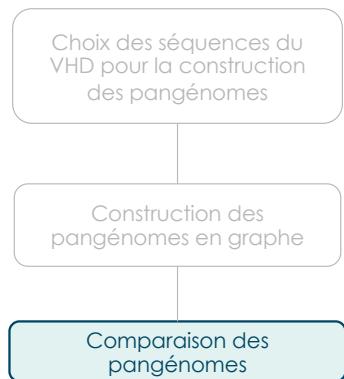


VISUALISATION AVEC BANDAGE

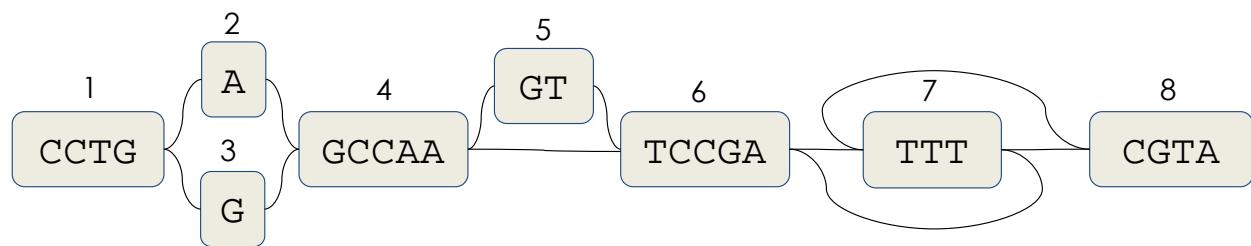


Bandage - /Users/lauraplayapariente/Downloads/graphe.gfa
Bandage - /Users/lauraplayapariente/Downloads/graphe.gfa



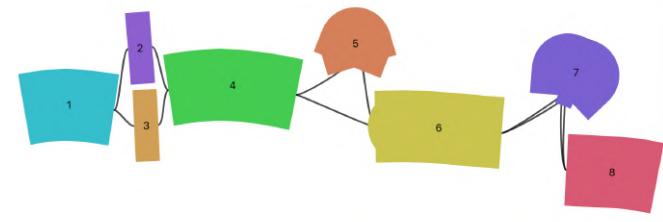
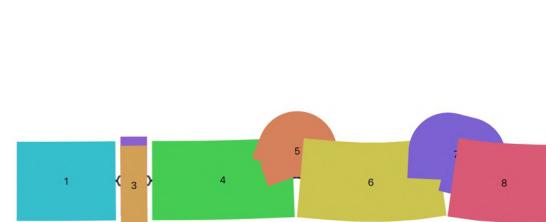


VISUALISATION AVEC BANDAGE



Bandage - /Users/lauraplayapariente/Downloads/graphe.gfa
Bandage - /Users/lauraplayapariente/Downloads/graphe.gfa

Bandage - /Users/lauraplayapariente/Downloads/graphe.gfa
Bandage - /Users/lauraplayapariente/Downloads/graphe.gfa



Introduction

Matériel et méthodes

Résultats

Conclusion

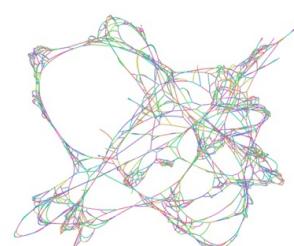
Choix des séquences du VHD pour la construction des pangénomes

Construction des pangénomes en graphe

Comparaison des pangénomes



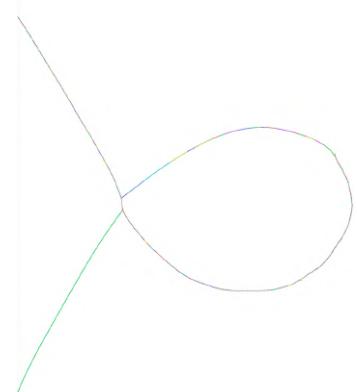
Cuttlefish



MiniGraph - Cactus



Pangenome Graph Builder





CRITERES

- Qualité des alignements

- % d'identité : aln.id
- % couverture : qsc



Cuttlefish



MiniGraph - Cactus



Pangenome Graph Builder

- Sur 44 séquences du VHD

- Référence
- [Validation](#)

Introduction

Matériel et méthodes

Résultats

Conclusion

CRITERES

- Qualité des alignements

- % d'identité : aln.id
- % couverture : qsc

- Sur 44 séquences du VHD

- Référence
- Validation



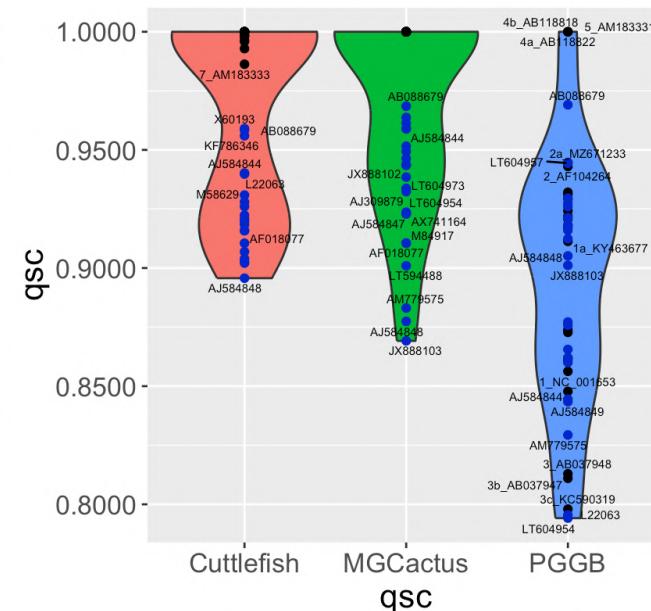
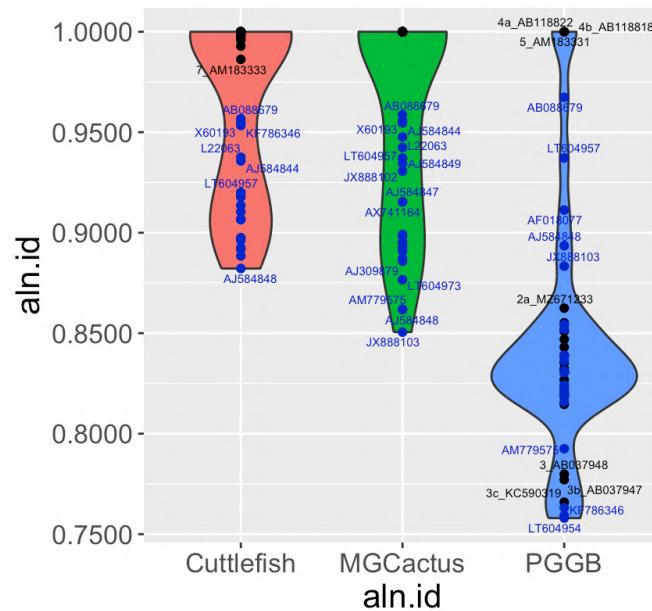
Cuttlefish



MiniGraph - Cactus



Pangenome Graph Builder



Introduction

Matériel et méthodes

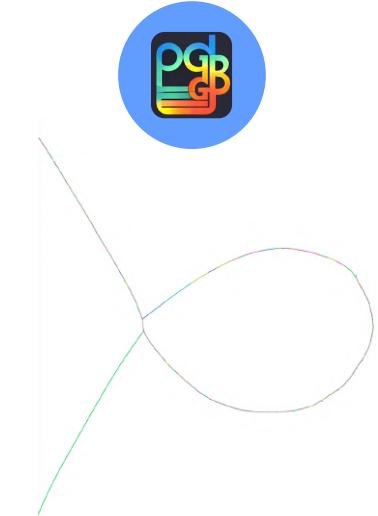
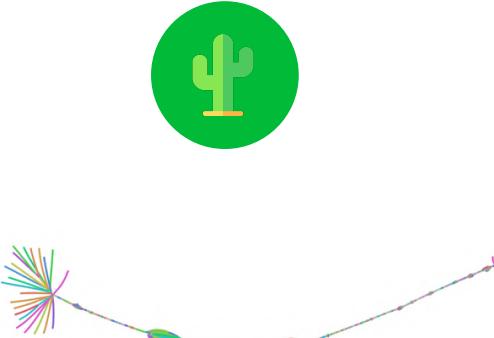
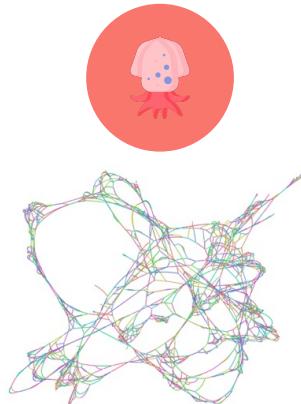
Résultats

