

Thesis, FRANZEN Rachelle

Auteur : Boanfa, Stefan-Laurentiu

Promoteur(s) : HAYETTE, Marie-Pierre

Faculté : Faculté de Médecine

Diplôme : Master en sciences biomédicales, à finalité spécialisée en biomédical data management

Année académique : 2025-2026

URI/URL : <http://hdl.handle.net/2268.2/25969>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

Décrypter la diversité génétique de l'*Echinococcus* par séquençage complet du mitogénome : comparaison des technologies Illumina et Nanopore à partir d'échantillons cliniques humains

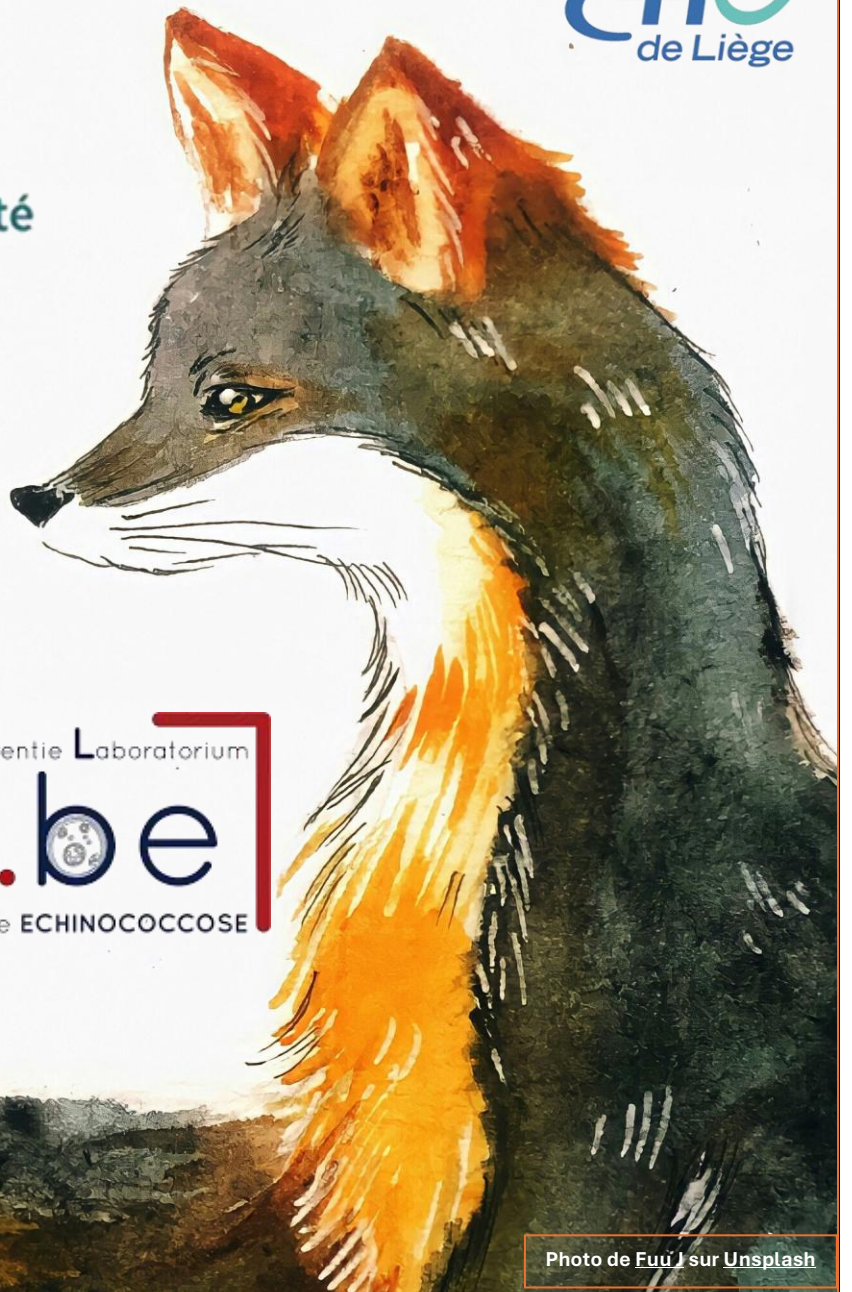


Photo de Fuu sur Unsplash



Année Académique : 2025-2026

Mémoire de Master en Sciences Biomédicales, à finalité spécialisée en
Biomedical Data Management

MÉMORANT : BOANFA STEFAN-LAURENTIU

SUPERVISEUR ACADÉMIQUE : PROFESSEURE HAYETTE MARIE-PIERRE

SUPERVISEURS TECHNIQUES : EGREK SABRINA, EL-MOUSSAOUI KHALID

REMERCIEMENTS

Je tiens tout d'abord à remercier la professeure Marie-Pierre Hayette pour son implication dans ce projet, sa confiance et l'opportunité qu'elle m'a offerte en m'accueillant au sein de son équipe. Ce mémoire n'aurait pas pu voir le jour sans son encadrement, son soutien et la dynamique de travail qu'elle a rendue possible. En permettant la mise en place de ce projet et en accordant à notre équipe les moyens nécessaires pour le mener à bien, elle m'a offert une expérience scientifique et humaine particulièrement importante. Cette étape a profondément marqué mon parcours, et je lui en suis très reconnaissant.

Je tiens à exprimer ma profonde reconnaissance envers Sabrina Egrek, scientifique au sein du laboratoire de microbiologie du CHU de Liège. Son accompagnement a été déterminant pour la réalisation de ce mémoire. Par son calme, sa bienveillance et sa grande implication, elle m'a permis d'évoluer dans un environnement de travail sain, stimulant et profondément formateur. La confiance qu'elle m'a accordée tout au long de ce projet a beaucoup compté pour moi. Elle m'a permis de progresser, d'oser davantage et de gagner en autonomie. Elle a su m'encadrer avec patience, tout en me laissant l'espace nécessaire pour apprendre par moi-même, y compris à travers mes erreurs. Cette manière de former, à la fois exigeante et humaine, a joué un rôle essentiel dans mon épanouissement au sein du laboratoire. Je garde également une reconnaissance particulière pour le temps considérable qu'elle a consacré à la correction de ce mémoire. Ses remarques, toujours constructives, ainsi que l'attention sincère portée à mon travail, m'ont permis d'améliorer ce manuscrit et de progresser dans ma manière de penser et de présenter un travail scientifique. Son implication restera pour moi l'un des éléments les plus marquants de cette expérience.

J'adresse également toute ma gratitude à Khalid El-Moussaoui pour son aide, son expertise et le partage généreux de ses connaissances. Sa maîtrise technique, sa curiosité scientifique et sa capacité à transmettre m'ont énormément apporté. Grâce à nos échanges, il m'a permis de mieux comprendre de nombreux aspects du travail de laboratoire, ainsi que d'élargir ma réflexion scientifique. Sa disponibilité, ses conseils et les nombreuses discussions m'ont permis de progresser bien au-delà du cadre strict de ce mémoire. Son accompagnement m'a poussé à dépasser mes limites et à gagner en rigueur. Sa passion pour la science et sa volonté de transmettre ont constitué une véritable source de motivation tout au long de ce projet.

TABLE DES MATIÈRES

CHAPITRE I : INTRODUCTION	p1
1.1) L'histoire évolutive d'un couple de parasites multicellulaires	p1
1.2) Taxonomie du genre <i>Echinococcus</i> et distribution mondiale	p2
1.3) Épidémiologie de l'EA et de l'EK en Europe et en Belgique	p3
1.4) Cycle parasitaire d' <i>E. multilocularis</i>	p3
1.5) Morphologie d' <i>E. multilocularis</i>	p5
1.6) L'EA : de la physiopathologie à la prise en charge du patient	p6
1.7) Cycle parasitaire d' <i>E. granulosus</i>	p7
1.8) Morphologie d' <i>E. granulosus</i>	p8
1.9) L'EK : de la physiopathologie à la prise en charge du patient	p9
1.10) Méthodes de diagnostic	p10
1.11) Typage moléculaire	p13
1.12) Évolution du typage moléculaire d' <i>E. multilocularis</i>	p15
CHAPITRE II : OBJECTIFS	p17
CHAPITRE III : MATÉRIEL ET MÉTHODES	p18
3.1) Table des échantillons	p18
3.2) Matériel et équipements communs	p18
3.3) Extraction de l'ADN	p19
3.4) Protocole de conception des amorces PCR	p20
3.5) Protocole de manipulation PCR	p21
3.6) Purification de l'ADN	p23
3.7) Quantification des amplicons purifiés	p25
3.8) Séquençage Illumina	p26
3.9) Séquençage Nanopore	p27
3.10) Nombre de séquençages et contrôles négatifs utilisés	p27
3.11) Pipeline bioinformatique Illumina et Nanopore	p28
3.12) Construction de l'arbre phylogénétique et du réseau d'haplotypes	p29
CHAPITRE IV : RÉSULTATS	p30
4.1) Séquençage Illumina iSeq100 d' <i>E. multilocularis</i>	p30
4.2) Séquençage Illumina iSeq100 d' <i>E. granulosus</i>	p31
4.3) Séquençage ONT MinION Mk1D d' <i>E. multilocularis</i> / Liège	p32
4.4) Séquençage ONT MinION Mk1D d' <i>E. granulosus</i> / Liège	p33
4.5) Séquençage ONT MinION Mk1D d' <i>E. multilocularis</i> / Besançon	p34
4.6) Séquençage ONT MinION Mk1D d' <i>E. granulosus</i> / Besançon	p35
4.7) Validation interne Illumina iSeq100	p36
4.8) Validation externe ONT MinION Mk1D	p37
4.9) Validation interne ONT MinION Mk1D	p38
4.10) Haplogroupes d' <i>E. multilocularis</i>	p39
4.11) Génotypes d' <i>E. granulosus</i>	p40
CHAPITRE V : DISCUSSION	p41
5.1) Optimisation de la méthode d'extraction d'ADN	p41

5.2) Validation des contrôles de l'Illumina iSeq100 et du Nanopore MinION Mk1D	p42
5.3) Analyse des métriques de séquençage : Illumina vs Nanopore.....	p42
5.4) Validation méthodologique du séquençage Nanopore et Illumina	p44
5.5) Diversité génétique <i>d'E. multilocularis</i> et <i>d'E. granulosus</i> en Belgique	p45
5.6) Comparaison du typage EmsB et du séquençage du mitogénome complet	p46

<u>CHAPITRE VI : CONCLUSIONS ET PERSPECTIVES</u>	p47
---	-----

<u>CHAPITRE VII : BIBLIOGRAPHIE</u>	p48
--	-----

<u>ANNEXES</u>	Non paginé
-----------------------------	------------

NOTE SUR L'UTILISATION DE L'IA

J'ai utilisé ChatGPT (version GPT-5.5 d'OpenAI) ainsi que Grammarly (version 14.1304.0) afin de corriger la syntaxe, la grammaire et l'orthographe de ce mémoire. Ces outils m'ont également permis de reformuler mes phrases afin d'en améliorer la structure et la clarté.

J'ai aussi utilisé Claude Opus 4.8 (Anthropic) pour m'aider à corriger et améliorer mes scripts R et Bash.

LISTE DES ABRÉVIATIONS

ADN : Acide désoxyribonucléique
ALL : *Ensemble des variants filtrés*
ALT : Allèle alternatif
ARN : Acide ribonucléique
BED : *Browser Extensible Data*
BLAST : *Basic Local Alignment Search Tool*
CE : Cystic echinococcosis
CHU : Centre Hospitalier Universitaire
CNR : Centre National de Référence
CNR-E : Centre National de Référence des Échinococcoses
CT-Scan : Tomodensitométrie
ddNTP : Didésoxynucléotide
DNA : *Deoxyribonucleic acid*
DOI : *Digital Object Identifier*
DP : *Depth*
dsDNA : ADN double brin
EA : Échinococcose alvéolaire
ECDC : *European Centre for Disease Prevention and Control*
EK : Échinococcose kystique
ELISA : *Enzyme-Linked Immunosorbent Assay*
E. multilocularis : *Echinococcus multilocularis*
EmsB : *Echinococcus multilocularis* satellite B
E. granulosus : *Echinococcus granulosus*
GC : Guanine-cytosine
HG : Haplogroupe
HG1 : Haplogroupe 1
HG2 : Haplogroupe 2
HG3 : Haplogroupe 3
HG3a : Sous-haplogroupe HG3a
HG3b : Sous-haplogroupe HG3b
HG3c : Sous-haplogroupe HG3c
HG4 : Haplogroupe 4
HS : High Sensitivity
ID : Identifiant
IDT : *Integrated DNA Technologies*
IgG : Immunoglobuline G
INDEL : Insertion-délétion
IRM : Imagerie par résonance magnétique
kb : Kilobase
LNRE : Laboratoire National de Référence des Échinococcoses
MAPQ : *Mapping Quality*
NCBI : National Center for Biotechnology Information
NGS : *Next-Generation Sequencing*
ONT : *Oxford Nanopore Technologies*
OMS : Organisation mondiale de la santé

PacBio : Pacific Biosciences

PAIR : Ponction-Aspiration-Injection-Réaspiration

PAS : Acide périodique de Schiff

pb : Paires de bases

PCR : Polymerase Chain Reaction

PET-Scan : Tomographie par émission de positrons

PNM : *Parasitic mass, Neighbouring organs, Metastasis* (Masse parasitaire, Organes voisins, Métastases)

qPCR : PCR quantitative

Q-score : Score de qualité Phred

QUAL : Qualité du variant

RCF : *Relative Centrifugal Force*

REF : Référence

s.l. : *sensu lato*

s.s. : *sensu stricto*

SMRT : Single-Molecule Real-Time

SNP : Single Nucleotide Polymorphism

SPRI : Solid Phase Reversible Immobilization

WB : Western Blot

G1 : Génotype 1 d'*Echinococcus granulosus*

G3 : Génotype 3 d'*Echinococcus granulosus*

G4 : Génotype 4 d'*Echinococcus granulosus*

G5 : Génotype 5 d'*Echinococcus granulosus*

G6-G7 : Génotypes 6 et 7 d'*Echinococcus granulosus*

G8-G10 : Génotypes 8 à 10 d'*Echinococcus granulosus*

cox1 : Gène mitochondrial codant la sous-unité 1 de la cytochrome c oxydase

nad1 : Gène mitochondrial codant la sous-unité 1 de la NADH déshydrogénase

rrnS : Gène mitochondrial codant la petite sous-unité de l'ARN ribosomal mitochondrial

RÉSUMÉS EN ANGLAIS ET EN FRANÇAIS

Advancing Echinococcus Mitogenome Sequencing: Illumina iSeq100 vs Nanopore MinION Mk1D Benchmarking and Future Methodological Refinements

Background: Complete mitochondrial genomes provide high-resolution phylogenetic markers for lineage assignment and population-structure analyses of *Echinococcus granulosus* (*E. granulosus*) and *Echinococcus multilocularis* (*E. multilocularis*), the parasitological agents of cystic and alveolar echinococcosis, respectively. However, whole-mitogenome sequencing from clinical material depends on robust laboratory workflows and an evidence-based choice of sequencing platform. Here, a benchmark was performed between Illumina and Oxford Nanopore Technologies (ONT) for *Echinococcus* whole-mitogenome sequencing of 20 Belgian and 6 French clinical samples.

Methods: Twenty-six clinical samples were analyzed, including 19 specimens from 15 alveolar echinococcosis patients and 7 specimens from 7 cystic echinococcosis patients. Mitogenomes were generated using a tiled PCR approach with overlapping amplicons (2.0–2.5 kb) spanning the mitochondrial genome, followed by sequencing on Illumina iSeq100 and Nanopore MinION Mk1D. Platforms were compared using consensus concordance, haplogroup assignment, and run-level details. Bioinformatic processing included read quality control, consensus generation, variant calling, and phylogenetic placement.

Results: Complete mitogenomes enabled clade assignment and an initial characterization of haplogroup diversity among Belgian cases. *E. multilocularis* isolates clustered within three European sub-haplogroups (HG3a, HG3b, and HG3c), whereas *E. granulosus* isolates segregated into the genotype 1 (G1). For matched isolates, Illumina and Nanopore produced concordant phylogenetic placement. Nanopore MinION MK1D facilitated improved resolution of repetitive regions, while Illumina delivered greater depth and stronger per-variant quality metrics.

Conclusions and Perspectives: Mitochondrial genomes provide a reliable framework for *Echinococcus* genotyping. Nanopore sequencing improves resolution in repetitive regions, while a better extraction method could enable a long-range PCR approach, thereby providing better characterization of the repetitive sequences of *E. granulosus*. Another long-term objective is to test this method on the main definitive hosts of *E. multilocularis* (red foxes, *Vulpes vulpes*) at a larger scale.

Évaluation comparative d'Illumina iSeq100 et de Nanopore MinION Mk1D pour le séquençage complet des mitogénomes d'*Echinococcus* et perspectives d'optimisation méthodologique

Contexte : Les génomes mitochondriaux complets constituent des marqueurs phylogénétiques robustes pour l'assignation des lignées et les analyses de structure des populations d'*Echinococcus granulosus* (*E. granulosus*) et d'*Echinococcus multilocularis* (*E. multilocularis*), agents parasitaires responsables respectivement de l'échinococcose kystique et de l'échinococcose alvéolaire. Cependant, le séquençage de mitogénomes complets à partir d'échantillons cliniques dépend de protocoles de laboratoire robustes et d'un choix raisonné de la plateforme de séquençage. Dans ce travail, les technologies Illumina et Oxford Nanopore Technologies (ONT) ont été comparées sur une cohorte de 20 échantillons belges et de 6 échantillons français.