Conclusion

CRITERES

- Topologie

- Nombre de nœuds



Introduction

Matériel et méthodes

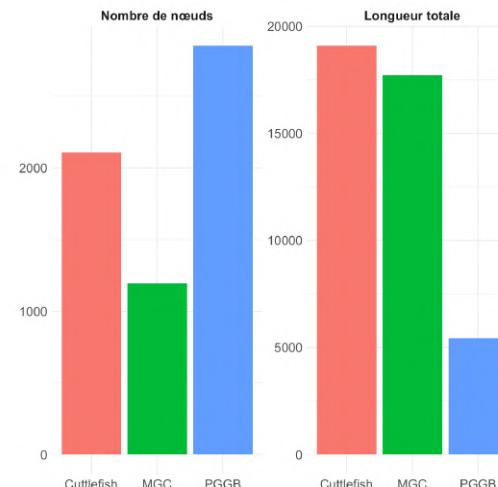
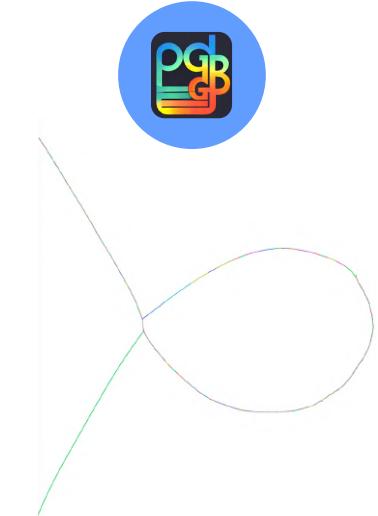
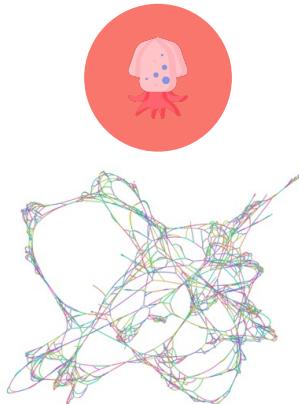
Résultats

Conclusion

CRITERES

- Topologie

- Nombre de nœuds
- Longueur totale



Introduction

Matériel et méthodes

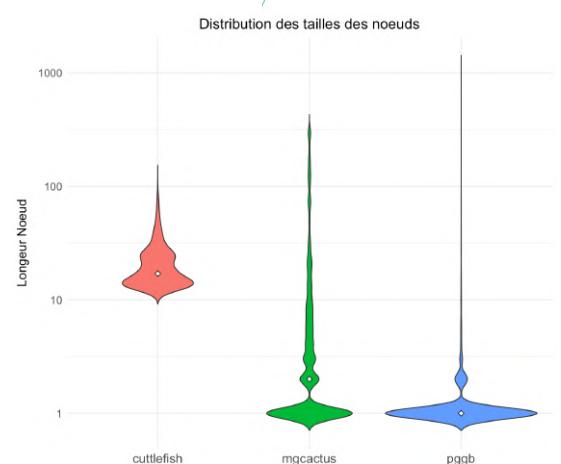
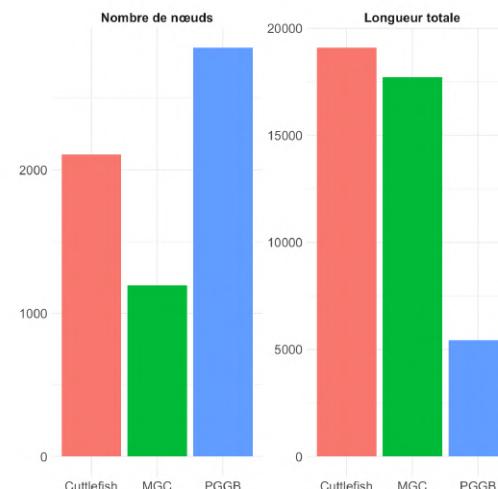
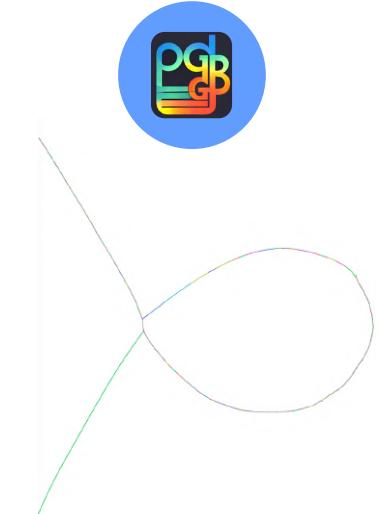
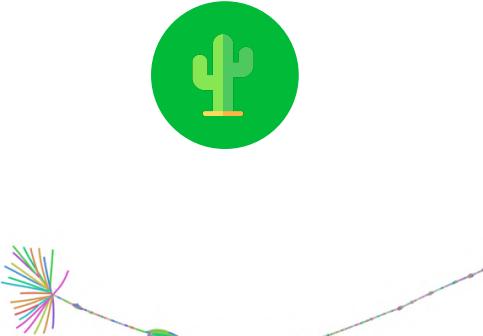
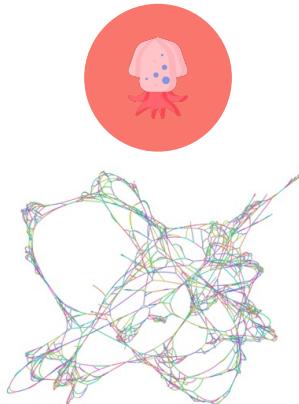
Résultats

Conclusion

CRITERES

• Topologie

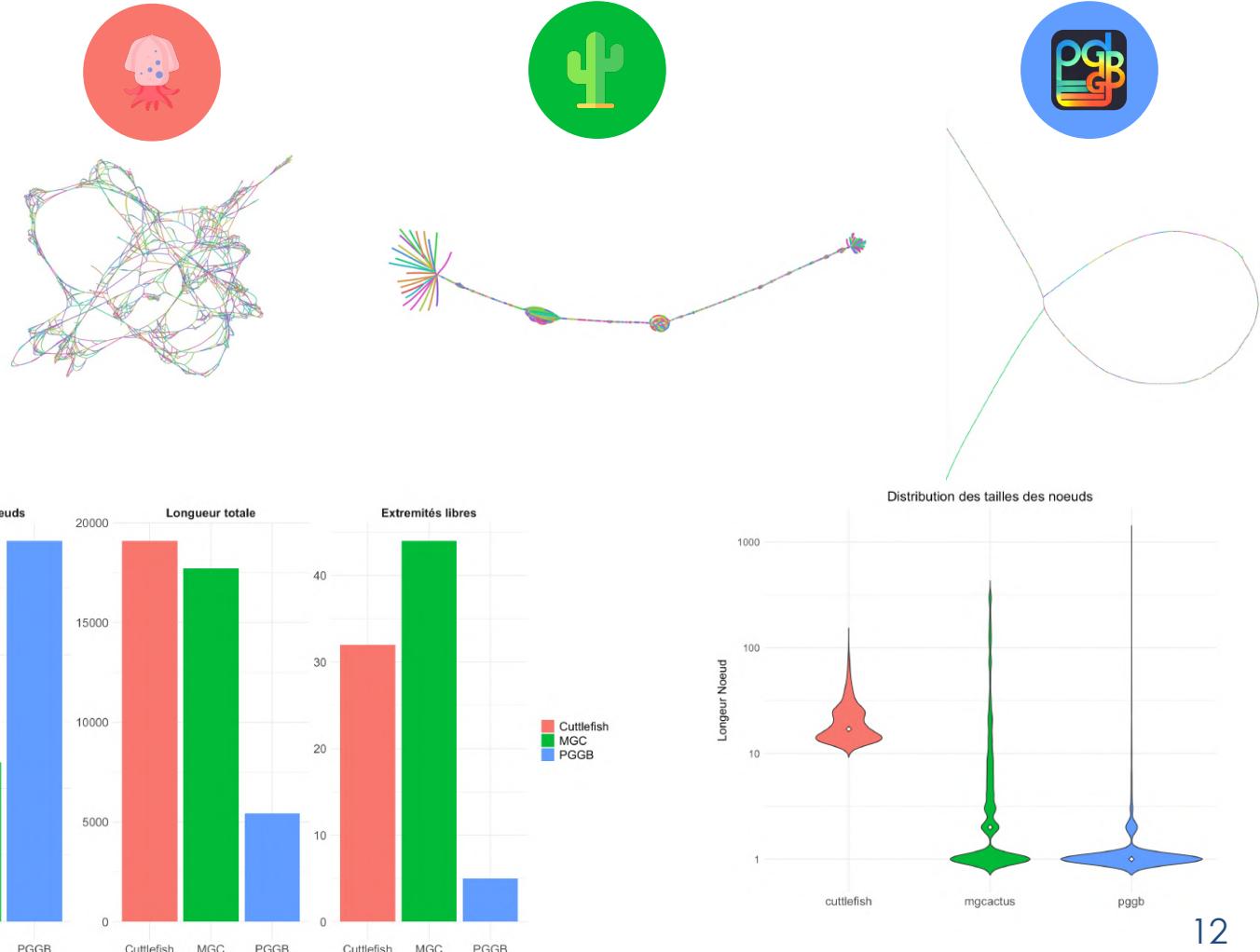
- Nombre de nœuds
- Longueur totale
- Distribution des tailles des nœuds



CRITERES

• Topologie

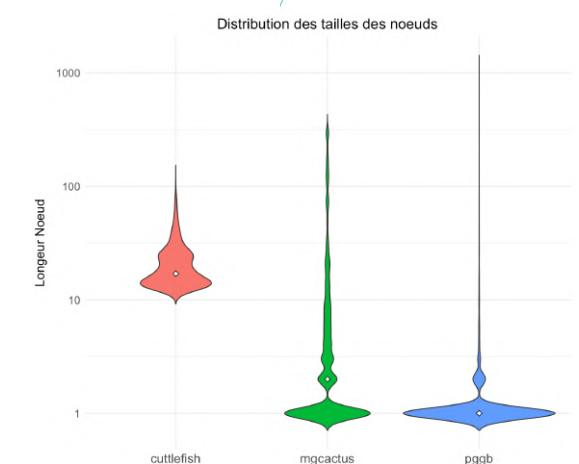
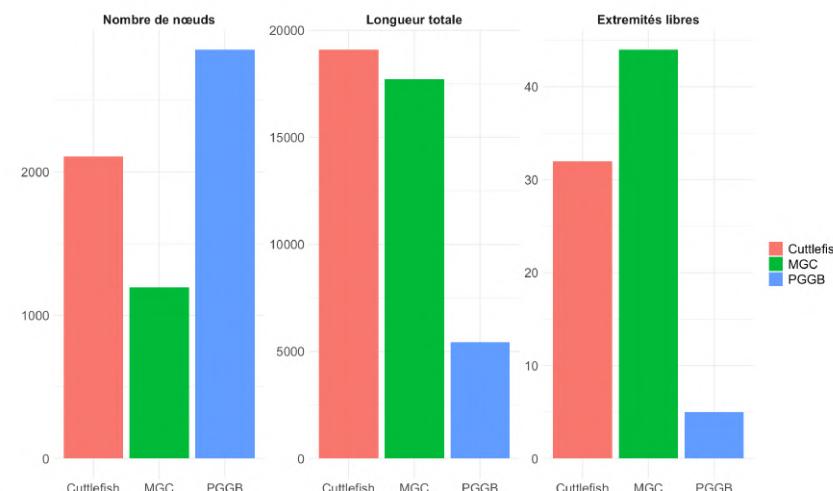
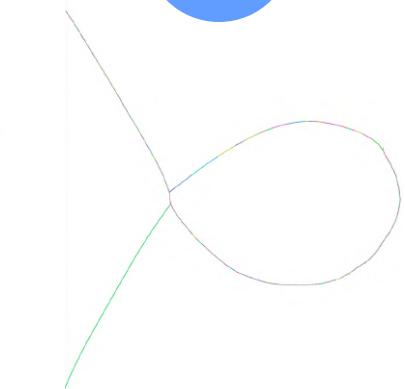
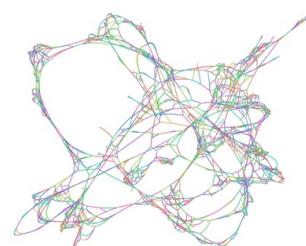
- Nombre de nœuds
- Longueur totale
- Distribution des tailles des nœuds
- Nombre d'extrémités libres



CRITERES

• Topologie

- Nombre de nœuds
- Longueur totale
- Distribution des tailles des nœuds
- Nombre d'extrémités libres
- Linéarité



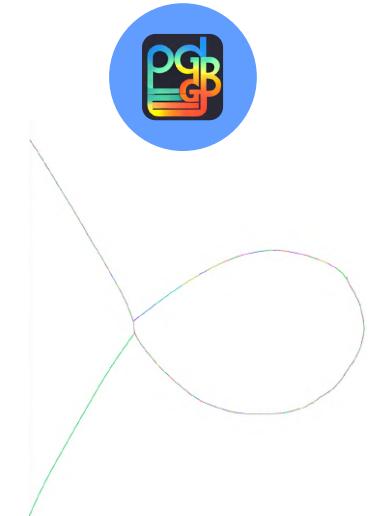
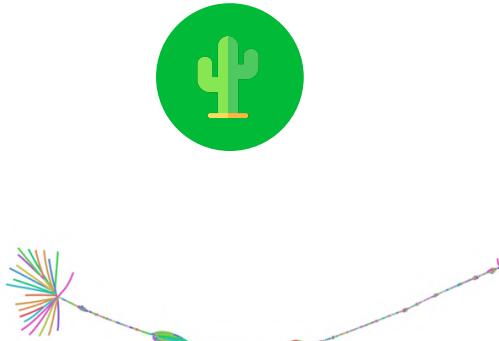
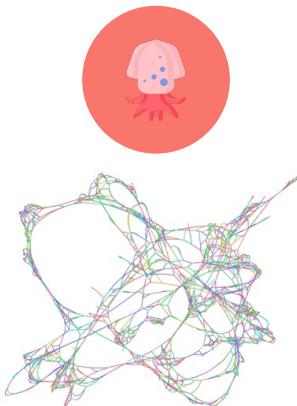
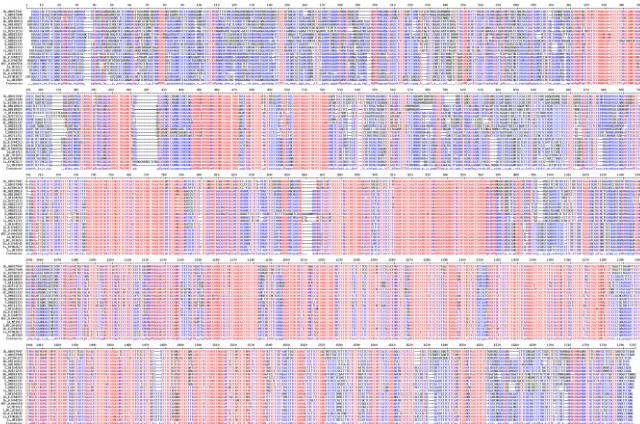
Introduction