Méthodes : Vingt-six échantillons cliniques ont été analysés, dont 19 prélèvements issus de 15 patients atteints d'échinococcose alvéolaire et 7 prélèvements issus de 7 patients atteints d'échinococcose kystique. Les mitogénomes ont été générés à l'aide d'une approche de PCR par tuilage, reposant sur des amplicons chevauchants de 2,0 à 2,5 kb couvrant l'ensemble du génome mitochondrial, suivie d'un séquençage sur Illumina iSeq100 et sur Nanopore MinION Mk1D. Les plateformes ont été comparées en termes de complétude, de concordance des consensus, d'assignation des haplogroupes et de caractéristiques propres à chaque run. Le traitement bioinformatique comprenait le contrôle de la qualité des lectures d'ADN (*reads*), la génération de consensus, l'appel de variants et le placement phylogénétique.

Résultats : Les mitogénomes complets ont permis d'assigner les lignées à des branches phylogénétiques. Les isolats d'*E. multilocularis* se regroupent en trois sous-haplogroupes européens, HG3a, HG3b et HG3c, tandis que les isolats d'*E. granulosus* se rattachent au génotype 1 (G1). Pour les isolats appariés, les technologies Illumina et Nanopore produisent un placement phylogénétique concordant. La technologie Nanopore offre une meilleure résolution des régions répétées, tandis que celle d'Illumina fournit une profondeur de séquençage plus élevée et des métriques de qualité par variant plus robustes.

Conclusions et perspectives : Les génomes mitochondriaux constituent un cadre fiable pour le génotypage d'*Echinococcus*. La technologie Nanopore améliore la résolution des régions répétées, tandis qu'une meilleure méthodologie d'extraction pourrait permettre une approche de PCR longue, offrant une caractérisation plus fine des séquences répétées d'*E. granulosus*. Ce projet a aussi pour but de préparer une analyse à grande échelle des principaux hôtes définitifs en Belgique d'*E. multilocularis* (les renards roux, *Vulpes vulpes*).

Chapitre I : Introduction



CHAPITRE I : INTRODUCTION

1.1) L'histoire évolutive d'un couple de parasites multicellulaires

L'histoire de l'échinococcose alvéolaire (EA) et kystique (EK) ne saurait être réduite à un simple enchaînement de découvertes scientifiques¹. Cette histoire s'inscrit dans un tissu de contextes et de bouleversements sociaux, de paradigmes biologiques en constante évolution et de rivalités intellectuelles^{1,2}. L'année 1852 marque une inflexion épistémologique : Ludwig Buhl, pathologiste bavarois, décrit une affection hépatique atypique, l'EA³. Trois ans plus tard, Rudolf Virchow réinterprète cette même lésion, cette fois à la lumière d'une observation rare : la présence de protoscolex (ou formes larvaires précoces de ténia) dans les vésicules parasitaires². Cette observation fortuite et hautement importante pour la trajectoire historique de la maladie ancre définitivement l'EA comme entité parasitaire^{1,2,3}. C'est ainsi que le regard change : d'une tumeur à une zoonose identifiable. Cette première localisation dans le sud de l'Allemagne, puis en Suisse et en Autriche, a figé pendant près d'un siècle la perception de l'EA comme une pathologie parasitaire européenne^{1,4}.

Par la suite, le dualisme parasitologique s'imposa lentement. Si R.L. Virchow voyait dans l'EA une variété du *Taenia*, d'autres, tels que A. Posselt et A. Morin¹, pressentaient qu'un autre parasite, *Echinococcus multilocularis* (*E. multilocularis*) était à l'œuvre (**figure 1.1**)¹. Ce conflit, qui a duré près d'un siècle, reflète à la fois la lenteur des preuves expérimentales et l'inertie des paradigmes scientifiques. Cependant, le débat ne fut pas seulement un conflit de microscopes. Il reflétait également des clivages culturels, linguistiques et géopolitiques. Les publications françaises, allemandes, russes, japonaises et chinoises n'étaient ni traduites ni citées, ce qui excluait, de fait, des pans entiers de savoir¹. Ainsi, les premiers cas humains chinois d'EA documentés dès les années 1950 restèrent invisibles aux yeux de la communauté anglophone⁵. L'ouverture politique de la Chine a permis à des chercheurs occidentaux comme Phil Craig de révéler l'ampleur du problème, car la région aujourd'hui la plus endémique (la Chine) fut longtemps absente de la carte épidémiologique mondiale^{1,5}.

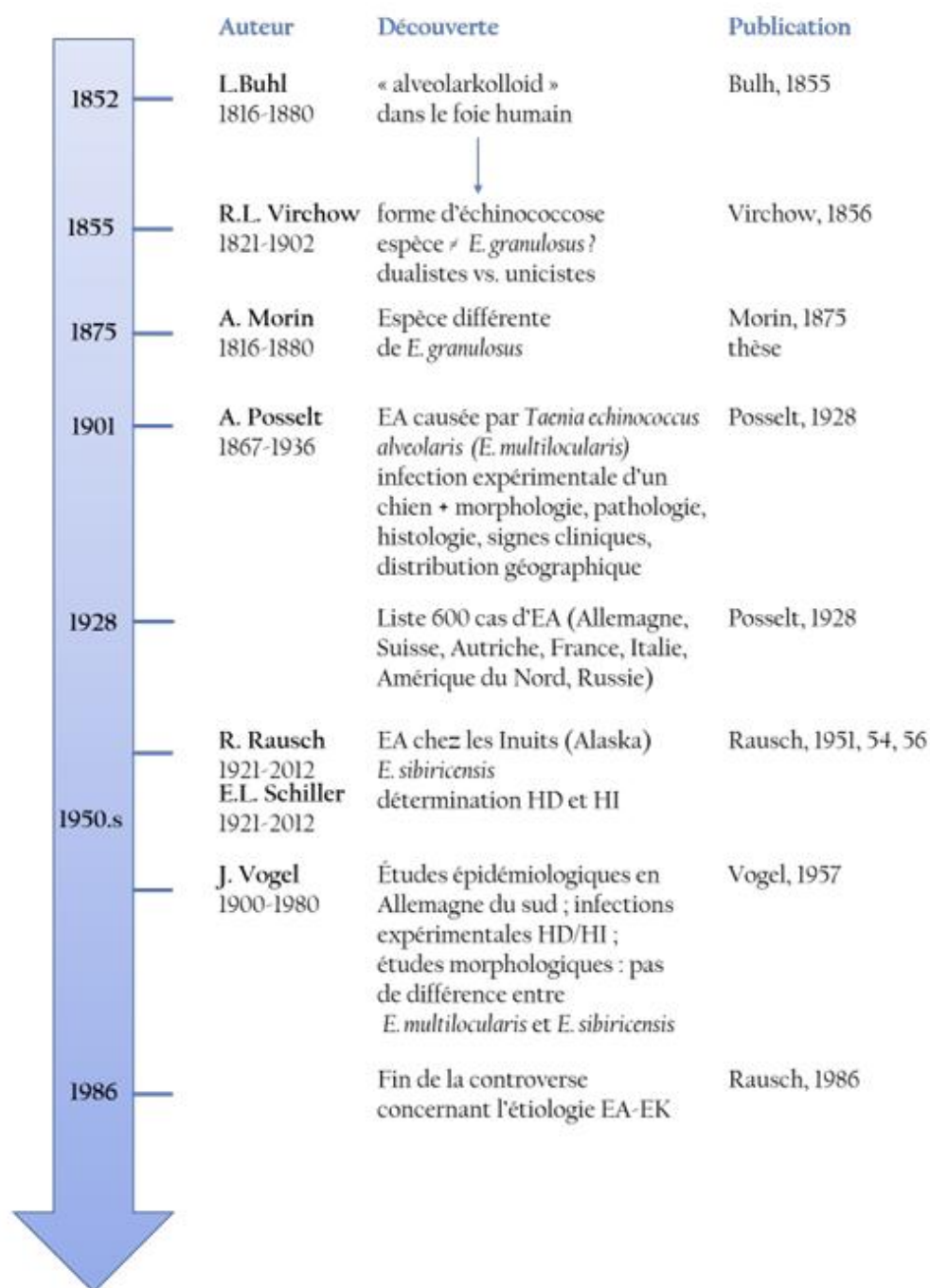


Figure 1.1 : Ligne du temps représentant l'évolution historique des découvertes scientifiques majeures sur les deux principales espèces d'*Echinococcus* : *E. multilocularis* et *E. granulosus*¹³.

L'EA a longtemps été perçue comme une maladie européenne ; cependant, elle est en réalité d'origine sibérienne¹. Par la suite, la progression des méthodes scientifiques de diagnostic, de dépistage et de surveillance des pathogènes a reconfiguré notre vision épidémiologique. Elle a mis en lumière l'influence de facteurs structurels, tels que la densité des rongeurs, les transformations agricoles, l'urbanisation des renards et les systèmes d'élevage intensifs, sur la dynamique du cycle parasitaire et la dissémination de ces deux agents pathogènes^{1,6,32}. Loin d'un modèle simpliste et linéaire, l'EA et l'EK s'inscrivent dans une écologie sociale et environnementale définie et influencée par un changement des habitudes humaines^{1,6,32}. C'est à ce titre qu'écrire une histoire de l'échinococcose ne consiste pas seulement à retracer la chronologie des cas, mais à interroger la manière dont les concepts scientifiques sont influencés par des contextes géographiques et sociaux.

1.2) Taxonomie du genre *Echinococcus* et distribution mondiale

Le genre *Echinococcus* regroupe des parasites appartenant à l'embranchement des Platyhelminthes, à la classe des Cestoda (Eucestoda) et à la famille des Taeniidae^{20,32}. La taxonomie du genre a longtemps été un terrain instable (classifications ambiguës des différentes espèces et du genre)^{20,32}. Ce n'est qu'avec l'intégration des données de séquences d'ADN qu'un arbre phylogénétique plus solide s'est imposé (1995)⁵⁵. *Echinococcus* est un genre qui comprend cinq espèces génétiquement confirmées (*E. multilocularis*, *E. granulosus*, *E. shiquicus*, *E. vogeli* et *E. oligarthra*), dont quatre sont responsables de pathologies chez l'humain (*E. granulosus*, *E. multilocularis*, *E. vogeli*, *E. oligarthra*) (**figure 1.2**)^{20,32,48}. Cette arborescence se traduit par une géographie très asymétrique des principales espèces, à savoir *E. granulosus* et *E. multilocularis* (**figure 1.3**)^{20,32,48}. L'EK, causée par *E. granulosus sensu lato* (s.l.), correspond à un complexe regroupant plusieurs sous-espèces et génotypes historiquement décrits du génotype 1 (G1) au génotype 10 (G10). Du génotype G1 au génotype G3, on parle du complexe *E. granulosus sensu stricto* (s.s.), puis, du génotype G4 au génotype G10, la dénomination des sous-espèces est *E. equinus* (G4), *E. ortleppi* (G5), *E. intermedius* (G6-G7) et *E. canadensis* (G8-G10)⁴⁸.

E. granulosus est présent dans les deux hémisphères, avec des zones d'endémie importantes en Chine occidentale, en Asie centrale, en Amérique du Sud, dans le bassin méditerranéen et en Afrique de l'Est (**figure 1.3**)³².

À l'inverse, *E. multilocularis* présente une distribution principalement limitée à l'hémisphère nord, notamment en Amérique du Nord, en Sibérie et en Europe centrale (**figure 1.3**)³².

Species	Strain/Genotype	Known Intermediate Hosts	Known Definitive Hosts	Infectivity to Humans	Disease
<i>Echinococcus granulosus</i>	Sheep/G1	Sheep (cattle, pigs, camels, goats, macropods)	Dog, fox, dingo, jackal and hyena	Yes	CE
	Tasmanian sheep/G2	Sheep (cattle?)	Dog, fox	Yes	CE
	Buffalo/G3	Buffalo (cattle?)	Dog, fox?	Yes	CE
<i>Echinococcus equinus</i>	Horse/G4	Horses and other equines	Dog	Probably not	CE?
<i>Echinococcus ortleppi</i>	Cattle/G5	Cattle	Dog	Yes	CE
<i>Echinococcus canadensis</i>	Cervids/G8,G10	Cervids	Wolves, dog	Yes	CE
<i>Echinococcus intermedius</i>	Camel/Pig/G6/G7	Camels, pigs, sheep	Dog	Yes	CE
<i>Echinococcus felidis</i>	Lion, dog	Warthog, (zebra, wildebeest, bushpig, buffalo, various antelope, giraffe Hippopotamus?)	Lion	?	-
<i>Echinococcus multilocularis</i>	Some isolate variation	Rodents, domestic and wild pig, dog, monkey, (horse?)	Fox, dog, cat, wolf, racoon-dog, coyote	Yes	AE
<i>Echinococcus shiquicus</i>	?	Pika and rodents	Tibetan fox and?	?	?
<i>Echinococcus vogeli</i>	None reported	Rodents	Bush dog	Yes	PE
<i>Echinococcus oligarthra</i>	None reported	Rodents	Wild felids	Yes	PE

Updated from References [2,3]. CE = Cystic Echinococcosis, AE = Alveolar Echinococcosis, PE = Polysystic

Figure 1.2 : Tableau illustrant les différentes espèces d'*Echinococcus*, ainsi que leurs hôtes respectifs et le type de pathologie associée⁴⁸.

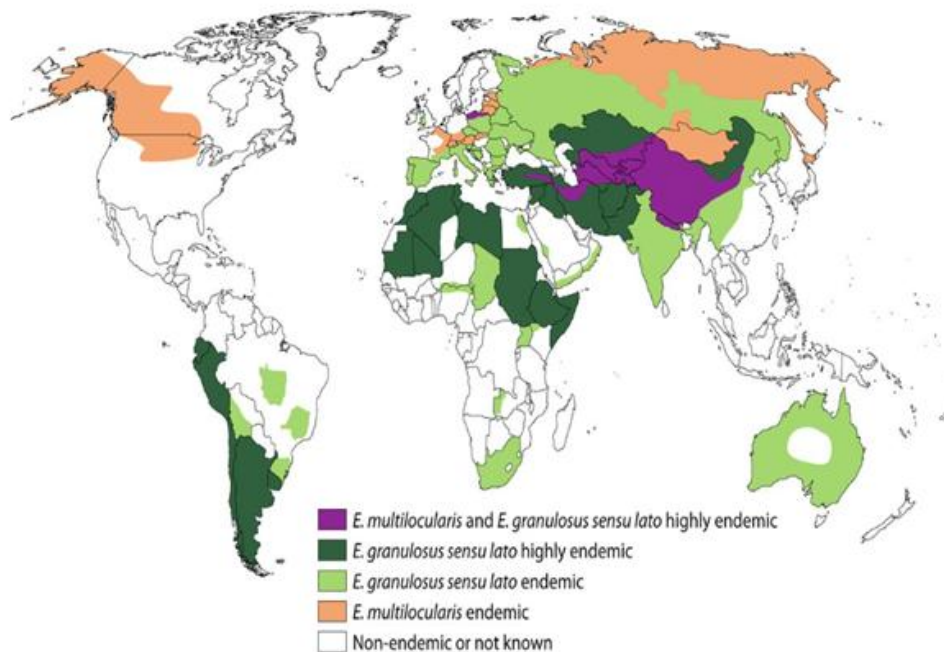


Figure 1.3 : Carte illustrant la distribution mondiale de cas humains d'*E. multilocularis* et d'*E. granulosus*³².

1.3) Épidémiologie de l'EA et de l'EK en Europe et en Belgique

En 2023, le centre européen de prévention et de contrôle des maladies (*European Centre for Disease Prevention and Control*, ECDC) a recensé 929 cas humains confirmés d'échinococcose, soit un taux de notification de 0,21 cas pour 100 000 habitants, ce qui correspond à une augmentation d'environ 10 % par rapport à 2022 (0,19/100 000 habitants)³³. Parmi les 624 cas pour lesquels l'espèce est renseignée, *E. granulosus s.l.* représente 67,3 % des cas et *E. multilocularis*, 32,7 %. Concernant les espèces animales, *E. multilocularis* a été détecté dans 9 États membres de l'UE ainsi que dans un pays non-membre (la Suisse)³³. En Europe, le renard roux (*Vulpes vulpes*), principal hôte définitif d'*E. multilocularis*, est l'espèce animale la plus fréquemment testée positive³³. Les taux de prévalence les plus élevés ont été rapportés en Pologne (31,2 %), en Tchéquie (22,0 %), en Allemagne (19,7 %), en Slovénie (17,6 %) et en France (15,2 %). En parallèle, *E. granulosus s.l.* est détecté par 12 États membres et un pays non-membre (la Norvège), dont la majorité des analyses concerne les ovins/caprins (11 543 410 animaux testés, 0,66 % positifs), les bovins (6 794 791 testés, 0,28 % positifs) et les porcs (55 162 890 testés, < 0,01 % positifs)³³. L'Italie et l'Espagne présentaient les plus fortes proportions de cas positifs pour ces différentes catégories d'animaux³³.