Matériel et méthodes

Résultats

Conclusion

ORIGINE DES SEQUENCES



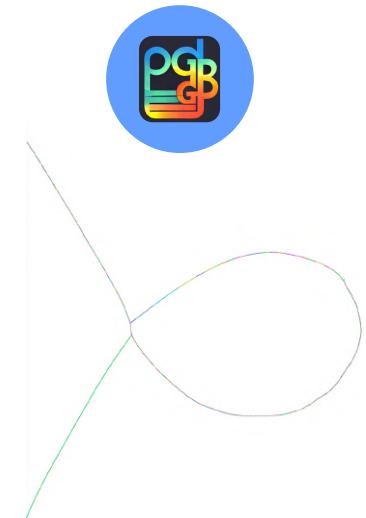
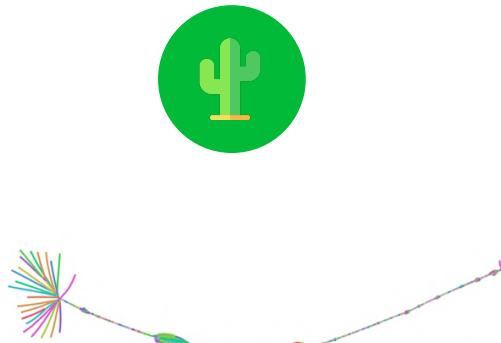
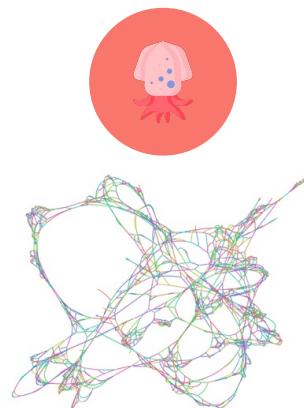
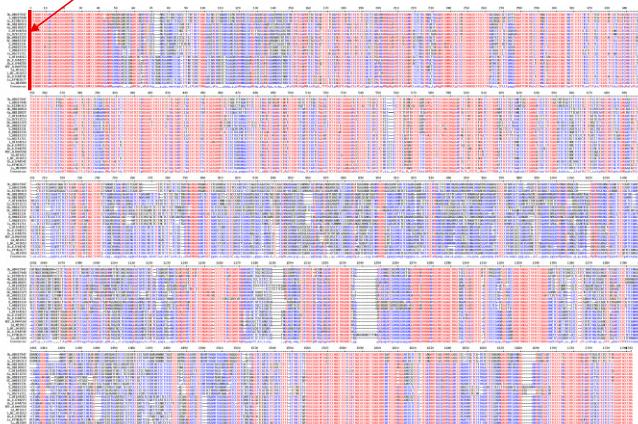
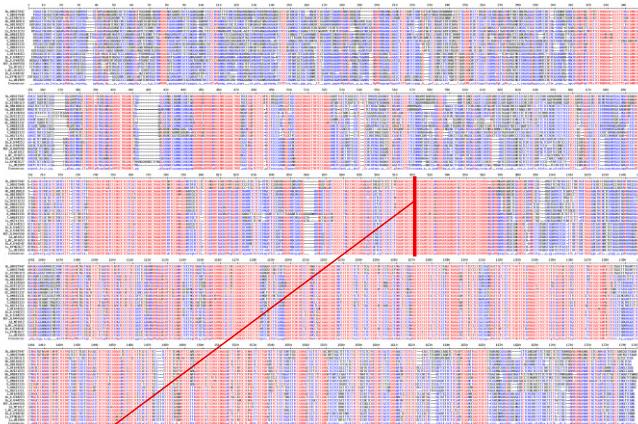
Introduction

Matériel et méthodes

Résultats

Conclusion

ORIGINE DES SEQUENCES



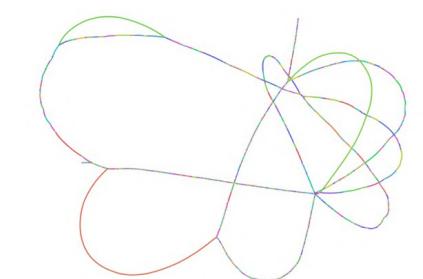
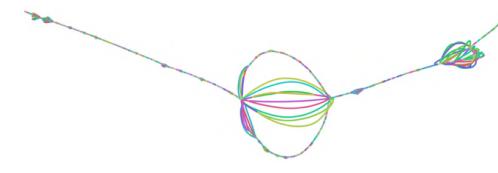
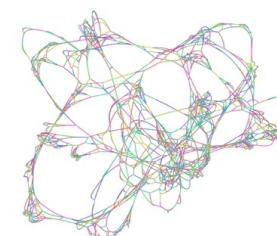
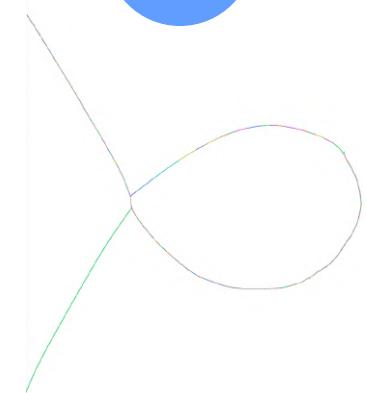
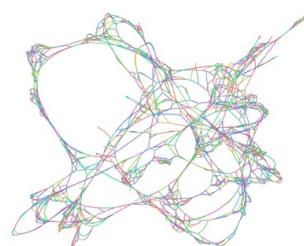
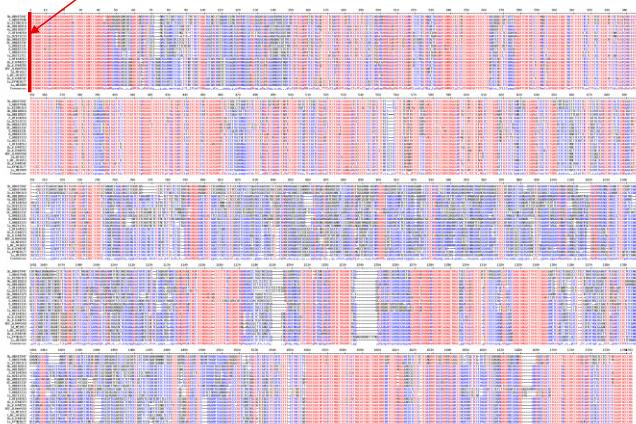
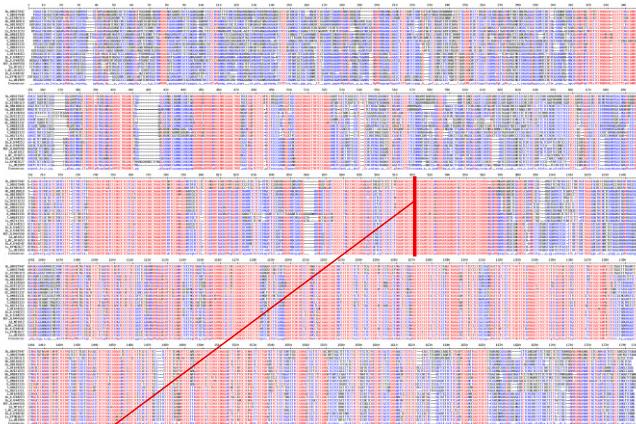
Introduction

Matériel et méthodes

Résultats

Conclusion

ORIGINE DES SEQUENCES







OBJECTIFS

CONCLUSION

PERSPECTIVES

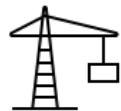
Introduction

Matériel et méthodes

Résultats

Conclusion

OBJECTIFS



Construction des premiers **pangénomes en graphe**
du VHD avec

- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish

CONCLUSION

PERSPECTIVES

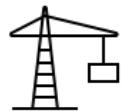
Introduction

Matériel et méthodes

Résultats

Conclusion

OBJECTIFS



- Construction des premiers **pangénomes en graphe du VHD** avec
- PGGB (Pangenome Graph Builder)
 - MiniGraph-Cactus
 - Cuttlefish

CONCLUSION

Pangénomes produits avec adaptation des paramètres d'alignement à la petite taille du VHD

PERSPECTIVES

OBJECTIFS



Construction des premiers **pangénomes en graphe du VHD** avec

- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish

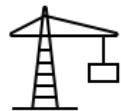
CONCLUSION

Pangénomes produits avec adaptation des paramètres d'alignement à la petite taille du VHD

PERSPECTIVES

Prise en compte de la circularité du génome

OBJECTIFS



Construction des premiers **pangenomes en graphe du VHD** avec

- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish



Comparaison des pangenomes produits



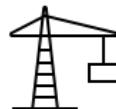
CONCLUSION

Pangenomes produits avec adaptation des paramètres d'alignement à la petite taille du VHD

PERSPECTIVES

Prise en compte de la circularité du génome

OBJECTIFS



Construction des premiers **pangenomes en graphe du VHD** avec

- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish



Comparaison des pangénomes produits



CONCLUSION

Pangenomes produits avec adaptation des paramètres d'alignement à la petite taille du VHD

Proposition de critères

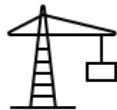
Qualité d'alignement

Topologie - Linéarité

PERSPECTIVES

Prise en compte de la circularité du génome

OBJECTIFS



Construction des premiers **pangénomes en graphe du VHD** avec

- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish



Comparaison des pangénomes produits



CONCLUSION

Pangénomes produits avec adaptation des paramètres d'alignement à la petite taille du VHD

Proposition de critères

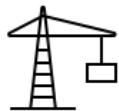
Qualité d'alignement

Topologie - Linéarité

PERSPECTIVES

Prise en compte de la circularité du génome

OBJECTIFS



Construction des premiers **pangénomes en graphe du VHD** avec

- PGGB (Pangenome Graph Builder)
- MiniGraph-Cactus
- Cuttlefish



Comparaison des pangénomes produits



CONCLUSION

Pangénomes produits avec adaptation des paramètres d'alignement à la petite taille du VHD

Proposition de critères

Qualité d'alignement

Topologie - Linéarité

PERSPECTIVES

Prise en compte de la circularité du génome

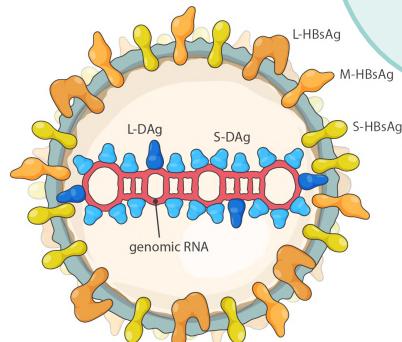
Enrichissement des critères

Test différents algorithmes d'alignement

Objectivation critère linéarité



Merci



REFERENCES

Charre, C., Regue, H., Dény, P., Josset, L., Chemin, I., Zoulim, F., and Scholtes, C. (2023). Improved hepatitis delta virus genome characterization by single molecule full-length genome sequencing combined with VIRIONT pipeline. *Journal of Medical Virology* 95, e28634. <https://doi.org/10.1002/jmv.28634>.

Chowdhury, S., Jacobsen, C., Depledge, D.P., Wedemeyer, H., Sandmann, L., and Kefalakes, H. (2025). Sequence analysis of the hepatitis D virus across genotypes reveals highly conserved regions amidst evidence of recombination. *Virus Evolution* 11, veaf012. <https://doi.org/10.1093/ve/veaf012>.

Guarracino A. poster 2021 :
https://andreaguarracino.github.io/posters/BoG2021_ThePanGenomeGraphBuilder_Poster_AndreaGuarracino.pdf

Hunt M, Letcher B, Malone KM, et al. Minos: variant adjudication and joint genotyping of cohorts of bacterial genomes. *Genome Biol.* 2022;23(1):147. Published 2022 Jul 5. doi:10.1186/s13059-022-02714-x

ICTV website : <https://ictv.global/report/chapter/kolmioviridae/kolmioviridae/deltavirus>

Karimzadeh H, Usman Z, Frishman D, Roggendorf M. Genetic diversity of hepatitis D virus genotype-1 in Europe allows classification into subtypes. *J Viral Hepat.* 2019;26(7):900-910. doi:10.1111/jvh.13086

Kaur H, Shannon LM, Samac DA. A stepwise guide for pangenome development in crop plants: an alfalfa (*Medicago sativa*) case study. *BMC Genomics.* 2024;25(1):1022. Published 2024 Oct 31. doi:10.1186/s12864-024-10931-w

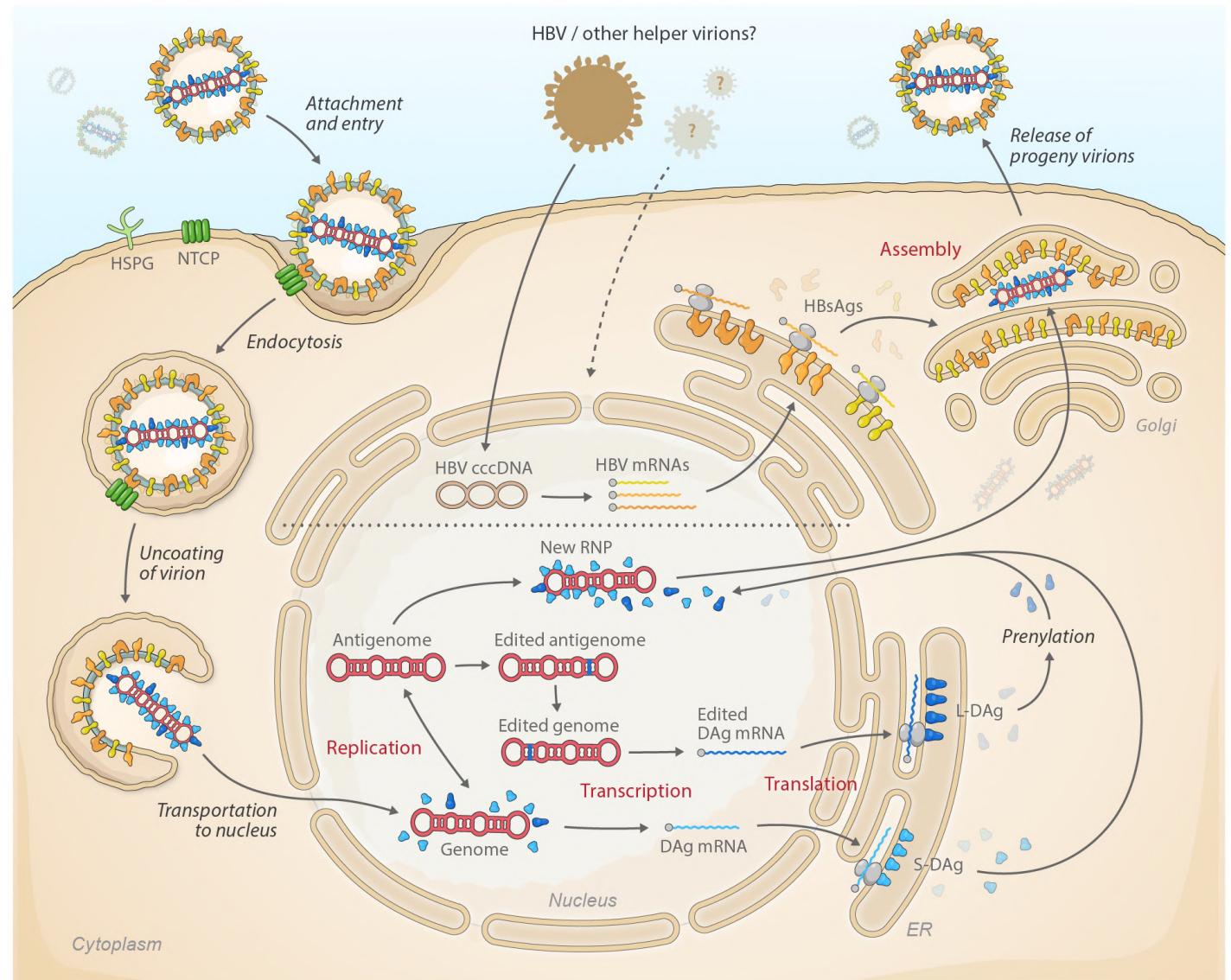
Lim, Y.-S.; Kim, W.R.; Dieterich, D.; Kao, J.-H.; Flaherty, J.F.; Yee, L.J.; Roberts, L.R.; Razavi, H.; Kennedy, P.T.F. Evidence for Benefits of Early Treatment Initiation for Chronic Hepatitis B. *Viruses* **2023**, *15*, 997. <https://doi.org/10.3390/v15040997>

Miao, Z., Zhang, S., Ma, Z., Hakim, M.S., Wang, W., Peppelenbosch, M.P., and Pan, Q. (2019). Recombinant identification, molecular classification and proposed reference genomes for hepatitis delta virus. *Journal of Viral Hepatitis* 26, 183–190. <https://doi.org/10.1111/jvh.13010>.