En 2024, la Belgique a notifié 11 nouveaux cas humains d'EA, soit une incidence de 0,09/100 000 habitants⁴⁰. Ce niveau est stable par rapport aux années 2022 et 2023, avec respectivement 12 et 10 cas⁴⁰. Parmi les 11 cas, dix sont autochtones (infections acquises en Belgique) et un a été contracté au Tibet⁴⁰. Tous les cas autochtones se concentrent en Wallonie, ce qui suggère une transmission géographiquement structurée à l'échelle nationale (**figure 1.4**)⁴⁰. Pour l'EK, 20 nouveaux cas sont rapportés sur l'ensemble du territoire, soit 0,17/100 000 habitants⁴⁰. L'incidence est stable par rapport à 2023, mais en progression par rapport à 2022 (0,11/100 000 habitants) et 2021 (0,08/100 000 habitants)⁴⁰. La dynamique épidémiologique est ici qualitativement différente, car 100 % des cas sont importés de régions endémiques, telles que l'Afrique du Nord, l'Europe de l'Est et le Moyen-Orient.

1.4) Cycle parasitaire d'*E. multilocularis*

1.4.1) Hôte définitif : une présence silencieuse

Chez l'hôte définitif, le plus souvent le renard roux (*Vulpes vulpes*) ou, très rarement, un coyote (*Canis latrans*), le parasite vit dans l'intestin grêle sous forme de ver adulte (**étape 1, figure 1.5**)^{8,9,10}. Ces vers, mesurant de 1,2 à 4,5 mm, sont constitués de 3 à 5 proglottides^{8,9,10}. Chaque ver adulte peut produire des centaines d'œufs embryonnés (**étape 2, figure 1.5**), excrétés dans les fèces de l'hôte définitif^{8,9,10}.

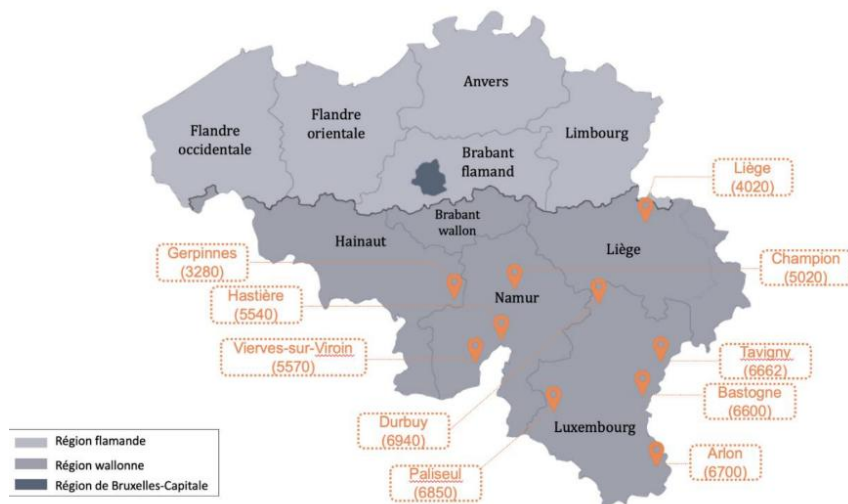


Figure 1.4 : Carte épidémiologique représentant la localisation des cas autochtones Belges d'*E. multilocularis* au cours de l'année 2024⁴⁰.

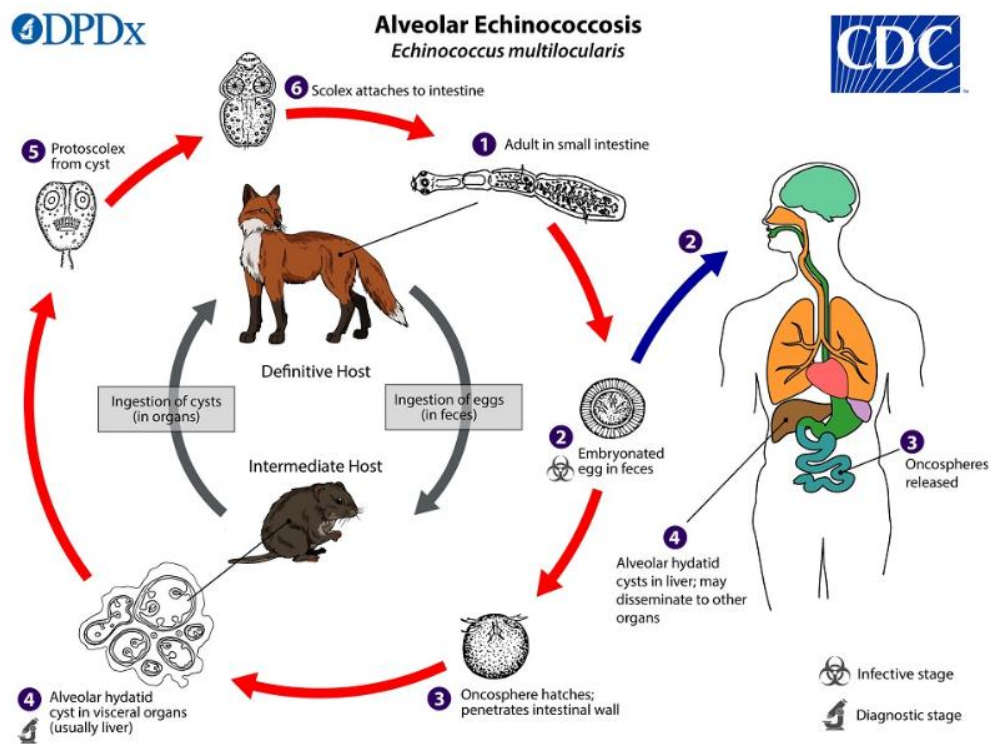


Figure 1.5 : Figure représentant le cycle d'infection d'*E. multilocularis* qui schématise le stage de développement larvaire du parasite et les différents hôtes successifs qu'il va infecter¹⁰.

Les œufs ainsi libérés (**étape 2, figure 1.5**) sont dotés d'un embryophore épais et radialement strié. Ils sont immédiatement infectieux et résistent longtemps dans l'environnement, parfois plusieurs mois, y compris dans des lieux et des sols extrêmement différents^{8,9,10,11}.

1.4.2) Hôte intermédiaire : l'entrée dans la latence proliférative

Lorsqu'un petit rongeur, typiquement un campagnol (*Arvicola amphibius*), un lemming ou un mulot, ingère accidentellement un œuf, celui-ci libère une larve hexacanthé (oncosphère) dans l'intestin (**étape 3, figure 1.5**)^{8,9,10}. L'oncosphère, équipée de six crochets chitineux (crochets qui permettent une bonne fixation à la muqueuse intestinale), traverse la paroi intestinale, rejoint le système porte et migre principalement vers le foie^{8,9,10}. Ensuite, l'oncosphère se métamorphose en métacestode alvéolaire. Ce métacestode est la forme larvaire pluricellulaire invasive (**étape 4, figure 1.5**)^{8,9,10}. Contrairement à l'unicysticité bien délimitée d'*E. granulosus* par une membrane fibreuse produite par l'hôte (l'adventice), la lésion d'*E. multilocularis* est multiloculaire. Cette lésion croît par bourgeonnement exogène et s'étend dans le parenchyme hépatique. De ce fait, l'EA rappelle le comportement des tumeurs malignes, tant par sa croissance que par son extension (**étape 4, figure 1.5**)^{8,9,10}. Au sein de cette lésion se forment des protoscolex, bien que leur présence ne soit pas systématique, surtout chez les hôtes aberrants (les humains) (**étape 5, figure 1.5**)^{8,9,10}.

1.4.3) Bouclage du cycle et synthèse évolutive

Le cycle n'est bouclé que si un carnivore prédateur consomme un rongeur parasité par des protoscolex (**étape 5, figure 1.5**)^{8,9,10}. Les protoscolex sont déversés dans l'intestin du nouvel hôte définitif, s'accrochent à la muqueuse intestinale par leur scolex et atteignent leur maturité en un à deux mois (**étape 6, figure 1.5**)^{8,9,10}. Ainsi, chaque étape du cycle d'*E. multilocularis* illustre une forme de dépendance écologique intégrée à un réseau d'interactions : le renard se nourrit du rongeur et partage ses espaces ; l'homme cultive les mêmes terres que celles fréquentées par ces rongeurs et ces renards^{8,9,10}. En exploitant les interactions étroites entre la faune sauvage, les animaux domestiques, l'environnement et les populations humaines, *E. multilocularis* s'inscrit pleinement dans la problématique *One Health*^{8,9,10}.

1.4.4) Cas des hôtes aberrants : impasse parasitaire

L'humain peut également ingérer les œufs par contact avec des fèces de renards ou avec des surfaces contaminées (par exemple, via la consommation de fruits sauvages comme des mûres ou des myrtilles). Il devient alors un hôte intermédiaire aberrant au sein d'un cycle sylvatique (forestier)^{8,9,10}. Chez l'humain, le cycle s'interrompt : la larve se développe, mais aucune transmission interhumaine n'est possible^{8,9,10}.

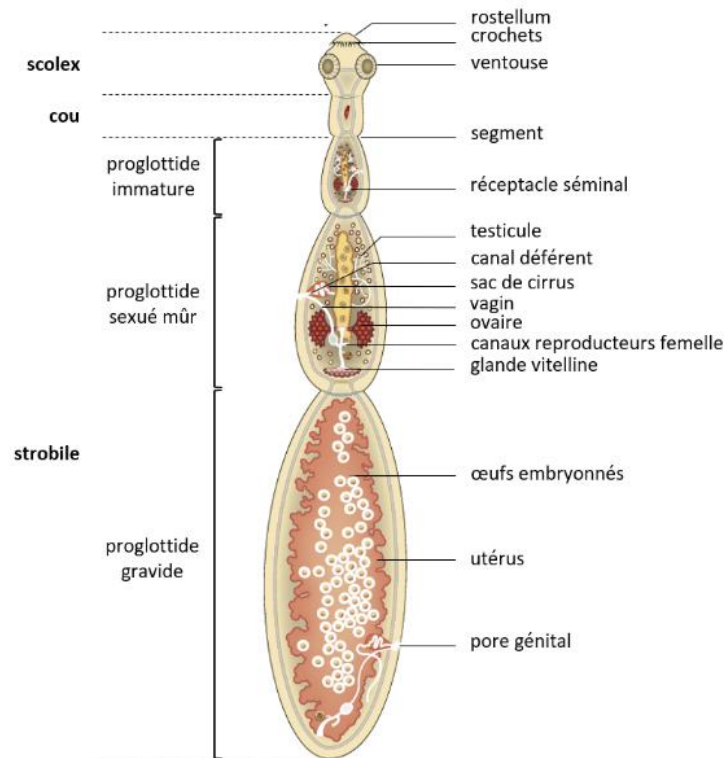


Figure 1.6 : Représentation du ver adulte d'*E. multilocularis*¹³.

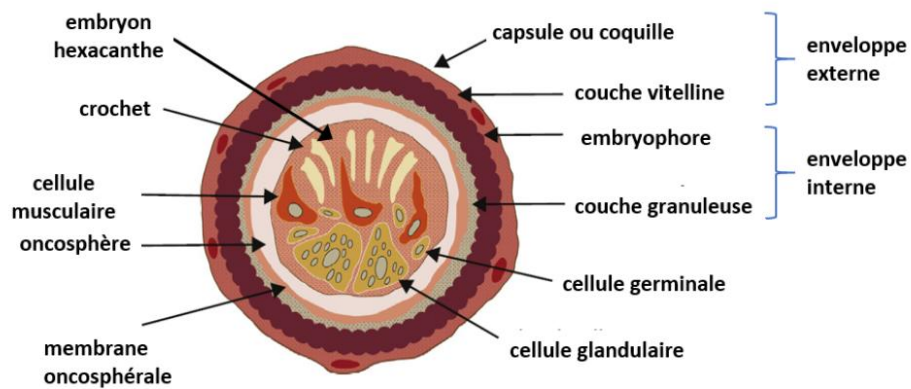


Figure 1.7 : Représentation d'un œuf d'*E. multilocularis*, typique d'un œuf de *Taenia*¹³.

1.5) Morphologie d'*E. multilocularis*

1.5.1) Morphologie du ver adulte et de l'œuf d'*E. multilocularis*

Le ver adulte est un cestode très court, de 1,2 à 4,5 mm, rarement composé de plus de cinq proglottides (**figure 1.6**)^{13,14}. Son scolex, situé à l'extrémité antérieure, est typique des Taeniidae^{13,14}. Il comporte quatre ventouses musculaires circulaires et un rostellum armé de deux couronnes concentriques de crochets chitineux, ce qui lui permet de s'ancrer à la muqueuse intestinale de l'hôte définitif (**figures 1.5 et 1.6**)^{13,14}. Au microscope, la distinction entre *E. multilocularis* et *E. granulosus* repose sur quelques caractéristiques morphologiques, telles que la taille du scolex, la forme des crochets et la localisation des pores génitaux. Ces critères demeurent toutefois insuffisants pour une identification certaine, qui nécessite généralement une confirmation moléculaire^{13,14,15}.

Les œufs d'*E. multilocularis*, comme ceux d'*E. granulosus*, mesurent environ 30 à 35 µm de diamètre. Ils présentent une coque épaisse striée radialement^{13,14,15} (**figure 1.7**). Chaque œuf renferme une larve hexacanthé (ou oncosphère) dotée de six crochets embryonnaires, difficilement visibles à l'observation au microscope.

1.5.2) Morphologie du métacestode : une prolifération organisée

C'est dans sa forme larvaire qu'*E. multilocularis* déploie toute sa particularité morphologique. Les lésions alvéolaires, qui se développent essentiellement dans le foie de l'hôte intermédiaire (rongeurs, humains ou, rarement, les chiens comme hôtes intermédiaires accidentels)^{8,9,10}, ne sont pas des cavités uniques bordées d'une membrane épaisse^{8,9,10}. Ce sont des masses infiltrantes, constituées d'une grande quantité d'alvéoles interconnectées, formés par bourgeonnement exogène (**figure 1.8.1**)^{8,9,10,13}. Chaque alvéole est une petite vésicule remplie de liquide, tapissée d'une membrane germinative pluripotente, capable de produire à la fois de nouvelles lésions et, dans certains cas, des protoscolex (**figure 1.8.1**)^{8,9,10}. Toutefois, chez les hôtes aberrants (les humains, les chiens), les protoscolex sont souvent absents et la prolifération résulte alors exclusivement de l'expansion tissulaire de la membrane germinative^{8,9,10,13}.

Ainsi, la lésion de l'EA est un organe à part entière, autorépliatif, capable de croître dans toutes les dimensions de l'espace qu'il occupe (**figure 1.8.1**)⁴⁹.

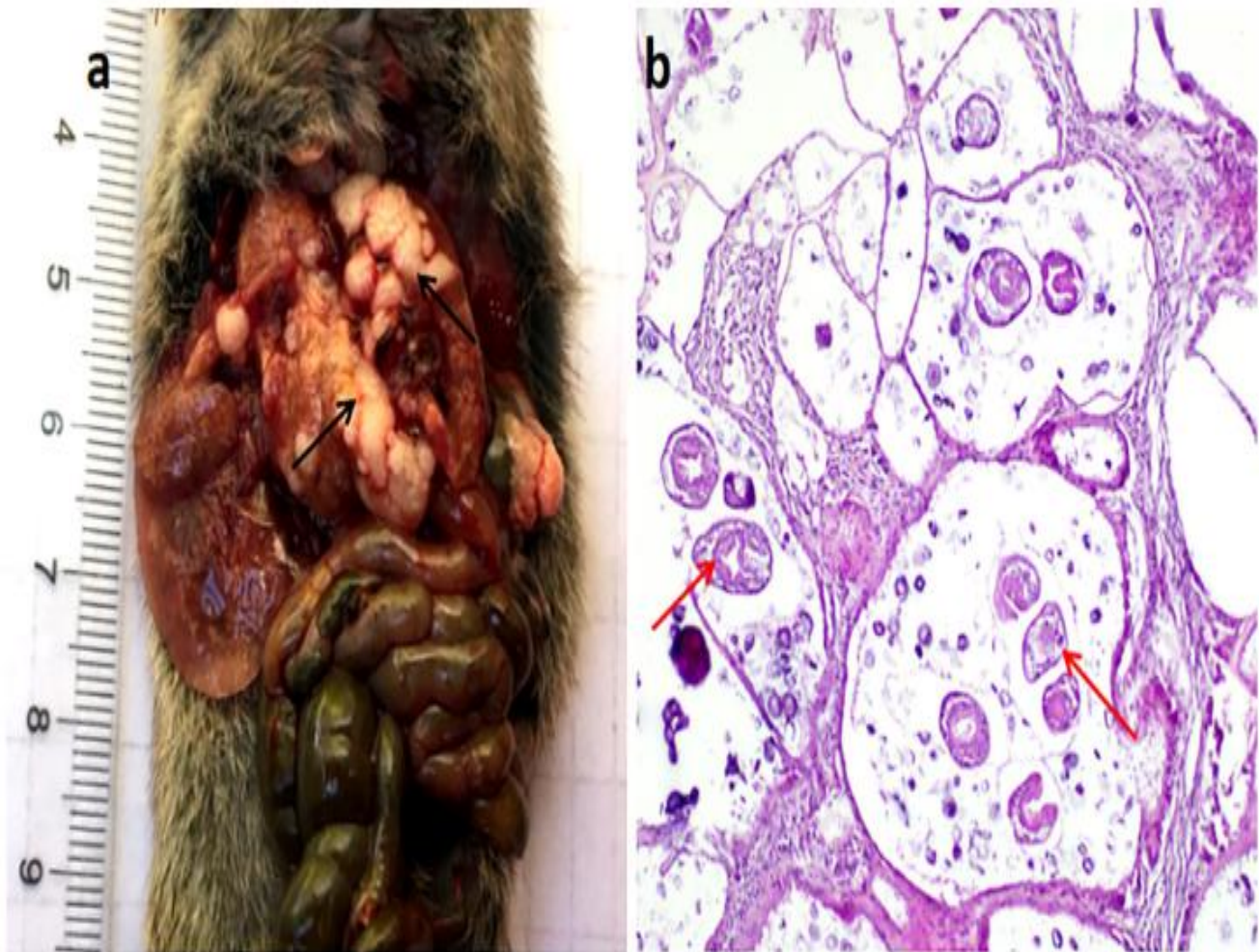


Figure 1.8.1 : Lésion hépatique chez un petit rongeur infecté par *E. multilocularis*. (a) Métacestode d'*E. multilocularis* observé dans le foie. (b) Coupe histologique hépatique colorée à l'hématoxyline-éosine montrant des lésions kystiques et la présence de protoscolex, indiqués respectivement par les flèches noire et rouge. Grossissement original : $\times 200^{49}$.