OMS website : <https://www.who.int/fr/news-room/fact-sheets/detail/hepatitis-d>

LE VIRUS DE L'HEPATITE D

cycle de vie



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

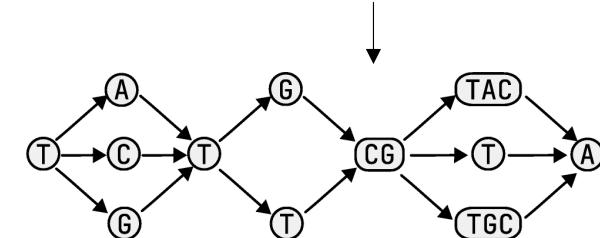
- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)

Référence

TATGCGTACA

Fichier VCF

Pos	Ref	Alt
2	A	C,G
4	G	T
7	TAC	T,TGC



(Hunt, 2022)



Variation Graph Toolkit



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)
 - Génomes découpés en k-mers uniques de taille k avec recouvrement de k-1

CCTGAGCCAAGTTCC



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

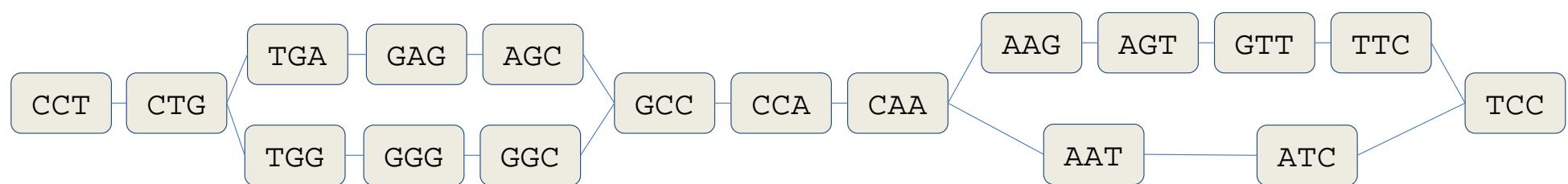
- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)
 - Génomes découpés en k-mers uniques de taille k avec recouvrement de k-1

CCTG**A**GCCAAG**T**TCC

CCT**G**GCCAAG**T**TCC

CCT**A**GCAA--TCC

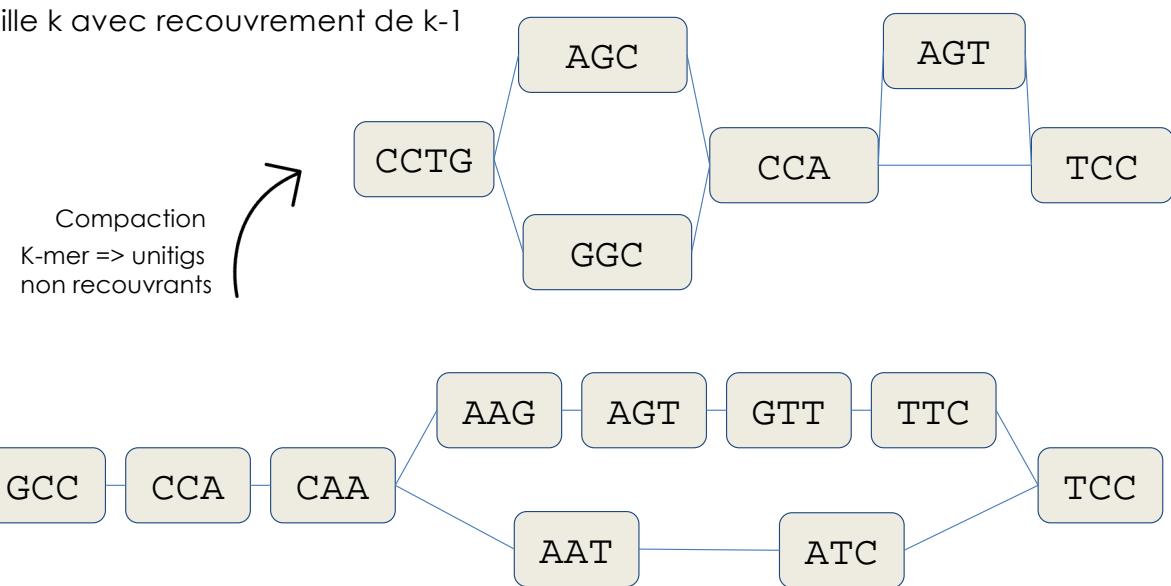
CCT**GG**CCA--TCC



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)
 - Génomes découpés en k-mers uniques de taille k avec recouvrement de k-1

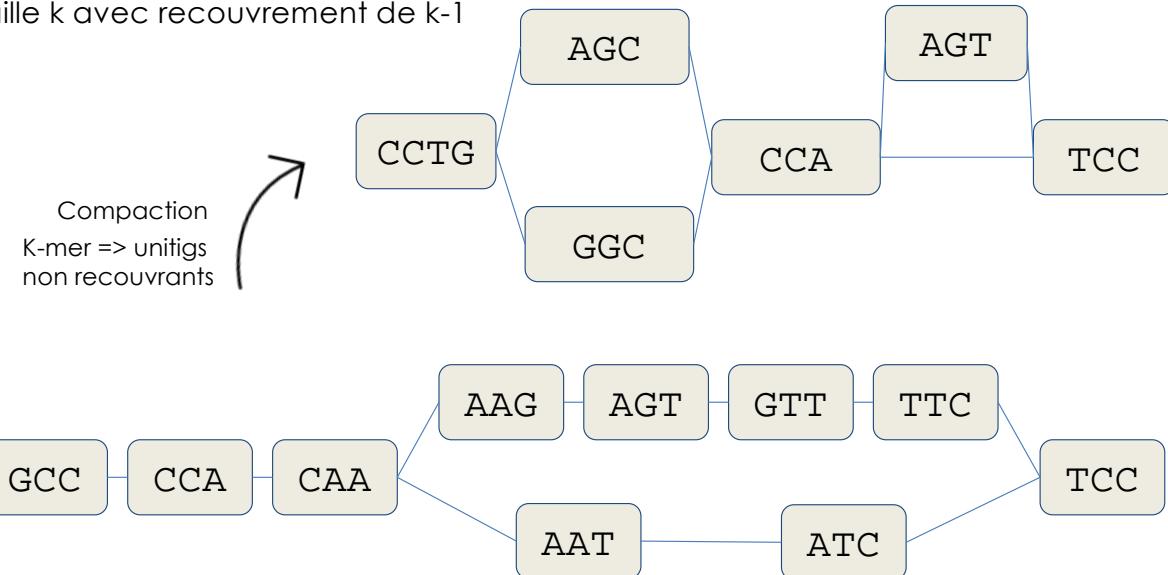
CCTG**A**GCCAAG**T**TCC
 CCTG**G**GCCAAG**T**TCC
 CCTG**A**GCAA--TCC
 CCTG**GG**GCAA--TCC



APPROCHES POUR LA CONSTRUCTION DE PANGÉNOMES EN GRAPHE

- A partir d'un génome de référence et une liste de variations (VCF)
- A partir d'alignements des génomes entiers
- Avec des graphes de De Bruijn compactés (cDBG)
 - Génomes découpés en k-mers uniques de taille k avec recouvrement de k-1

CCTG**A**GCCAAG**T**TCC
 CCTG**GG**GCCAAG**T**TCC
 CCTG**A**GCCA--TCC
 CCTG**GG**GCAA--TCC



Pantools, SplitMEM, Bifrost, TwoPaCo, Cuttlefish, ...

[nature](#) > [nature methods](#) > [brief communications](#) > [article](#)

Brief Communication | Published: 21 October 2024

Building pangenome graphs

Erik Garrison , Andrea Guerracino, Simon Heumos, Flavia Villani, Zhigui Bao, Lorenzo Tattini, Jörg Hagmann, Sebastian Vorbrugg, Santiago Marco-Sola, Christian Kubica, David G. Ashbrook, Kaisa Thorell, Rachel L. Rusholme-Pilcher, Gianni Liti, Emilio Rudbeck, Agnieszka A. Golicz, Sven Nahnsen, Zuyu Yang, Moses Njagi Mwaniki, Franklin L. Nobrega, Yi Wu, Hao Chen, Joep de Ligt, Peter H. Sudmant, ... Pjotr Prins + Show authors

[Nature Methods](#) 21, 2008–2012 (2024) | [Cite this article](#)

10k Accesses | 85 Citations | 30 Altmetric | [Metrics](#)

Abstract

Pangenome graphs can represent all variation between multiple reference genomes, but current approaches to build them exclude complex sequences or are based upon a single reference. In response, we developed the PanGenome Graph Builder, a pipeline for constructing pangenome graphs without bias or exclusion. The PanGenome Graph Builder uses all-to-all alignments to build a variation graph in which we can identify variation, measure conservation, detect recombination events and infer phylogenetic relationships.