1.6) L'EA : de la physiopathologie à la prise en charge du patient

Au stade de métacestode, le parasite impose sa propre logique architecturale au foie^{12,13,18,32,34}. La masse larvaire se constitue d'abord d'un réseau de microvésicules bourgeonnantes, dont la surface externe est une couche laminée acellulaire, riche en glucides, qui joue le rôle de barrière protectrice et d'écran antigénique^{13,32}. Sous cette enveloppe se déploie un tégument syncytial, prolongé par des microtriches, ancrées dans la couche laminée, qui assurent les échanges avec le milieu environnant^{12,13,18,32,34}. Ensuite, une couche germinative vivante concentre les éléments essentiels à la croissance : cellules musculaires et conjonctives, réserves de glycogène, cellules nerveuses et, surtout, un contingent de cellules souches capables de régénérer en permanence de nouvelles vésicules^{12,32}. Avec le temps, cette structure larvaire prolifère de manière exogène et diffuse, infiltrant progressivement le parenchyme hépatique^{12,13,18,32,50} (**figure 1.8.2 A**). La lésion chronique typique se présente comme une grande cavité centrale nécrotique, remplie d'un matériel blanc, amorphe, entourée d'une fine couche fibreuse dense^{13,32}. Face à cette invasion, le foie ne reste pas passif. Autour des vésicules se mettent en place de larges granulomes constitués de macrophages, de lymphocytes T et de myofibroblastes. Ainsi, un tissu conjonctif réactionnel se développe, s'épaissit, se fibrose, tentant de cloisonner la progression du métacestode^{13,32,50} (**figure 1.8.2 B**). Cette réaction contribue aussi à la destruction progressive du parenchyme hépatique, à la déformation de l'architecture portale et, *in fine*, à l'insuffisance hépatique^{13,32}.

Les principaux traitements restent la résection chirurgicale par hépatectomie partielle (enlever précisément la partie du foie où se trouvent les lésions) et la prescription d'un traitement par albendazole en complément après et avant l'opération. Dans le cas où le patient ne peut pas suivre un traitement chirurgical, l'infectiologue peut prescrire uniquement l'albendazole, sans aucune procédure chirurgicale supplémentaire^{13,32}. Cependant, l'hépatotoxicité de ce médicament est très élevée et peut poser problème chez un patient présentant un foie fortement atteint par l'EA. Un second traitement est le mébendazole^{13,32}. Celui-ci est utilisé lorsque le patient présente une résistance au traitement classique (l'albendazole) ou ne le tolère pas correctement. Pour finir, la période d'incubation asymptomatique de l'EA se situe entre 5 et 15 ans, avec un taux de mortalité supérieur à 90 % chez les patients diagnostiqués plus de 10 ans après l'infection^{13,32}.

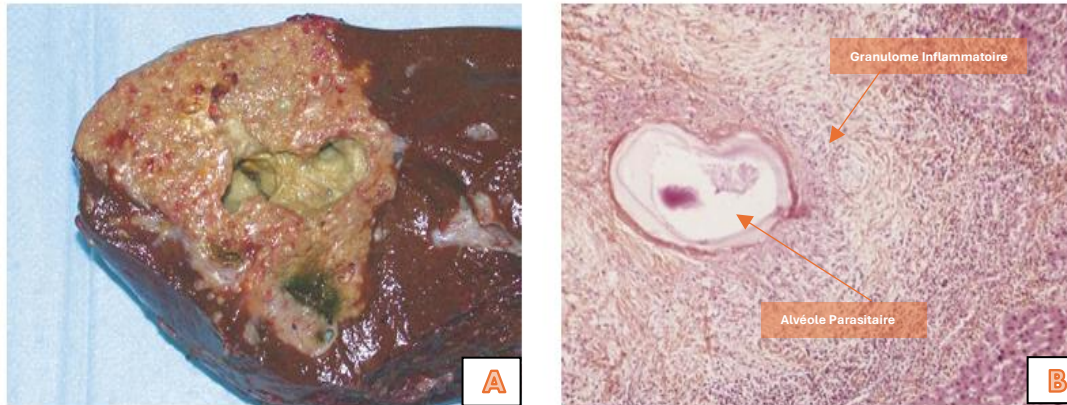


Figure 1.8.2 : Aspects macroscopiques et microscopiques des lésions d'EA hépatique⁵⁰.

- A. Coupe macroscopique du parenchyme hépatique révélant de multiples alvéoles parasitaires et une volumineuse cavité de nécrose centroparasitaire.
- B. Coupe histologique d'une lésion jeune en zone périphérique. Cette coupe met en évidence une intense réaction granulomateuse périparasitaire avec persistance de travées hépatocytaires (coloration : hématoxyne-éosine-safran).

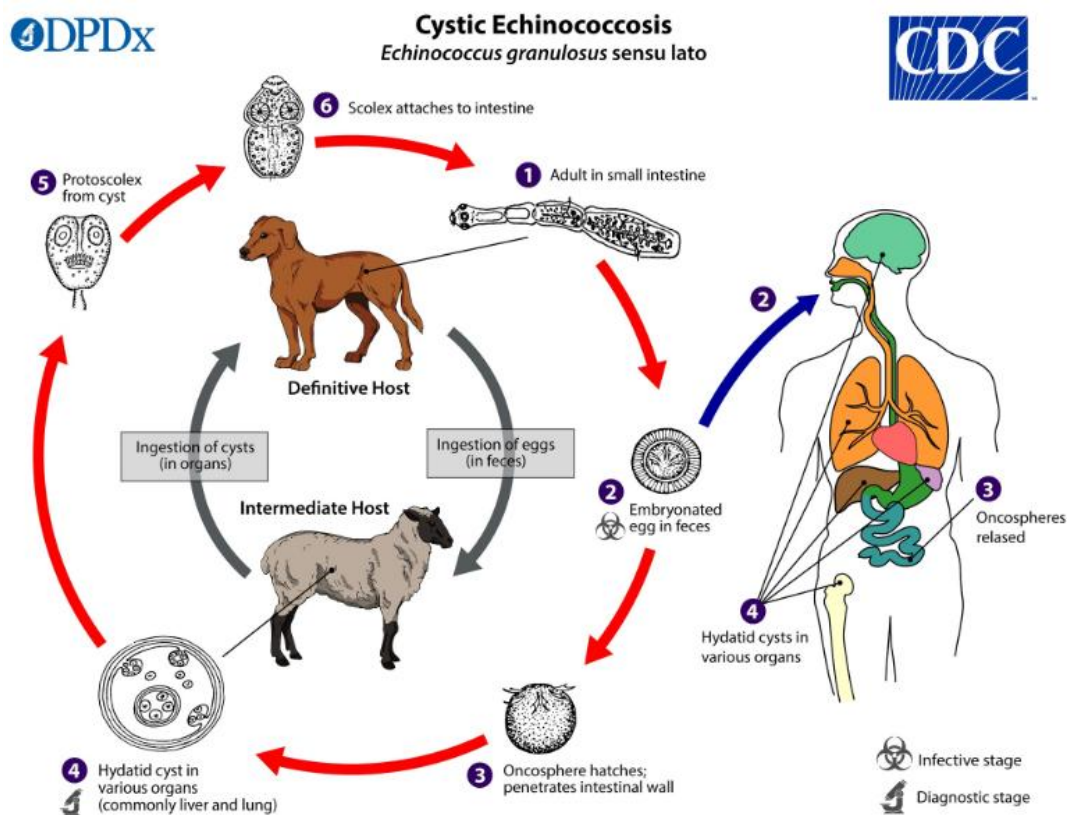


Figure 1.9 : Représentation du cycle d'infection d'*E. granulosus* avec une représentation des différents stades du développement larvaire et les différents hôtes qui se succèdent¹⁰.

1.7) Cycle parasitaire d'*E. granulosus*

1.7.1) Hôte définitif : transmission intestinale, émission d'œufs, dissémination

Chez l'hôte définitif, principalement le chien dans le cycle rural, mais aussi d'autres canidés selon les contextes, le ver adulte se fixe dans l'intestin grêle (**étape 1, figure 1.9**)^{10,32,35}. Il s'y fixe par son scolex (ventouses et rostre armé), s'insère au contact de la muqueuse et peut se loger entre les villosités jusqu'aux cryptes de Lieberkühn^{10,32,35}. Le ver ne provoque pas de lésions majeures dans la plupart des infections ordinaires^{10,32,35}. Sa fonction dominante n'est pas l'invasion, mais la production d'œufs^{10,32,35}. Les segments gravides assurent l'émission d'œufs, excrétés avec les fèces de l'animal, et ces œufs, présents dans l'environnement, sont infectieux dès leur libération (**étape 2, figure 1.9**)^{32,35}.

À ce stade, la transmission devient un fait de dispersion : les œufs contaminent le sol, les plantes, l'eau, le pelage de l'animal et peuvent être transportés mécaniquement par la poussière et la pluie^{32,35}. Le résultat n'est pas une contamination homogène, mais une pollution environnementale^{32,35}. Ainsi, tout comme *E. multilocularis*, *E. granulosus* présente une dynamique *One Health*.

1.7.2) Hôte intermédiaire : oncosphère, veine porte, métacestode kystique

L'hôte intermédiaire est typiquement un animal d'élevage (mouton, chèvre, bovin, porc, etc.) dans le cycle rural, ou, dans certains contextes, un cycle sylvatique (forestier) peut impliquer des cervidés et les deux cycles peuvent coexister ou s'interpénétrer^{32,35}. L'infection débute par l'ingestion d'œufs. Dans l'intestin, l'œuf libère une oncosphère (embryon hexacanthé), qui traverse la paroi intestinale et atteint la circulation porte. Le foie agit alors comme premier filtre. Une partie des oncosphères y est retenue ; d'autres peuvent atteindre les poumons et, plus rarement, d'autres tissus par dissémination hématogène (**étape 3, figure 1.9**)^{32,35}.

Dans l'organe cible, l'oncosphère évolue en métacestode sous la forme d'un kyste hydatique (**étape 4, figure 1.9**)^{32,35}. Ce kyste est une structure de croissance lente, organisée autour d'une paroi spécialisée (membrane laminée) et d'une zone proliférative (membrane germinative) à partir de laquelle se forment des capsules proligères. À l'intérieur de ces capsules se développent des protoscolex invaginés, unités biologiques capables, si elles sont ingérées par un canidé, de redonner l'adulte intestinal (**étape 5, figure 1.9**)^{32,35}.

1.7.3) Bouclage du cycle : protoscolex, évagination, maturité en quelques semaines

Le cycle est bouclé lorsque l'hôte définitif consomme des tissus parasités de l'hôte intermédiaire, tels que les viscères infectés^{32,35}. C'est ici que la biologie rejoint la logistique : sans accès

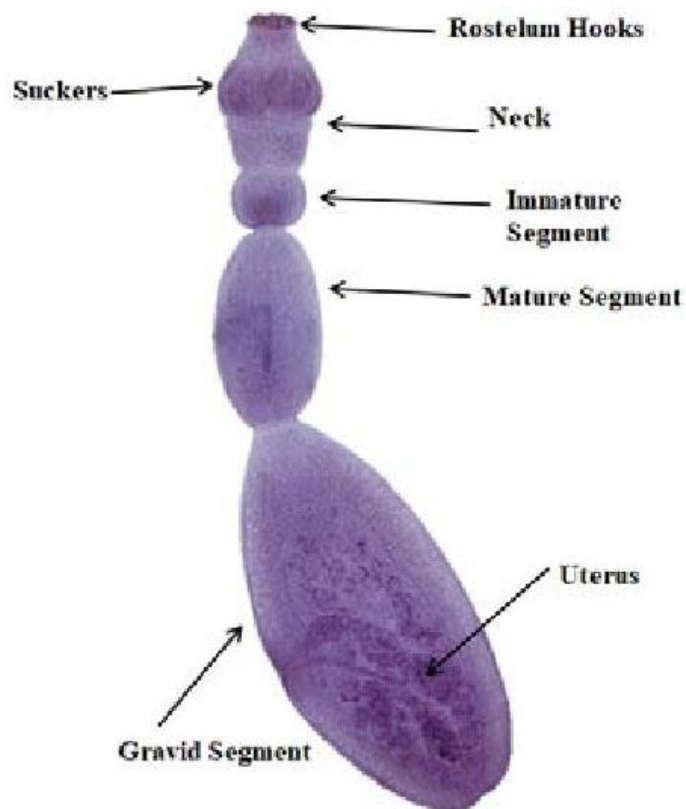


Figure 1.10 : Représentation du ver adulte d'*E. granulosus*³⁵.

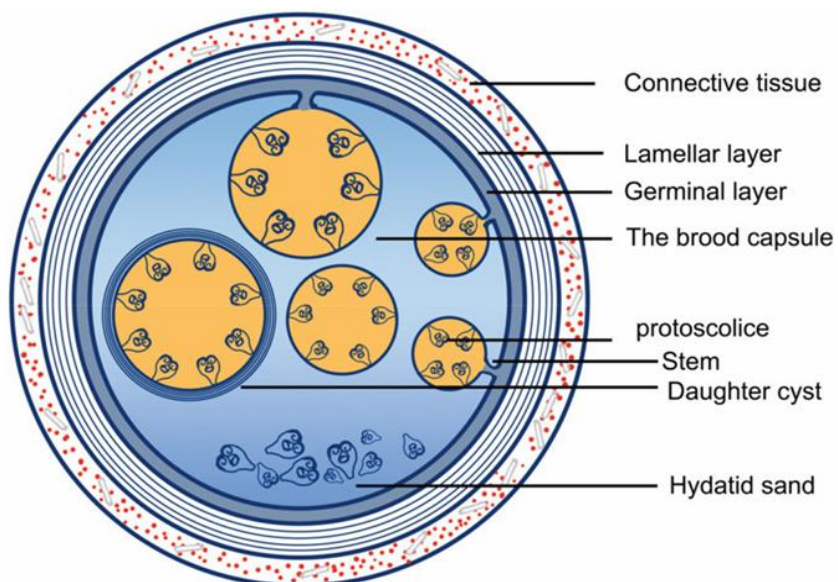


Figure 1.11 : Représentation du métacestode d'*E. granulosus*⁵¹.

des chiens aux abats (abattage non contrôlé), la boucle se rompt³⁵. Après ingestion, les protoscolex s'éversent (évagination), s'attachent à la muqueuse intestinale et se développent en vers adultes. La maturité de ces vers est atteinte en quelques semaines (32 à 80 jours ; **étape 6, figure 1.9**)^{32,35}.

1.7.4) Humain : hôte intermédiaire accidentel, impasse de transmission, pathologie

L'humain s'infecte par ingestion d'œufs par la voie féco-orale, par contact direct avec des hôtes définitifs excréteurs ou avec des matrices environnementales contaminées (le sol, les végétaux et l'eau)^{32,35}. Sur le plan biologique, il s'agit d'un hôte intermédiaire accidentel. La formation du kyste est possible, mais l'infection humaine ne favorise pas la continuité du cycle, ni même la transmission interhumaine^{32,35}.

1.8) Morphologie d'*E. granulosus*

1.8.1) Le ver adulte : un appareil reproducteur segmenté

Le ver adulte est un petit cestode de 3 à 7 mm (**figure 1.10**)³⁵. Il est organisé en trois unités morphologiques : scolex, cou court et strobile segmenté (**figure 1.10**)^{32,35,36,51}. Le scolex est globuleux et porte quatre ventouses ; il est armé d'un rostellum muni d'une double couronne de crochets. Cette armature assure l'ancrage à la muqueuse intestinale et stabilise le parasite dans un milieu mécanique (péristaltisme) où l'adhérence est très importante^{32,35,36}. Le strobile comporte 3 à 5 segments. Le dernier segment, gravidé, est disproportionné par rapport à l'ensemble : il concentre l'utérus rempli d'œufs^{32,35,36,51}.

1.8.2) Le métacestode : le kyste hydatique en tant qu'organe parasitaire

La forme larvaire, correspondant au métacestode (**figure 1.11**), se manifeste par la formation d'un kyste hydatique. Ce dernier est le plus souvent uniloculaire et globalement sphérique, mais sa taille et son retentissement dépendent du site d'implantation et du temps. Sur le plan structural, la paroi du kyste s'organise en trois couches qui matérialisent la frontière parasite-hôte^{36,51} :

- **Péricyste (l'adventice)** : couche d'origine hôte, fibreuse, dense et réactionnelle. Elle protège, cloisonne et participe à l'isolement immunitaire et tissulaire du parasite^{36,51}.
- **Membrane laminée (ectocyste)** : couche acellulaire, produite par le parasite, épaisse, constituée d'un complexe polysaccharide–protéine. Elle joue un rôle d'interface et de barrière ; son épaisseur peut augmenter avec l'âge du kyste^{36,51}.
- **Membrane germinative (endocyste)** : couche vivante, fine mais essentielle. Elle régule la perméabilité, assure la production du liquide hydatique^{35,36,51}.

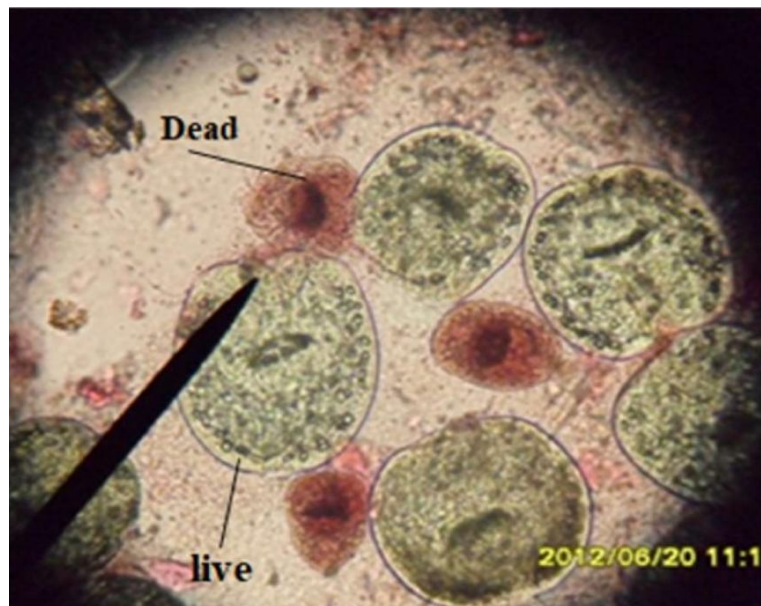


Figure 1.12 : Représentation des protoscolix viables et non viables de l'*E. granulosus*³⁶.

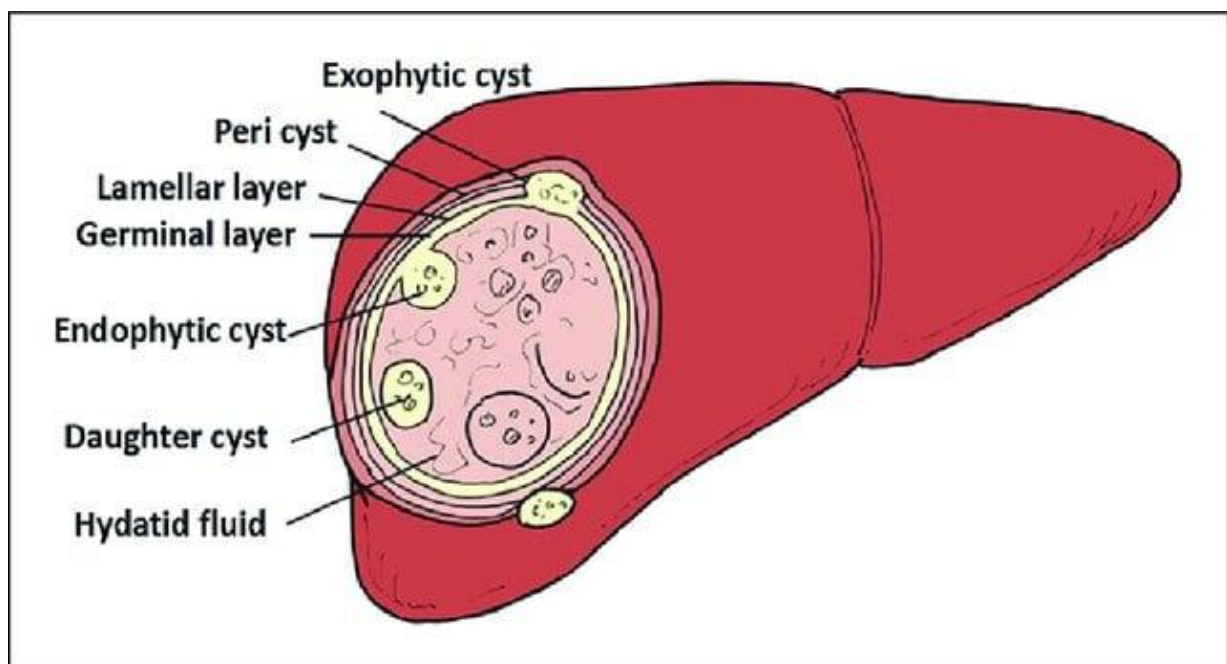


Figure 1.13 : Représentation de la physiopathologie de l'EK dans le foie³⁷.

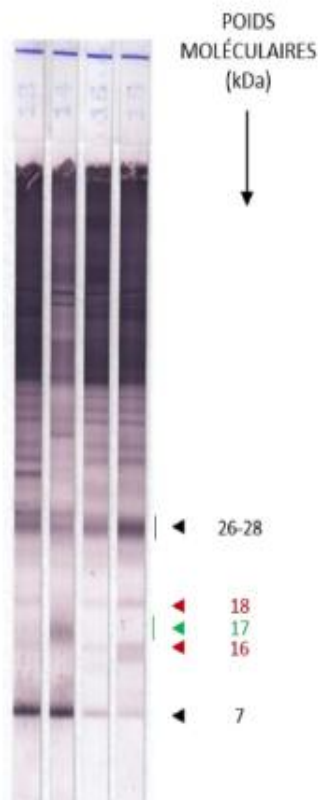
1.8.3) Capsules proligères, protoscolex et fertilité du kyste.

À partir de la membrane germinative se forment des capsules proligères (petites vésicules), véritables unités de reproduction asexuée, contenant des protoscolex invaginés^{32,35,36} (**figure 1.12**). Ces protoscolex n'ont de sens évolutif que s'ils sont ingérés par un hôte définitif chez lequel ils pourront s'attacher à la muqueuse intestinale et donner un ver adulte^{32,35,36}.

1.9) L'EK : de la physiopathologie à la prise en charge du patient

La pathogénie du kyste se manifeste principalement dans des organes tels que le foie (75,2 % des cas)⁴¹ et le poumon (22,4 % des cas)⁴¹. Des localisations plus rares peuvent également être observées, notamment dans la cavité abdominale et pelvienne (5,2 %), la rate (1,0 %), le rein (0,4 %), le cerveau (0,4 %), les os (0,2 %), les muscles et la peau (0,2 %), l'ovaire (0,06 %), ainsi que le médiastin et le cœur (0,03 %)⁴¹. D'autres localisations exceptionnelles ont également été décrites, telles que l'œil, la moelle épinière, le pancréas, la vessie ou le testicule, chacune représentant moins de 0,1 % des cas⁴¹. L'EK procède d'un double mouvement : croissance expansive du métacestode et réponse tissulaire de l'hôte^{32,37}. Dans le foie, la croissance du kyste comprime le parenchyme et déclenche une réaction inflammatoire qui aboutit à la formation d'un péricyste entièrement constitué de tissu hôte (**figure 1.13**)^{32,37}. Ce péricyste est fibreux, résistant et piège au sein de son épaisseur les structures biliaires. Les canaux biliaires, d'abord étirés à la surface du kyste en expansion, finissent parfois engainés par cette enveloppe puis peuvent fistuliser, ouvrant une communication kysto-biliaire^{32,37}. Sur le plan architectural, la rigidité fibreuse du péricyste empêche souvent l'effondrement complet de la cavité résiduelle après évacuation du contenu kystique^{32,37}. La formation du péricyste, caractéristique du foie et de la rate (1% des cas)⁴¹, n'est pas reproduite de la même façon dans d'autres tissus tels que le poumon ou le cerveau (0,4% des cas)^{37,41}. Une rupture kystique peut également survenir, avec un risque de dissémination du pathogène et un choc anaphylactique probable^{37,41}.

Ainsi, pour soigner un patient atteint de l'EK, une cystectomie totale ou partielle peut être réalisée (procédé chirurgical d'ablation totale ou partielle de l'adventice parasitaire), ou une hépalectomie partielle, ou un procédé PAIR (Ponction-Aspiration-Injection-Réaspiration) qui consiste en la réalisation d'une ponction dans le kyste afin d'aspirer le liquide hydatique et d'injecter une solution salée hypertonique pour tuer l'ensemble des parasites présents^{17,32,57}. Cependant, cette procédure PAIR comporte des risques, tels que la rupture du kyste, le choc anaphylactique ou l'hypernatrémie, liés à l'utilisation de la solution saline. Un traitement à l'albendazole, associé au praziquantel, est également prescrit au patient. La période d'incubation asymptomatique est également comprise entre 5 et 15 ans^{17,32,57}.



Diagnostic du genre *Echinococcus* sp.

→ Présence des bandes extrêmes à 7 et/ou 26-28 kDa

Diagnostic différentiel de l'espèce :

→ *Echinococcus multilocularis* : bandes à 16 et/ou 18 kDa

→ *Echinococcus granulosus* : bande à 17 kDa

Figure 1.14 : Représentation du test diagnostique WB au sein du LNRE⁴⁰.

1.10) Méthodes de diagnostic

1.10.1) Méthodes de sérologie

Au sein du Laboratoire National de Référence des échinococcoses (LNRE), rattaché au Service de Microbiologie Clinique du Centre Hospitalier Universitaire de Liège, le dépistage sérologique de l'EA repose sur un test immunoenzymatique de type ELISA (*Enzyme-Linked Immunosorbent Assay*)⁴⁰. C'est une méthode immunoenzymatique permettant de détecter et de quantifier des anticorps de type immunoglobulines G (IgG) dirigés contre les antigènes Em2 et Em18, spécifiques d'*E. multilocularis* révélés par un signal enzymatique colorimétrique⁴⁰. Les résultats sont exprimés sous forme d'index. Au LNRE, une zone grise volontairement élargie a été définie, de sorte qu'un index compris entre 0,6 et 1,1 est considéré comme douteux, tandis qu'un résultat supérieur ou égal à 1,1 est interprété comme positif. Malgré l'emploi d'antigènes reconnus pour leur bonne spécificité, des réactions croisées demeurent possibles, notamment avec *E. granulosus* ou d'autres ténias⁴⁰. Ce test pour l'EA est le test fabriqué par la firme Bordier Affinity Products, qui présente une sensibilité de 83 % et une spécificité de 98 %⁴⁴. Pour finir, la sérologie de dépistage ne peut être interprétée seule. Un test de confirmation par Western Blot (WB) est systématiquement réalisé en cas de résultat positif ou douteux.

Dans le cadre de l'EK au sein du LNRE, le dépistage sérologique est assuré par le test Hydatidose VirClia IgG (Vircell), une méthode immunologique permettant de détecter ou de quantifier les anticorps IgG anti-*E. granulosus* (antigène provenant du fluide hydatique) grâce à une réaction antigène–anticorps révélée par un signal lumineux de chimioluminescence⁴⁰. Le test VirClia présente une sensibilité de 100 % et une spécificité de 77 %, et les résultats sont exprimés sous forme d'index, avec une zone grise non concluante comprise entre 0,9 et 1,1. Le résultat est donc considéré comme positif si l'index est supérieur à 1,1⁴⁰. Là encore, tout résultat positif ou équivoque, ainsi que toute situation dans laquelle le contexte clinico-radiologique demeure fortement évocateur malgré une négativité sérologique, doit conduire à la réalisation d'un Western Blot de confirmation⁴⁰.

Le test WB occupe ainsi une place centrale dans l'algorithme diagnostique. Il s'agit d'un test qualitatif fondé sur la détection d'IgG dirigées contre des antigènes spécifiques d'*E. granulosus*, d'*E. multilocularis*, ainsi que contre des antigènes communs au genre *Echinococcus*. Cet examen permet de renforcer la fiabilité diagnostique, notamment dans les situations de discordance entre la sérologie de dépistage et les arguments cliniques ou radiologiques, car il présente une sensibilité et une spécificité de 100 % (**figure 1.14**)⁴⁰.

WHO classification	Stage	Features
CL	Active	Unilocular cystic lesion without a well-defined wall (double wall sign is not seen); no internal debris or septations (without evident signs of non-parasitic aetiology); seen during the initial phase of hydatid development; requires diagnostic tests to differentiate from other cystic lesions
CE1	Active	Unilocular cystic lesion with a well-defined wall; fine internal echoes/debris may be seen representing hydatid sand (pathognomonic): "Snowflake sign"; "double wall sign" (pathognomonic): Two echogenic lines seen representing the parasite larval cyst and the host-derived fibrous cyst capsule, separated by minimal hypoechoic fluid. However, may not always be seen, and care should be taken not to dismiss a CE1 cyst on the basis of not seeing this sign; internal septae/membranes not seen
CE2	Active	Multivesicular cystic structure with daughter vesicles; rosette/Honey-comb pattern: Daughter cysts fill the entire mother cyst; spoke-wheel pattern: Daughter vesicles arranged peripherally with minimal central high-density matrix
CE3a	Transitional	Detachment of endocyst from pericyst: Can be due to decreasing intracystic pressure, degeneration, host response, trauma, or response to therapy; Produces serpentine linear floating membranes: "Water-lily sign" (highly specific sign); detachment may be partial or complete
CE3b	Transitional	Few daughter cysts in a predominantly avascular solid matrix: "Swiss cheese" because of significant solid component it can mimic malignancy
CE4	Inactive	Heteroechoic solid avascular mass with no daughter cysts: Pseudotumour/ball-of-wool/canicular/cerebroid appearance; due to solid nature: Mimics malignancy (requires contrast-enhanced CT/MRI: Showing lack of enhancement to differentiate from malignancy)
CE5	Inactive	Thick, densely calcified walls (complete or near complete); partial wall calcification does not always indicate parasite death; however, densely calcified cysts may be assumed to be inactive; wall calcification is most extensively found in, but not limited to CE4 and CE5; heterogeneous avascular solid content (when acoustic shadow allows visualization)

WHO: World health organization; CL: Cystic lesions; CE: Cystic echinococcosis; CT: Computed tomography; MRI: Magnetic resonance imaging.

Figure 1.15 : Tableau précisant la classification de l'EK selon l'OMS et les différentes technologies d'imagerie médicale¹⁷.

1.10.2) Biologie moléculaire

Les méthodes moléculaires constituent un complément diagnostique essentiel, en particulier lorsqu'un prélèvement tissulaire ou liquidien est disponible. Elles peuvent être appliquées à des biopsies ainsi qu'à divers liquides biologiques, notamment la bile, le liquide pleural, les expectorations et le lavage broncho-alvéolaire, mais ne sont pas validées sur le sang. Le LNRE dispose actuellement de deux approches de *Polymerase Chain Reaction (PCR)* complémentaires sur le plan de la sensibilité⁴⁰. En première intention, une PCR multiplexe en temps réel, développée et validée en interne, est privilégiée pour sa rapidité d'exécution. Cette méthode permet la détection simultanée d'*E. multilocularis* et d'*E. granulosus* s.s. (sous-ensemble du complexe *E. granulosus* s.l. le plus fréquemment retrouvé)⁴⁰. En cas de résultat négatif, une PCR multiplex classique est réalisée en seconde intention^{40,45}. Cette méthode permet de détecter le complexe *E. granulosus* s.l. par amplification d'une région du gène mitochondrial codant la petite sous-unité de l'ARN ribosomal mitochondrial (*rrnS*), tandis qu'*E. multilocularis* est ciblé séparément via le gène mitochondrial codant la sous-unité 1 de la NADH déshydrogénase (*nad1*)^{40,45}. Néanmoins, lorsque l'identification précise du génotype d'*E. granulosus* s.l. est requise, un séquençage post-amplification de Sanger demeure nécessaire⁴⁰.

Plus récemment, des recherches prometteuses portent sur le développement d'une PCR quantitative fondée sur les microARN (petits ARN non codants, généralement d'environ 20–25 nucléotides, qui participent à la régulation de l'expression des gènes en inhibant la traduction ou en favorisant la dégradation des ARN messagers cibles), présents dans le sang³⁹. Cette technique vise à renouveler la stratégie de diagnostic et de suivi chez un patient infecté par l'*E. multilocularis*³⁹. Cette approche, ciblant des microARN relargués par le foie, détecte dans le sérum des signatures associées à l'infection par *E. multilocularis* (marqueurs de dégradation du foie caractéristiques de l'EA). Cette méthode non invasive, fondée sur la quantification des microARN hépato-spécifiques par PCR quantitative (qPCR), ouvre la voie à un diagnostic précoce et potentiellement innovant pour le suivi thérapeutique et l'assignation des stades de progression de la maladie³⁹.

1.10.3) Imagerie Médicale

Les méthodes d'imagerie occupent une place essentielle dans le diagnostic, la classification et le suivi des échinococcoses^{16,17,32,38}. L'échographie abdominale, la tomodensitométrie (CT-Scan), l'imagerie par résonance magnétique (IRM) et la tomographie par émission de positrons au fluorodésoxyglucose (PET-Scan) permettent de détecter les lésions, d'en préciser la taille, la localisation, l'extension et les rapports anatomiques, notamment en vue d'une prise en charge chirurgicale^{16,17,32,38}. Dans l'EK, l'échographie (**figure 1.15**) constitue l'examen de référence et permet de classer les lésions selon les classifications de l'OMS (Organisation Mondiale

de la Santé), fondées sur leur aspect morphologique^{17,32,38}. La tomодensitométrie et l'IRM complètent cette analyse en offrant une meilleure caractérisation anatomique, notamment dans les localisations complexes^{17,32,38}. Dans l'EA, l'IRM et le CT-Scan sont également indispensables à l'évaluation de l'extension lésionnelle et au classement selon la classification PNM de l'OMS qui prend en compte la localisation parasitaire (P), l'envahissement des structures voisines (N) et la présence éventuelle de métastases (M) (**figure 1.16**)^{16,32,38}. Enfin, le PET-Scan présente un intérêt particulier dans le suivi de l'échinococcose alvéolaire, en permettant d'apprécier l'activité métabolique des tissus inflammatoires péri-parasitaires et, par conséquent, l'activité évolutive des lésions après traitement, à la suite de l'injection d'un traceur radioactif 3 heures avant la procédure^{32,38,56}. L'échographie peut également contribuer au diagnostic et au suivi clinique d'un patient.