Minigraph- Cactus

nature biotechnology

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

[nature](#) > [nature biotechnology](#) > [articles](#) > [article](#)

Article | Published: 10 May 2023

Pangenome graph construction from genome alignments with Minigraph-Cactus

[Glenn Hickey](#)✉, [Jean Monlong](#), [Jana Ebler](#), [Adam M. Novak](#), [Jordan M. Eizenga](#), [Yan Gao](#), [Human Pangenome Reference Consortium](#), [Tobias Marschall](#), [Heng Li](#) & [Benedict Paten](#)✉

[Nature Biotechnology](#) **42**, 663–673 (2024) | [Cite this article](#)

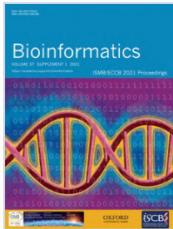
33k Accesses | **207** Citations | **1281** Altmetric | [Metrics](#)

Abstract

Pangenome references address biases of reference genomes by storing a representative set of diverse haplotypes and their alignment, usually as a graph. Alternate alleles determined by variant callers can be used to construct pangenome graphs, but advances in long-read sequencing are leading to widely available, high-quality phased assemblies. Constructing a pangenome graph directly from assemblies, as opposed to variant calls, leverages the graph's ability to represent variation at different scales. Here we present the Minigraph-Cactus pangenome pipeline, which creates pangenomes directly from whole-genome alignments, and demonstrate its ability to scale to 90 human haplotypes from the Human Pangenome Reference Consortium. The method builds graphs containing all forms of genetic variation while allowing use of current mapping and genotyping tools. We measure the effect of the

Bioinformatics

Issues Advance articles Submit ▾ Alerts About ▾ Bioinformatics ▾



Volume 37, Issue
Supplement_1
July 2021

JOURNAL ARTICLE

Cuttlefish: fast, parallel and low-memory compaction of de Bruijn graphs from large-scale genome collections

Jamshed Khan , Rob Patro 

Bioinformatics, Volume 37, Issue Supplement_1, July 2021, Pages i177–i186,
<https://doi.org/10.1093/bioinformatics/btab309>

Published: 12 July 2021

 PDF  Split View  Cite  Permissions  Share ▾

Article Contents

- Abstract
- 1 Introduction
- 2 Related work
- 3 Preliminaries
- 4 Algorithm
- 5 Results
- 6 Conclusion
- Acknowledgements

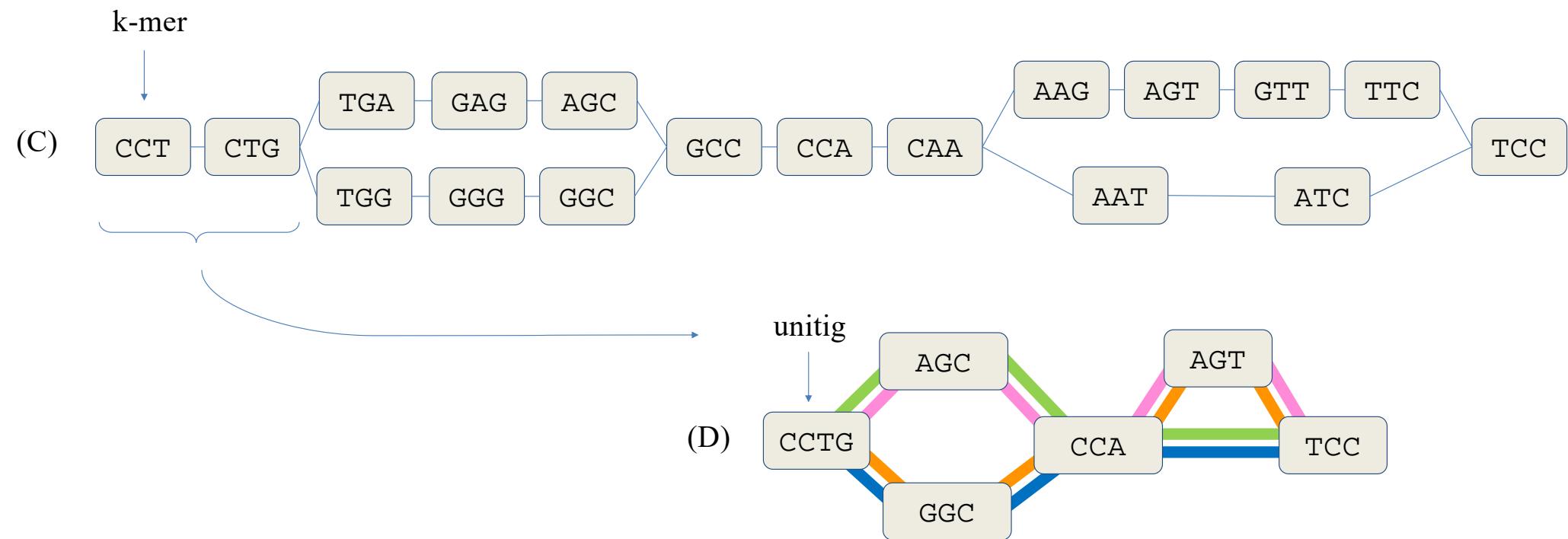
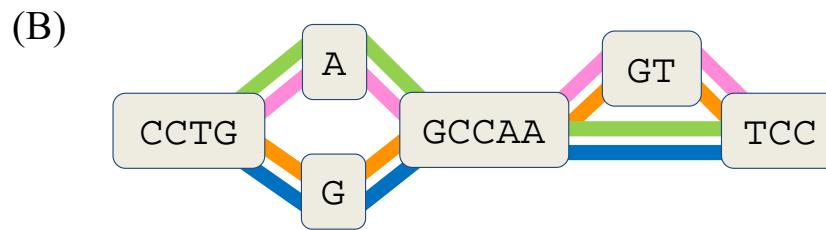
Abstract

Motivation

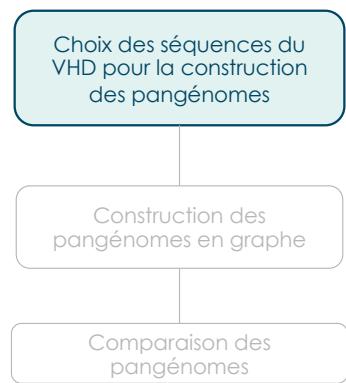
The construction of the compacted de Bruijn graph from collections of reference genomes is a task of increasing interest in genomic analyses. These graphs are increasingly used as sequence indices for short- and long-read alignment. Also, as we sequence and assemble a greater diversity of genomes, the colored compacted de Bruijn graph is being used more and more as the basis for efficient methods to perform comparative genomic analyses on these genomes. Therefore, time- and memory-efficient construction of the graph from reference sequences is an important problem.

(A)

```
CCTGAGCCAAGTTCC
CCTGGGCCAACAGTTCC
CCTGAGCCAAT--TCC
CCTGGGCCAAAT--TCC
```



SEQUENCES VHD



Génotype VHD	ICVT 2025 Genus Deltavirus	Miao 2019 ¹⁷ Sous-type numéros d'accession	Chowdhury 2025 ¹⁸ Table S5	Charre 2023 ¹⁹ Table S1
VHD-1 AJ000558 (*)	AF104263	1a	JX888100 KY463677	NC_001653
		1b	JX888098 KJ744242 KJ744255	KJ744223
		1c		M84917
		1d		M58629
		1e		M21012
VHD -2	AF104264	2a	X60193	MZ671233
		2b	AJ309879	X60193
VHD -3	AB037948	3A	LT604954	L22063
		3b	AB037947	LT604954
		3c	KC590319	KF786346 RC
VHD -4	AF018077	4a	AF018077	AB118822
		4b	AB118818	AB088679
VHD -5	AM183331	5a	JX888103	LT604957
		5b	AM183331	AJ584848
VHD -6	AX741164	6a	AJ584847	AJ584847
		6b	JX888102	AM183329
		6C	AM183332	
VHD -7	AM183333	7a	AJ584844	MG711711
		7B	AM183333	AJ584844
VHD -8	AM183330	8a	AJ584849	AM183333
		8B	LT594488	AJ584849

Introduction

Matériel et méthodes

Résultats

Conclusion

OUTILS



Linux Ubuntu 22.04.3

Choix des séquences du VHD pour la construction des pangénomes

Construction des pangénomes en graphe

Comparaison des pangénomes

Outils	Version	Paramètres
PGGB	v0.7.4	-p 77 -s 1000
MiniGraph Cactus	v2.9.9	--noSplit --permissiveContigFilter --configFile vhd.xlm
Cuttlefish	v2.2	-k 13