PNM classification of alveolar echinococcosis

P	Hepatic localisation of the <i>Parasite</i>
P X	Primary tumor cannot be assessed
P 0	No detectable tumor in the liver
P 1	Peripheral lesions without proximal vascular and/or biliar involvement
P 2	Central lesions with proximal vascular and/or biliar involvement of one lobe ^a
P 3	Central lesions with hilar vascular or biliar involvement of both lobes and/ or with involvement of two hepatic veins
P 4	Any liver lesion with extension along the vessels ^b and the biliary tree
N	Extra hepatic involvement of neighbouring organs [diaphragm, lung, pleura, pericardium, heart, gastric and duodenal wall, adrenal glands, peritoneum, retroperitoneum, parietal wall (muscles, skin, bone), pancreas, regional lymph nodes, liver ligaments, kidney]
N X	Not evaluable
N 0	No regional involvement
N 1	Regional involvement of contiguous organs or tissues
M	The absence or presence of distant <i>Metastasis</i> [lung, distant lymph nodes, spleen, CNS, orbital, bone, skin, muscle, kidney, distant peritoneum and retroperitoneum]
M X	Not completely evaluated
M 0	No metastasis ^c
M 1	Metastasis

^a For classification, the plane projecting between the bed of the gall bladder and the inferior vena cava divides the liver in two lobes.

^b Vessels mean inferior vena cava, portal vein and arteries.

^c Chest X-ray and cerebral CT negative.

Figure 1.16 : Tableau représentant la classification de l'EA, selon la classification du PNM¹⁶.

1.11) Typage moléculaire

1.11.1) Séquençage de première génération

Les méthodes de séquençage de première génération (**figure 1.17**) reposaient sur le clivage chimique de l'ADN ou sur l'utilisation de nucléotides terminateurs de chaîne, suivis d'une séparation manuelle des fragments par électrophorèse^{42,43,52}. La méthode de Sanger s'est rapidement imposée comme la référence, en raison de sa robustesse et de sa simplicité d'interprétation^{42,43,52}. Son principe repose sur l'incorporation aléatoire de didésoxynucléotides (ddNTP), dépourvus de groupement hydroxyle en 3', ce qui interrompt l'élongation de la chaîne naissante lors de leur incorporation^{42,43}. La lecture de la séquence est ensuite déduite à partir de la taille des fragments obtenus sur un gel^{42,43}. Cette méthode a d'abord été réalisée avec des marqueurs radioactifs, puis améliorée grâce à l'introduction de fluorophores, ce qui a permis une détection automatisée plus fiable^{42,43}. Malgré son importance historique, le séquençage de Sanger reste une méthode à faible débit, avec des lectures généralement inférieures à 1 kb, ce qui le rend peu adapté au séquençage rapide de grands génomes^{42,43,52}.

1.11.2) Séquençage de deuxième génération

L'apparition du séquençage de deuxième génération (**figure 1.17**), également appelé *Next-Generation Sequencing* (NGS), a profondément modifié ce paysage en introduisant le principe du séquençage à haut débit^{42,43,52}. À partir du milieu des années 2000, des plateformes comme Roche, puis Illumina et Thermo Fisher, ont rendu possible la production simultanée de millions, voire de milliards de lectures au cours d'un seul run de séquençage^{42,43}. Le changement conceptuel majeur est le passage d'un séquençage fragment par fragment à un séquençage de haut débit. Cette évolution a considérablement réduit le coût et le temps de séquençage de grands génomes^{42,43}.

Parmi les technologies de deuxième génération, la firme Illumina s'est imposée comme le leader du marché^{42,43}. Le principe de séquençage proposé par cette entreprise (**figure 1.18**) repose sur une amplification en pont sur *Flowcell* et sur une chimie de séquençage par synthèse utilisant des terminateurs fluorescents réversibles. Les fragments d'ADN (ou *reads*), munis d'adaptateurs, s'hybrident avec des oligonucléotides fixés à la surface de la *Flowcell*, puis forment des clusters après plusieurs cycles d'amplification^{42,43,52}.

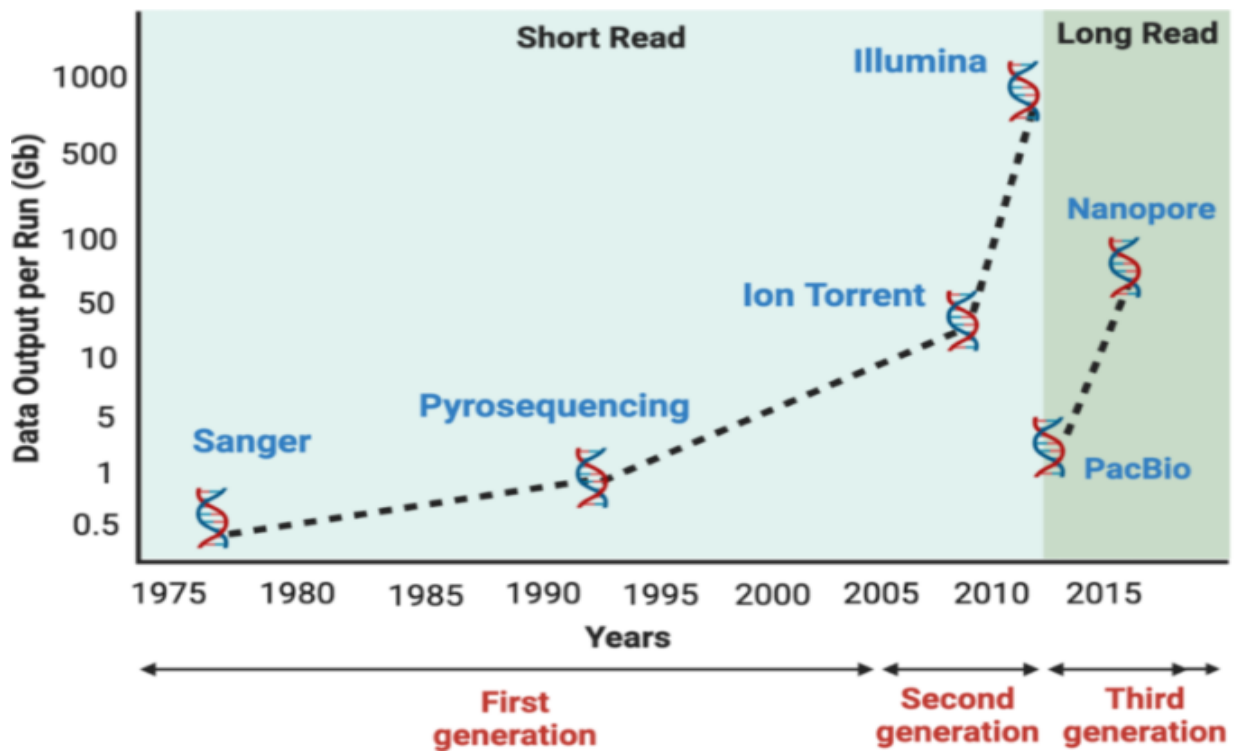


Figure 1.17 : Évolution des technologies de séquençage⁴³.

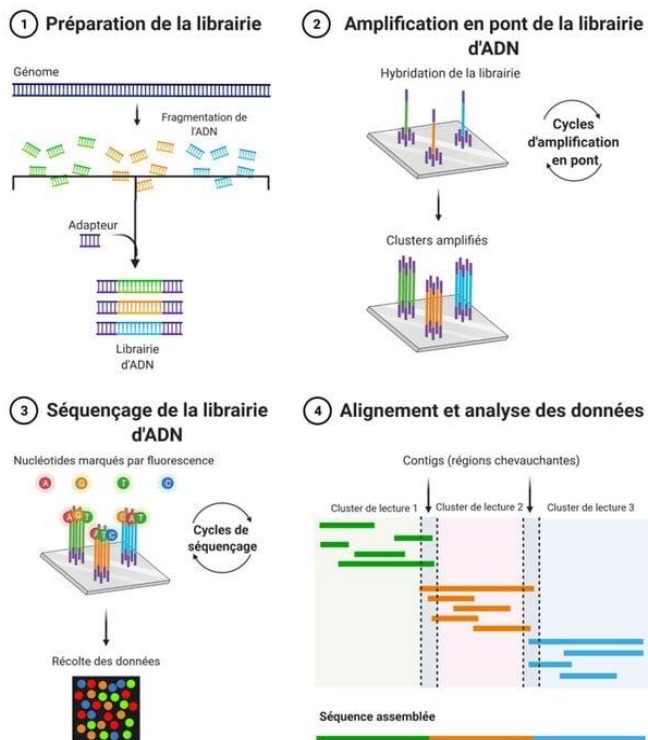


Figure 1.18 : Principe du séquençage Illumina⁵².

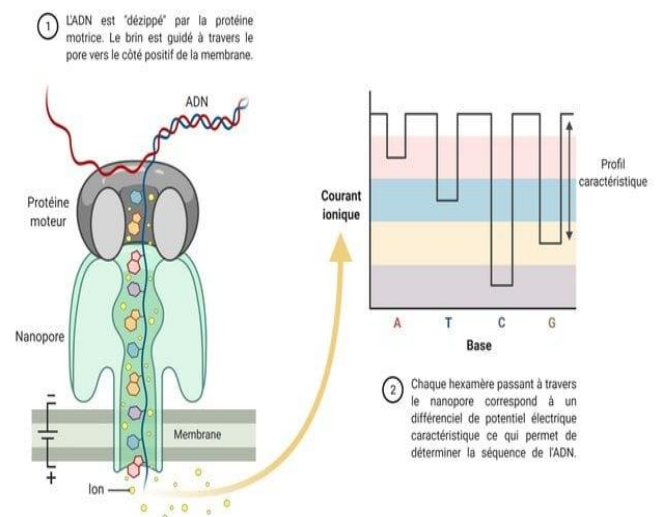


Figure 1.19 : Principe du séquençage Nanopore⁵².

Lors de la phase de lecture, un seul nucléotide modifié est incorporé à chaque cycle ; il est identifié par la fluorescence émise. Après cela, le fluorophore et le groupement bloquant en 3' sont retirés afin de permettre le cycle suivant^{42,43,52}. Cette approche offre une très grande capacité de production de données, avec des millions de lectures en quelques heures et à un coût par base nettement inférieur à celui du séquençage de première génération^{42,43}. En contrepartie, les lectures restent relativement courtes, ce qui complique l'assemblage *de novo* et l'analyse des régions répétées ou des variants structuraux complexes^{42,43}.

1.11.3) Séquençage de troisième génération

Le séquençage de troisième génération (**figure 1.17**), ou *single-molecule sequencing*, marque un nouveau changement de paradigme^{42,43}. Contrairement aux approches précédentes, il ne nécessite pas d'amplification préalable des fragments sur la *Flowcell*, ce qui réduit les biais introduits par l'amplification et permet d'analyser de plus faibles quantités d'ADN^{42,43}. Il présente aussi l'avantage de mieux préserver les signatures moléculaires intrinsèques, notamment certaines modifications de bases comme la méthylation de l'ADN^{42,43}. Son autre atout majeur réside dans la longueur des lectures, très supérieure à celle des technologies de séquençage de deuxième génération, ce qui facilite l'assemblage génomique, la résolution des régions répétées et l'identification des variants structuraux^{42,43,52}. En revanche, ces méthodes restent historiquement associées à des taux d'erreur bruts plus élevés (surtout avec la technologie Nanopore)^{42,43}.

Les deux grandes approches de troisième génération les plus connues restent le séquençage de Pacific Biosciences (PacBio) et le séquençage d'Oxford Nanopore Technologies (ONT)^{42,43}. La technologie *Single-Molecule, Real-Time* (SMRT) repose sur l'observation en temps réel de l'incorporation de nucléotides fluorescents dans des nanocavités appelées *zero-mode waveguides*, où une ADN polymérase (Phi29 modifiée par PacBio) synthétise un brin complémentaire^{42,43,52}. Le séquençage Nanopore (**figure 1.19**), quant à lui, repose sur un principe totalement différent : une molécule d'ADN simple brin est entraînée à travers un nanopore membranaire par une protéine motrice, et chaque hexamère (enchaînement de 6 nucléotides) perturbe le courant ionique selon un profil caractéristique, ce qui permet d'inférer la séquence^{42,43,52}. Cette technologie se distingue par sa très grande longueur de lecture, pouvant atteindre plusieurs dizaines, voire des centaines de kilobases selon les conditions, ce qui en fait un outil particulièrement puissant pour l'étude des génomes complexes, riches en séquences répétitives^{42,43,52}.

1.12) Évolution du typage moléculaire d'*E. multilocularis*

Les premiers travaux, menés au début des années 1990, s'appuyaient sur trois approches : la PCR, le séquençage de Sanger et la comparaison de gènes mitochondriaux à l'aide d'arbres phylogénétiques pour clarifier les relations phylogénétiques^{20,21,22}. Ainsi, la première diversité génétique mesurée repose sur des génomes mitochondriaux incomplets. Le premier séquençage porte sur 366 pb (sur 1 608 pb) de la sous-unité 1 du gène Cytochrome C Oxydase (cox1) et le deuxième porte sur 471 pb (sur 894 pb) de la sous-unité 1 du gène NADH Déshydrogénase (nad1)^{13,21,22}.

L'obstacle majeur n'a pourtant pas été seulement la longueur des fragments, mais bien la comparaison des résultats entre les études. Des travaux antérieurs n'utilisent pas de gènes complets et combinent divers marqueurs. Même lorsque cox1 ou nad1 sont réutilisés, les longueurs divergent d'un article à l'autre^{20,21,22,23}. À cela s'ajoute une nomenclature inconstante des génotypes/haplotypes²². Dans ces conditions, l'évaluation globale de la diversité devient fragile, parce que les jeux de données mitochondriaux ne s'additionnent pas proprement²².

Une seconde trajectoire s'est structurée à partir de marqueurs nucléaires, principalement des séquences microsatellitaires, avec un marqueur devenu central : *Echinococcus multilocularis* satellite B (EmsB)^{22,24,25,26,27}. Ce dernier est localisé sur le chromosome 5 et présente environ 40 copies en tandem. Le typage EmsB repose sur l'amplification par PCR du marqueur EmsB. Les fragments amplifiés, de tailles variables, sont ensuite séparés par électrophorèse capillaire, générant un profil de pics caractérisé par la taille des fragments et leur intensité de fluorescence. Ce profil moléculaire est utilisé comme empreinte génétique pour comparer les isolats entre eux à partir de matrices de similarité ou d'analyses de *clustering*^{22,24,25,26,27}. Ses avantages sont constamment décrits (fort pouvoir discriminant, applicable à divers types d'échantillons), mais ses limites sont tout de même importantes (interprétation difficile et sous-estimation de la diversité)^{22,24,25,26,27}. Il devient également clair que des échantillons multirégionaux sont nécessaires pour construire des dendrogrammes robustes, puisque le marqueur et la méthode statistique influencent la structure des clusters en fonction des variations propres aux échantillons inclus, ce qui a nourri des critiques et suscité un appel à des approches plus fondées sur l'analyse de séquences d'ADN²².

C'est précisément là qu'intervient le séquençage complet du génome mitochondrial, en tant que changement d'échelle majeur dans l'étude génétique d'*E. multilocularis* car on passe de l'analyse de fragments à celle de gènes entiers, puis à l'exploration du génome mitochondrial complet (mitogénome)^{22,28,29,30}. Historiquement, les études ont évolué de fragments isolés vers des séquences de plus en plus longues, jusqu'à l'acquisition récente du mitogénome complet.^{28,29,30} La méthode principale de séquençage repose essentiellement sur la méthode NGS

d'Illumina, avec pour unité de comparaison les *Single Nucleotide Polymorphisms* (SNPs)²². Le mitogénome complet d'*E. multilocularis* d'une longueur de 13 738 pb est présenté comme stratégiquement utile pour réancrer des marqueurs plus anciens et combiner des données nouvelles et historiques dans un même espace comparatif^{22,28,29,30}.

À cette résolution, des résultats deviennent concrets et cumulables. Les analyses fondées sur les mitogénomes complets confirment l'existence de trois clades répartis sur trois continents (Europe, Asie et Amérique du Nord) et montrent que la diversité est plus élevée qu'on ne le supposait quelques années auparavant^{22,28,29,30}. Des travaux récents utilisent une exploration mitochondriale complète d'*E. multilocularis* chez des patients français (13 735–13 740 pb)^{22,28,29,30}. Ils décrivent 58 haplotypes regroupés en quatre haplogroupes (HG1–HG4) (**figure 1.20**), avec une subdivision de HG3 en micro-haplogroupes et une assignation géographique (HG1 rapporté notamment en Alaska, St. Lawrence Island, Yakoutie, Svalbard ; HG2 observé en Asie, Amérique du Nord, Europe ; HG3 en Europe (HG3a transatlantique) et HG4 réduit à un isolat d'Olkhon Island)^{22,28,29,30}.

Ainsi, la variabilité du mitogénome de l'*E. multilocularis* et de l'*E. granulosus* reste encore assez méconnue en Belgique. C'est pourquoi un travail visant à caractériser la complexité génétique de ces deux espèces parasitaires s'impose afin de cartographier la distribution haplotypique de notre population et de la comparer à la littérature.

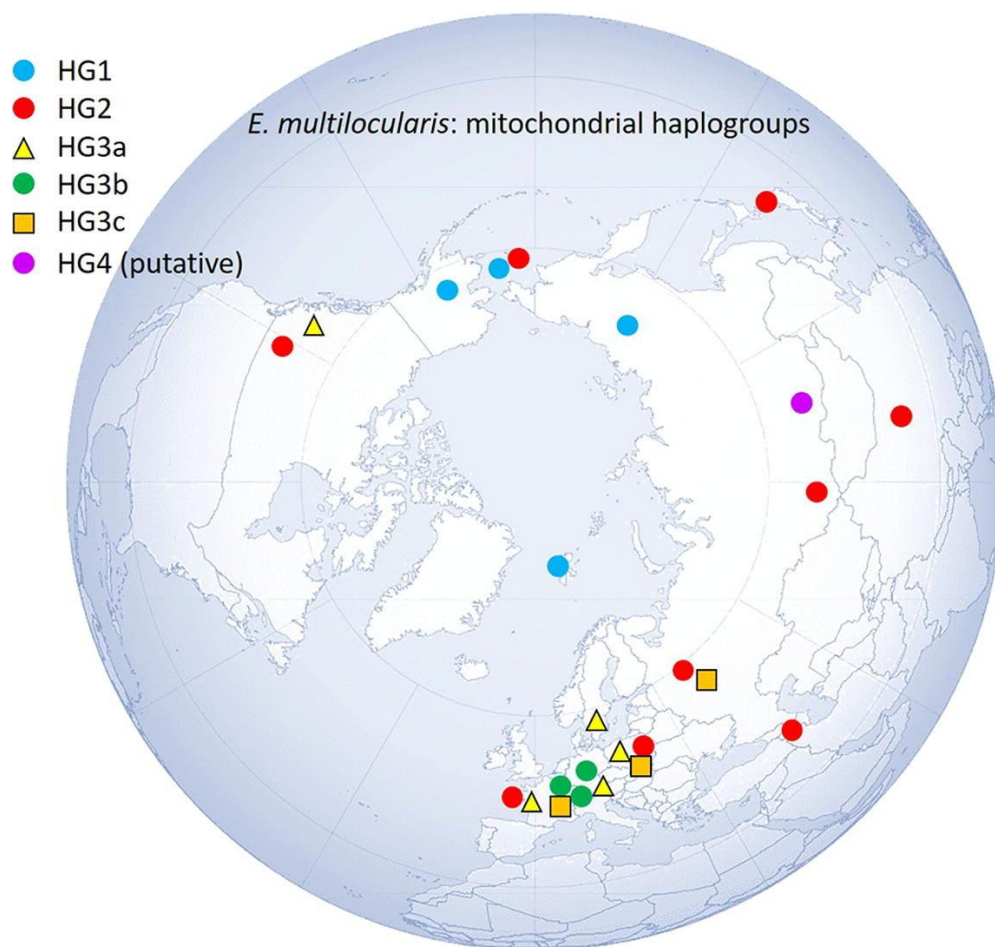


Figure 1.20 : Cartographie mondiale des différents haplogroupes majeurs d'*E. multilocularis*²⁹.

Chapitre II : Objectifs



CHAPITRE II : OBJECTIFS

L'objectif principal de ce mémoire est de développer et d'évaluer de nouvelles approches de séquençage applicables à la caractérisation moléculaire d'*E. granulosus* et d'*E. multilocularis*, en remplacement ou en complément des méthodes actuellement utilisées au laboratoire. Jusqu'à présent, la caractérisation reposait principalement sur le typage du microsatellite EmsB par analyse de fragments en électrophorèse capillaire. Cette méthode génère un profil de pics propre à chaque isolat, mais n'en fournit qu'une image partielle de la diversité génétique : faute de séquences nucléotidiques, elle ne permet aucune reconstruction phylogénétique.

Le présent travail vise donc à mettre en place des stratégies fondées sur le séquençage du mitogénome complet, afin d'obtenir une résolution plus fine des lignées mitochondriales circulant chez les patients pris en charge en Belgique.

Pour atteindre l'objectif, la première étape du travail consiste à développer un *workflow* de séquençage du mitogénome complet à l'aide de la technologie Illumina sur l'appareil iSeq100. Cette approche repose sur l'amplification de plusieurs fragments chevauchants par PCR, suivie d'un séquençage par NGS. Ces fragments chevauchants couvrent l'ensemble du génome mitochondrial. L'enjeu consiste à vérifier la faisabilité de cette stratégie sur des échantillons cliniques, à évaluer la qualité des séquences obtenues et à déterminer si cette approche permet d'assigner les isolats à des *clades* ou à des haplogroupes mitochondriaux. Un deuxième objectif est d'implémenter au laboratoire une approche basée sur la technologie MinION Mk1D d'Oxford Nanopore et de l'appliquer au séquençage mitochondrial. Ce volet permettra d'évaluer les performances de cette technologie en termes de qualité de lecture, de profondeur de séquençage, de robustesse analytique et de capacité à produire des consensus exploitables pour l'analyse phylogénétique. Une attention particulière sera portée à son intérêt potentiel pour la résolution de régions répétitives, afin de déterminer si les lectures longues constituent un avantage réel pour l'analyse du mitogénome. Le troisième objectif vise à construire 2 pipelines bioinformatiques (pour le séquençage Illumina et Nanopore) afin de valider un consensus fiable à partir des fichiers FASTQ générés.

Ces trois objectifs permettront de comparer les approches Illumina et Oxford Nanopore sur les plans analytique, technique et économique. Cette comparaison portera notamment sur la qualité des données produites, la profondeur obtenue, la facilité de mise en œuvre au laboratoire, ainsi que la pertinence de chaque technologie pour une utilisation future en recherche.



Chapitre III : Matériel et Méthodes



CHAPITRE III : MATÉRIEL ET MÉTHODES

3.1) Table des échantillons (annexe 1)

Cette étude a été réalisée sur une cohorte de 20 échantillons (ID1 – ID20) issus de 16 patients pris en charge au CHU de Liège. Au sein de notre service de microbiologie clinique, tous ces prélèvements ont été confirmés positifs par PCR pour *E. granulosus* (4 échantillons issus de 4 patients) ou pour *E. multilocularis* (16 échantillons issus de 12 patients). De plus, 6 extractions d'ADN, issues de prélèvements cliniques (ID21 – ID26) de 6 patients distincts, nous ont été transmises par le Centre National de Référence pour les échinococcoses (CNR-E) du CHU de Besançon (France). Les échantillons sont répartis équitablement entre *E. granulosus* (ID24, ID25 et ID26) et *E. multilocularis* (ID21, ID22 et ID23). Ainsi, la cohorte totale de l'étude comprend 26 échantillons issus de 22 patients. Pour plus d'informations, veuillez consulter l'**annexe 1**.

3.2) Matériel et équipements communs

Ci-dessous sont listés les consommables, réactifs et équipements utilisés pour l'ensemble des manipulations

3.2.1) Consommables :

- Embouts de pipette 20 µL, 200 µL et 1000 µL (tips) : (Rainin) Bio-Clean Ultra Pipette Tips (racks, 960 embouts).
- Tubes 2 mL (vissés) : Screw-Top Tubes with O-Ring Cap (Sigma-Aldrich), 2 mL, autoportants, en polypropylène naturel.
- Tubes de 2 mL (alternatifs) : Sarstedt, 2 mL, autoportants, polypropylène naturel, boîte de 500.
- Microtubes ADN à faible adsorption : Eppendorf DNA LoBind Tubes, 2.0 mL, PCR clean, résistance jusqu'à 25 000 × g, lot de 250.

3.2.2) Réactifs :

- Eau pour biologie moléculaire : Eau pour biologie moléculaire (Applichem), flacon de 500 mL (Labelians), réf. A73980500.

3.2.3) Équipements :

- Vortex: Heidolph 036130000 Vibrating Test Tube Shaker / Vortex Mixer.
- Microcentrifugeuse : Eppendorf Centrifuge 5425 R.
- Centrifugeuse de table : Benchmark Scientific BSC1006-B-E.

- Pipettes P2, P10, P20, P100, P200, P1000 (uni ou multicanaux) : Rainin Pipet-Lite Single- or Multiple-Channel Pipette.

3.3) Extraction de l'ADN

3.3.1) Objectifs :

L'objectif de cette technique est d'extraire rapidement de l'ADN à partir d'un fragment de biopsie hépatique (tissu calcifié ou nécrosé) ou d'un prélèvement liquide, tout en limitant l'exposition de l'ADN aux DNases présentes dans l'échantillon.

3.3.2) Principe :

Le principe de cette technique consiste en une lyse mécanique à l'aide de billes en céramique, suivie d'une lyse enzymatique par la protéinase K et d'une purification sur une colonne de silice.

3.3.3) Conditions de sécurité / prévention des contaminations :

- Réaliser l'ensemble des étapes sous hotte à flux laminaire.
- Porter les équipements de protection individuelle (EPI) adaptés (gants, tablier) et utiliser du matériel stérile ou à usage unique.

3.3.4) Matériel, réactifs et équipements :

- Kit d'extraction d'ADN : High Pure PCR Template Preparation Kit (Roche, Bâle, Suisse), version 28 (11796828001).
- Homogénéisateur à billes : Bead Ruptor 12 (Omni Homogenizers).
- Thermomixeur : Eppendorf ThermoMixer Comfort + ThermoBlock 2 mL
- Isopropanol : 2-propanol $\geq 99,0$ %, conforme USP (Sigma-Aldrich).
- Tubes avec billes de céramique : Hard Tissue Homogenizing Mix, 2.8 mm Ceramic (tubes renforcés de 2 mL, 50/paquet) (Revvity).
- Spectrophotomètre : Thermo Scientific NanoDrop One/OneC (UV-visible à microvolume, Wi-Fi).
- Centrifugeuse compatible microtubes 2 mL et colonnes (RCF réglable).

3.3.5) Procédure :

3.3.5.1) Préparation de l'échantillon :

- Placer dans un tube contenant des billes en céramique un fragment de biopsie (quelques mm de diamètre) ou 200 μ L de prélèvement liquide.
- Ajouter 300 μ L de tampon de lyse.
- Vortexer brièvement pour homogénéiser.

3.3.5.2) Homogénéisation mécanique :

- Placer le tube dans le Bead Ruptor réglé à $6 \text{ m} \cdot \text{s}^{-1}$ pendant une minute.
- Centrifuger 1 min à vitesse maximale, ou jusqu'à disparition complète de la mousse.

3.3.5.3) Lyse enzymatique :

- Transférer la totalité du volume dans un tube conique standard de 2 mL (sans billes).
- Ajouter 40 µL de protéinase K.
- Incuber 10–15 min à 56 °C dans un thermomixeur.

3.3.5.4) Fixation sur colonne de silice :

- Ajouter 200 µL de Binding Buffer, puis vortexer.
- Incuber 10 min à 70 °C.
- Pendant l'incubation, préchauffer aussi le tampon d'élution à 70 °C.
- Assembler la colonne de silice et le tube collecteur fournis par la firme Roche.
- Déposer l'intégralité du mélange sur la colonne.
- Centrifuger 1 min à 8 000 × g (RCF).
- Placer la colonne de silice dans un nouveau tube collecteur.

3.3.5.5) Lavages :

- Ajouter 500 µL de Removal Buffer sur la membrane.
- Centrifuger 1 min à 8 000 × g (RCF).
- Ajouter 500 µL de Wash Buffer. Centrifuger 1 min à 8 000 × g (RCF).

Répéter l'étape précédente une seconde fois (2 lavages au Wash Buffer au total).

- Centrifuger pendant 10 s à vitesse maximale (étape de séchage de la membrane).

3.3.5.6) Élution :

- Ajouter 200 µL de tampon d'élution préchauffé directement sur la membrane.
- Laisser en contact pendant 3 min à température ambiante.
- Centrifuger 1 min à 8 000 × g (RCF).
- Récupérer l'ADN élué dans un tube Eppendorf stérile.

3.3.5.7) Quantifier :

Quantifier et mesurer la qualité de l'ADN à partir de 2 µL d'ADN extrait à l'aide d'un spectrophotomètre (Thermo Scientific™ NanoDrop™ One/One^C).

3.4) Protocole de conception des amorces PCR

Ce protocole de PCR a été développé afin de concevoir des amorces pour amplifier le mitogénome de *I.E. granulosus* s.s. et de *I.E. multilocularis*. Plusieurs amplicons sont générés pour couvrir l'intégralité du mitogénome.

3.4.1) Conception des amorces :

Plusieurs outils ont été utilisés seuls ou en combinaison pour concevoir des amorces :

- Clone Manager version 12
- FastPCR version 7.0.05

3.4.2) Vérification sur génome de référence et contre l'ADN humain :

La vérification des primers se fait sur un génome de référence (*E. multilocularis* : AB018440.2 et *E. granulosus* : NC_044548.1) des parasites étudiés et sur l'ADN humain, afin de :

- Confirmer l'alignement attendu sur la séquence cible et limiter le risque d'amplifications non spécifiques.

3.4.3) Outils :

- NCBI Primer-BLAST (Primer designing tool — NCBI/NIH) (version 2026)
- BLAST (Basic Local Alignment Search Tool) (version 2026)

3.4.4) Analyse des dimères et contrôle qualité des amplicons :

Vérifier la formation potentielle de homodimères et hétérodimères, ainsi que la qualité théorique des amplicons par l'utilisation des outils suivants :

- PrimerQuest (PCR & qPCR Primer Design Tool) — IDT (version 2026)
- Multiple Primer Analyzer — Thermo Fisher Scientific (version 2026)

3.4.5) Visualisation des alignements et simulation in silico :

Visualiser l'alignement des amorces sur un génome de référence et simuler un rendu PCR (ex. taille attendue sur gel d'agarose 1 ou 2 %) par l'utilisation de l'outil SnapGene 8.0.3

3.4.6) Étape optionnelle :

Réaliser un alignement multiple avec l'algorithme MAFFT (v7.526) sur plusieurs séquences du mitogénome et choisir ses amorces manuellement en regardant les régions conservées sur Aliview version 1.31

3.5) Protocole de manipulation PCR

Ce protocole de PCR classique a été développé afin d'amplifier les régions ciblées par nos réactions, tout en tenant compte des contraintes de qualité et de quantité de l'ADN obtenu après extraction.

3.5.1) Matériel, réactifs et consommables :

3.5.1.1) Amorces (primers) (annexes 3 et 4).

Les séquences des amorces sont disponibles dans les annexes 3 et 4. Celles-ci ont toutes été commandées auprès d'Integrated DNA Technologies (IDT) (Belgique). Les séquences pour amplifier le mitogénome de l'*E. multilocularis* et *E. granulosus* issus de la thèse de Séverine Lallemand (référence bibliographique 13) ont été utilisés (**Annexe 4**). De plus, d'autres amorces pour amplifier le mitogénome de l'*E. granulosus* ont été conçus et testés (**annexe 3**).

3.5.1.2) Réactifs PCR :

- Polymérase / Master Mix: Applied Biosystems™ AmpliTaq Gold™ 360 Master Mix (reference: 15594324)
- ADN matrice (template DNA)
- Eau PCR : Mini Plasco Aqua ad iniectabilia (B. Braun)

3.5.1.3) Thermocycleur :

- Thermocycleur : Applied Biosystems™ SimpliAmp™ Thermal Cycler

3.5.1.4) Électrophorèse / révélation :

- Gels: Invitrogen™ E-Gel™ Agarose 2% Gel avec SYBR™ Safe DNA Gel Stain Thermo Fisher
- Appareil d'électrophorèse et de révélation : Invitrogen™ E-Gel™ Power Snap Electrophoresis Device de chez Thermo Fisher.

3.5.1.5) Marqueurs de poids moléculaire (1 kb) :

- Thermo Scientific™ GeneRuler 1 kb DNA Ladder

3.5.1.6) Consommables :

- Microtubes PCR : barrettes PCR 4 puits, 200 µL (Sarstedt)

3.5.2) Manipulations.

3.5.2.1) Préparation du mélange réactionnel :

- Mélanger les composants de la réaction PCR dans les microtubes PCR (barrettes à 4 puits) (**Tableau 3.1**).

Composants	Volume (µL) x 1
AmpliTaq Gold 360 Master Mix (2×)	10
Amorce forward (1 µM)	5
Amorce reverse (1 µM)	5
ADN matrice	5
Total	25

Tableau 3.1 : Liste illustrant l'ensemble des réactifs utilisés pour la PCR ainsi que les volumes pipetés par échantillon.