Bandage

v0.8.1

.GFA

gfatools

v0.5-r292

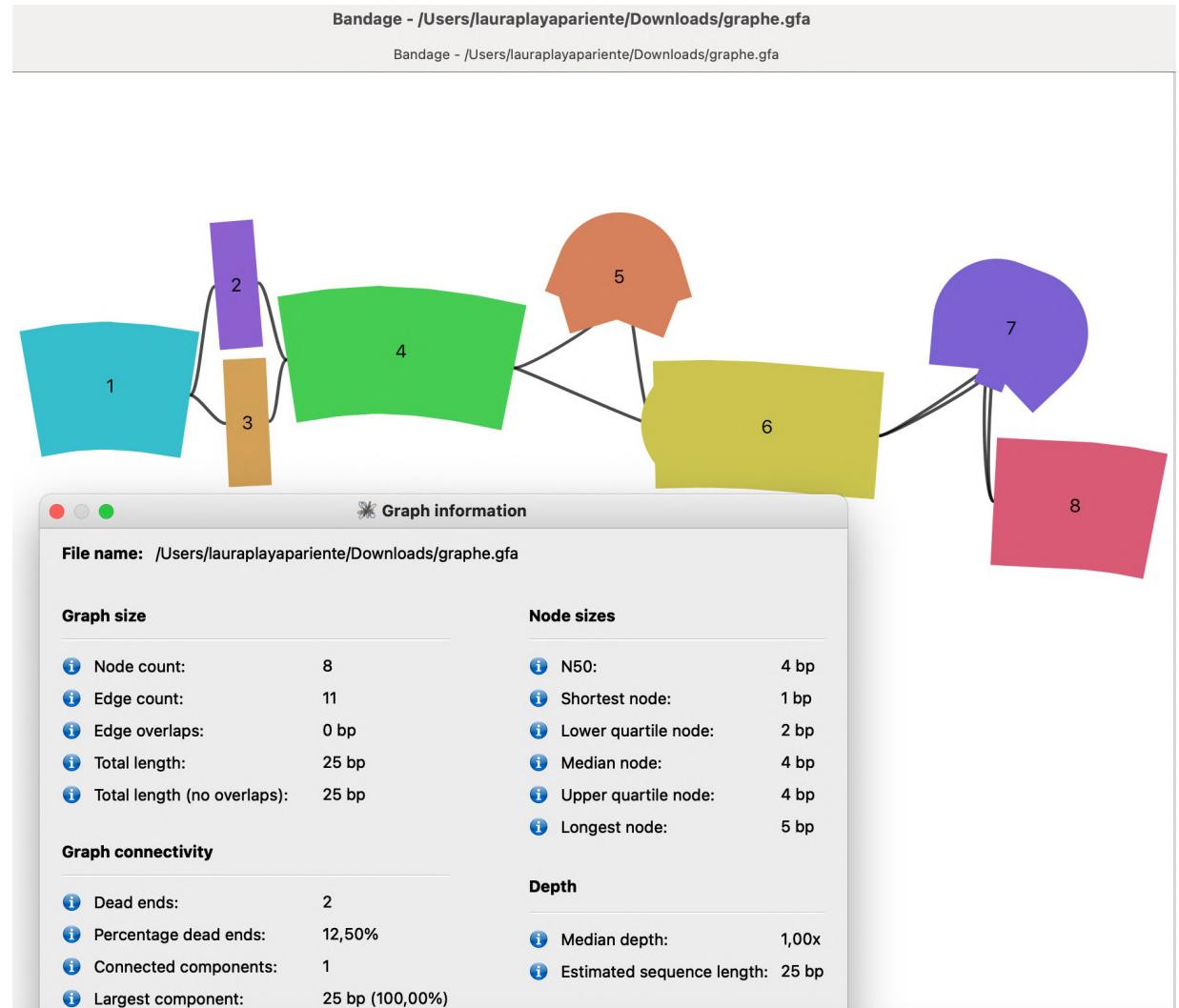
PGGE (Pangenome Graph Evaluator)

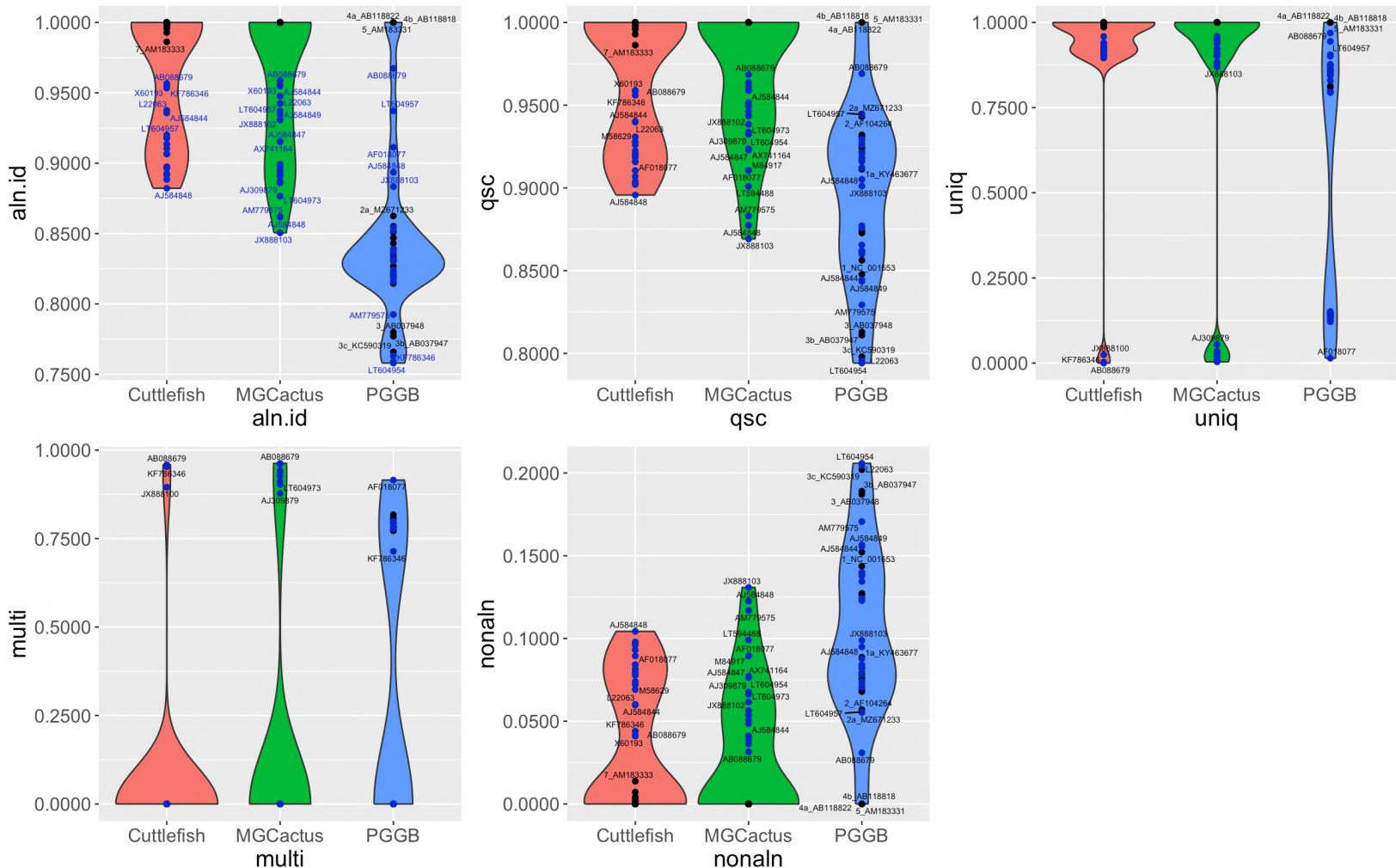
.CSV

- (A) CCTG**A**GCCAA--TCCGA**AAA**CGTA
 (B) CCTG**A**GCCAA**GT**TCCGA**TTT**CGTA
 (C) CCTGGGCCAA**GT**TCCGA**AAA**CGTA
 (D) CCTGGGCCAA--TCCGA**TTT**CGTA

MSA length : 24 b

- Nombre de noeuds : 8
- Longueur totale : 25
- Extrémités libres : 2





Introduction

Matériel et méthodes

Résultats

Conclusion

Master 1 Bioinformatics internship

Pangenome Graph for Hepatitis D Virus (HDV)



Mai – July
2025



INSTITUT
D'HÉPATOLOGIE
DE LYON



Xavier Grand
Research Engineer
Senior Bioinformaticien

- Co-infection with HBV
- Different clinical presentations
 - Acute -> mild, severe, fulminant
 - Chronic -> increased risk of developing cirrhosis and hepatocellular carcinoma

HDV genome

- ssRNA genome 1.7 kb
- Highly variable, possible recombination
 - 8 genotypes - 2 to 5 subtypes per genotype
- Large differences in replicative capacity in vitro

HDV
Genotype
subtype
variant



- Epidemiology
- Pathogenesis
- Treatment response