3.5.2.2) Programmer le thermocycleur et déposer les micro-tubes PCR (Tableau 3.2) :

A) <i>E. granulosus</i>	Température	Durée	Nombre de cycles
Dénaturation initiale	95 °C	5 min	1
Dénaturation	95 °C	20 s	40
Hybridation	55 °C	20 s	40
Élongation	65 °C	4 min	40
Élongation finale	65 °C	7 min	1
Conservation	4 °C	∞	—

B) <i>E. multilocularis</i>	Température	Durée	Nombre de cycles
Dénaturation initiale	95 °C	10 min	1
Dénaturation	98 °C	20 s	40

Hybridation	55 °C	15 s	40
Élongation	65 °C	3 min	40
Élongation finale	65 °C	5 min	1
Conservation	4 °C	∞	—

Tableau 3.2 : Programme des thermocycleurs pour la PCR (A) *E. granulosus* et (B) *E. multilocularis*¹³.

3.5.2.3) Gels PCR :

- À la fin du programme PCR, récupérer les tubes PCR.
- Prendre le gel d'agarose à 2 % (E-Gel) prêt à l'emploi et identifier les puits (échantillons, contrôles, marqueur).
- Charger 10 µL d'eau PCR dans tous les puits et compléter avec 10 µl d'eau PCR dans les puits non utilisés (volume maximal par puits : 20 µL).
- Charger 10 µL d'amplicon dans les puits correspondants.
- Déposer 10 µL du marqueur de poids moléculaire dans le premier puits du gel.
- Placer le gel dans le système d'électrophorèse et lancer la migration selon les paramètres du fabricant pendant 20 minutes.
- À la fin de la migration, récupérer les images et les captures du gel, puis les enregistrer sur une clé USB.

3.6) Purification de l'ADN

Cette méthode de purification a été conçue pour assurer une purification rapide, peu coûteuse et compatible avec la préparation de bibliothèques Illumina et Nanopore, par captation d'ADN à l'aide de billes magnétiques à base de carboxylate, selon le principe de l'immobilisation réversible sur phase solide (SPRI).

3.6.1) Matériel :

3.6.1.1) Consommables :

- Strips low-bind: Applied Biosystems™ MicroAmp™ 8-Tube Strips (0,2 mL)
- Réservoirs : Costar® 50 mL Reagent Reservoirs, stériles (5/sachet)

3.6.1.2) Réactifs :

- Billes de purification: ABL DeepChek NGS Clean Up Beads (50 mL)
- Eau pour biologie moléculaire : Water for Molecular Biology (PanReac Appli-Chem, A7398.0500)
- Éthanol : éthanol 100 % qualité biologie moléculaire (Sigma-Aldrich) (E7023)

3.6.1.3) Équipement :

- Portoir magnétique pour strips 0,2 mL: 0.2 mL PCR Strip Magnetic Separator (24 tubes) (Permagen)

3.6.2) Manipulations (purification sur billes magnétiques).

3.6.2.1) Fixation de l'ADN sur les billes magnétiques :

- Pipeter 10-15 µL d'amplicons dans des strips *low-bind* à l'aide de la pipette multicanaux.
- Vortexer les billes (utilisées pour accrocher l'ADN) afin de les suspendre complètement.
- Ajouter un volume équivalent de billes (15 µL) à chaque échantillon à l'aide de la pipette multicanaux, puis homogénéiser par pipetage (*up-and-down*).
- Incuber 3 min à température ambiante.

3.6.2.2) Capture magnétique et élimination du surnageant :

- Placer les strips sur le portoir magnétique et attendre que les billes s'agrègent contre la paroi (jusqu'à la clarification du surnageant).
- Prélever et éliminer le surnageant sans perturber le culot des billes.

3.6.2.3) Lavages à l'éthanol 80 % :

- Laisser sécher les billes à l'air libre pendant ≤ 30 s (sans surséchage).
- Pendant ce temps, préparer de l'éthanol à 80 % et le verser dans un réservoir stérile.
- Ajouter 200 µL d'éthanol à 80 % dans chaque microtube à l'aide de la pipette multicanaux.
- Laisser le strip sur le portoir magnétique, puis effectuer une brève permutation ou une inversion contrôlée.
- Retirer l'éthanol à l'aide de la pipette multicanaux.
- Répéter l'étape de lavage.
- S'assurer qu'il ne reste aucun résidu d'éthanol (inhibiteur fréquent des réactions enzymatiques, notamment de la PCR et de certaines étapes de préparation de librairies).

3.6.2.4) Élution :

- Ajouter 20 µL d'eau pour la biologie moléculaire.
- Resuspendre complètement les billes par pipetage (*up-and-down*) jusqu'à obtention d'une suspension homogène brunâtre.
- Incuber 5 min à température ambiante pour permettre l'élution de l'ADN.
- Replacer les strips sur le portoir magnétique et attendre l'agrégation complète des billes (jusqu'à l'obtention d'une solution claire).

- Transférer 18 µL du surnageant (ADN purifié) dans de nouveaux strips low-bind.

3.7) Quantification des amplicons purifiés

Cette étape est essentielle pour déterminer la concentration d'ADN post-PCR après purification, afin d'effectuer la normalisation, première étape requise avant la préparation des bibliothèques de séquençage.

3.7.1) Matériel :

- Strips à 8 positions compatibles avec Qubit: Invitrogen™ Qubit™ Flex Assay Tube Strips, Thermo Fisher Scientific.
- Kit de quantification : Invitrogen™ Qubit™ 1X dsDNA HS Assay Kit, Thermo Fisher Scientific (Q33231).
- Fluoromètre : Invitrogen™ Qubit™ Flex Fluorometer.

3.7.2) Méthode (dosage Qubit dsDNA HS) :

3.7.2.1) Préparation des étalons (standards) :

- Prendre 2 strips (un pour l'étalon bas et un pour l'étalon haut).
- Distribuer 190 µL de réactif de dosage (working solution) dans chaque puits des strips à l'aide de la pipette multicanaux.
- Ajouter 10 µL d'étalon bas dans le strip correspondant.
- Ajouter 10 µL de l'étalon haut dans l'autre strip correspondant.
- Mélanger par pipetage (up-and-down) en évitant les bulles.

3.7.2.2) Préparation des échantillons :

- Distribuer 198 µL de réactif de dosage dans chaque puits destiné aux échantillons.
- Ajouter 2 µL d'ADN (amplicons purifiés) dans chaque puits.
- Mélanger soigneusement par pipetage (*up-and-down*).

3.7.2.3) Lecture au fluoromètre :

- Mesurer les strips dans le Qubit Flex en commençant par les étalons.
- Sélectionner l'essai dsDNA HS (ADN double brin, haute sensibilité), puis indiquer un volume d'échantillon de 2 µL.
- Relever les concentrations affichées par l'appareil (en ng/µL).

3.8) Séquençage Illumina

3.8.1) Matériel :

3.8.1.1) Consommables :

- Strips low-bind: Applied Biosystems™ MicroAmp™ 8-Tube Strips (0,2 mL)
- Réservoirs : Costar® 50 mL Reagent Reservoirs, stériles (5/sachet)
- Flowcell : *Flowcell* Illumina iSeq100

3.8.1.2) Réactifs :

- Billes de purification: ABL DeepChek NGS Clean Up Beads (50 mL)
- Eau pour biologie moléculaire : Water for Molecular Biology (PanReac Appli-Chem, A7398.0500).
- Éthanol : éthanol 100 % qualité biologie moléculaire (Sigma-Aldrich).
- Kit de préparation des librairies : DeepChek NGS Library Preparation V2, Deep Check Adapters (24) V2 (116B24).

3.8.1.3) Équipement :

- Portoir magnétique pour strips 0,2 mL: 0.2 mL PCR Strip Magnetic Separator (24 tubes) (Permagen)
- Séquenceur : Illumina iSeq100.

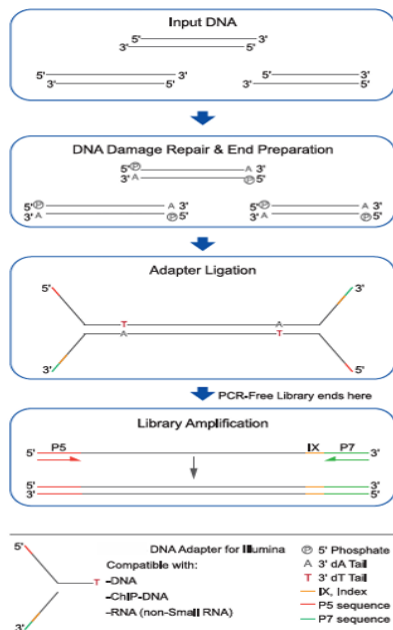


Figure 3.1 : Protocole DeepChek NGS ABL⁵⁴.

3.8.2) Protocole :

Le protocole utilisé pour cette manipulation est DeepChek® NGS Library Preparation + Adapters (RUO) – User Guide – 01/2024 : <https://www.abldiagnostics.com/products/deepchek-ngs-library-preparation-kit/>⁵⁴

Le protocole (**figure 3.1**) débute par une normalisation de l'ADN à 3 ng/μL pour l'ensemble des amplicons. Il est ensuite impératif de regrouper, pour chaque patient, tous les amplicons dans un même tube.

Ensuite, une étape de fragmentation de l'ADN est réalisée, puis la ligation d'adaptateurs, qui contiennent également un code-barres (barcode) attribué à chaque patient. L'étape suivante consiste à purifier les échantillons barcodés, puis à amplifier la librairie. Une nouvelle purification est ensuite effectuée, suivie d'une étape de normalisation à 2 ng/μL.

Enfin, la qualité des librairies est vérifiée sur un gel d'agarose avant de procéder à la dilution finale et au chargement de la librairie sur la *Flowcell* (pour plus de détails, se référer à l'insert du kit).

3.9) Séquençage Nanopore

3.9.1) Matériel :

3.9.1.1) Consommables :

- Strips low-bind: Applied Biosystems™ MicroAmp™ 8-Tube Strips (0,2 mL)
- Réservoirs : Costar® 50 mL Reagent Reservoirs, stériles (5/sachet)
- Flowcell : MinION/GridION Flowcell (DNA) FLO-MIN114/ version : R10.4.1

3.9.1.2) Réactifs :

- Eau pour biologie moléculaire : Water for Molecular Biology (PanReac AppliChem, A7398.0500)
- Éthanol : éthanol 100 % qualité biologie moléculaire (Sigma-Aldrich)
- Kit de préparation des bibliothèques : Rapid Barcoding Kit 24 V14 (SQK-RBK114.24)

3.9.1.3) Équipements :

- Portoir magnétique pour tubes : DynaMag-2 magnetic rack
- Séquenceur : MinION MK1D Nanopore

3.9.2) Protocole :

Le protocole utilisé pour cette manipulation est disponible sur : <https://store.Nanoporetech.com/us/rapid-barcoding-sequencing-kit-24-v14.html>⁵³

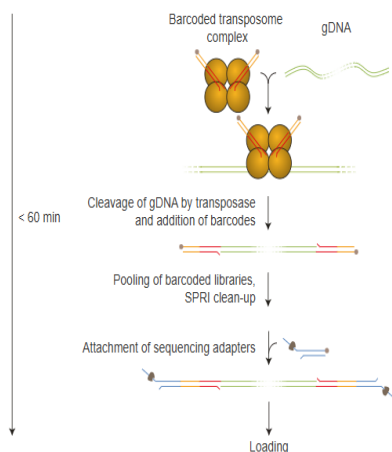


Figure 3.2 : Protocole Rapid Barcoding Kit Nanopore⁵³.

Le protocole (figure 3.2) commence par une normalisation à 6 ng/μL lorsqu'on travaille sur des amplicons. Ensuite, les amplicons de chaque patient sont regroupés. L'ensemble des amplicons d'un même patient est barcodé par un complexe de transposase qui clive l'ADN et y insère les barcodes correspondants.

Une fois ces étapes effectuées, les échantillons de tous les patients sont regroupés, puis l'ADN est purifié. Après cette purification, les adaptateurs sont ligués, puis la protéine motrice est attachée. Enfin, la bibliothèque est chargée sur la Flowcell du MinION (pour plus de détails, se référer à l'insert du kit⁵³).

3.10) Nombre de séquençages réalisés et contrôles négatifs utilisés

Au total, cinq séquençages ont été réalisés dans le cadre de ce travail : trois runs Nanopore et deux runs Illumina. En complément, les fichiers FASTQ issus de la technologie Illumina NextSeq500, correspondants aux échantillons ID21 à ID26, ont été fournis par le Centre National de Référence des Échinococcoses (CNR-E) du CHU de Besançon, en France.

Pour chaque batch d'extraction d'échantillons d'*E. multilocularis* et d'*E. granulosus*, un contrôle négatif constitué d'eau de biologie moléculaire a été ajouté. Ce contrôle a suivi l'ensemble du *workflow* expérimental, de l'extraction à la PCR et à la purification. Après amplification, les produits PCR issus d'un même contrôle négatif ont été regroupés en un pool unique, puis identifiés par un *barcode* spécifique lors de la préparation des bibliothèques de séquençage. Cette stratégie permettait de séquencer séparément le signal des contrôles négatifs et d'évaluer la présence éventuelle de contaminations introduites au cours du workflow, notamment lors de l'extraction, de l'amplification ou de la purification. Ce contrôle sera dénommé « contrôle négatif d'extraction ».

Un contrôle négatif supplémentaire, constitué d'eau de biologie moléculaire, a également été ajouté lors de la préparation des bibliothèques, uniquement pour le run Illumina. Ce blanc avait pour objectif d'évaluer les contaminations résiduelles susceptibles d'être introduites lors de la préparation des bibliothèques ou du séquençage. Ce contrôle sera dénommé « blanc de séquençage ».

3.11) Pipeline bioinformatique Illumina et Nanopore

Les lectures Illumina et Nanopore ont d'abord été filtrées pour éliminer les séquences de faible qualité. Pour les données Illumina, le prétraitement a été effectué avec **fastp (v1.0.1)**. Les adaptateurs ont été retirés automatiquement. Les extrémités de faible qualité ont été coupées dès que le score moyen est tombé en deçà du score Q20. Les lectures trop courtes (< 50 pb), trop riches en bases de faible qualité ou contenant plus de cinq bases indéterminées ont été écartées. Les queues poly-G et poly-X ont également été supprimées. La qualité des lectures, avant et après le nettoyage, a été vérifiée à l'aide de **FASTQC (v0.12.1)**. Pour les données Nanopore, le filtrage a été effectué avec **Chopper (v0.11.0)**. Seules les lectures supérieures à 500 pb ont été conservées. Un seuil de qualité moyen, Q15 ou Q10, a été utilisé en fonction de la perte de lecture. La qualité des lectures filtrées a ensuite été évaluée à l'aide de **NanoPlot (v1.46.2)**, notamment la distribution des longueurs, les scores de qualité et le rendement total. Les lectures nettoyées ont ensuite été alignées sur les génomes de référence. Pour *E. multilocularis*, la référence utilisée était AB018440.2. Pour *E. granulosus*, deux références ont été utilisées : NC_044548.1, qui contient les régions répétitives du mitogénome, et PQ874683.1, utilisée pour produire des consensus compatibles avec les analyses phylogénétiques. Les lectures Illumina ont été alignées avec **bwa-mem (v0.7.19-r1273)**. Les lectures Nanopore ont été alignées avec **minimap2 (v2.30-r1287)**, avec le préréglage map-ont adapté aux longues lectures. Les lectures Illumina ont été conservées si leur score de mapping (MAPQ) était supérieur ou égal à 20. Pour les lectures Nanopore, un seuil de MAPQ ≥ 10 a été utilisé, car les lectures longues restent souvent informatives. Les alignements ont ensuite été triés et indexés avec **samtools (v1.23)**. Plusieurs métriques ont été calculées pour chaque échantillon : la proportion

de lectures alignées, la fraction du génome couverte et la profondeur moyenne de séquençage. Ces métriques ont été obtenues avec samtools flagstat, samtools coverage et samtools depth. L'appel de variants a été réalisé séparément selon la technologie. Pour les données Illumina, les variants ont été appelés avec **freebayes (v1.3.10)**, en mode haploïde. Les positions retenues devaient présenter une qualité de mapping et de base d'au moins Q20. L'allèle alternatif devait représenter au moins 25 % des lectures et être soutenu par au moins dix lectures. Les variants ont ensuite été filtrés avec **bcftools (v1.23)** selon une profondeur locale (DP) > 30. Les seuils de qualité étaient de QUAL > 150 pour les polymorphismes (SNPs) et de QUAL > 50 pour les insertions et délétions. Pour les données Nanopore, les variants ont été appelés avec **Clair3 (v2.0.1)**. Le modèle utilisé était adapté à la chimie r1041_e82_400bps_hac. L'analyse a été réalisée en mode haploïde, avec l'option --haploid_precise. Les variants ont ensuite été filtrés avec bcftools selon une profondeur locale (DP) ≥ 30 et une qualité (QUAL) ≥ 15. Pour les deux technologies, les positions couvertes à moins de 30× ont été considérées comme insuffisamment fiables. Les séquences consensus ont ensuite été générées avec bcftools consensus. Les variants validés ont été intégrés à la référence. Cette étape évite d'attribuer une base lorsque les données ne sont pas suffisantes. L'ensemble des métriques de qualité a finalement été résumé dans un rapport **MultiQC (v1.33)** (voir l'Annexe II).

3.12) Construction de l'arbre phylogénétique et du réseau d'haplotypes

Les séquences consensus ont d'abord été regroupées dans un fichier unique. Elles ont ensuite été alignées avec **MAFFT (v7.526)**. Cette étape permet de comparer les positions homologues entre les séquences. L'option --localpair, correspondant à la méthode L-INS-i, a été utilisée. Le nombre maximal d'itérations a été fixé à 1000 afin d'améliorer la qualité de l'alignement. L'arbre phylogénétique a ensuite été construit avec **IQ-TREE (v3.0.1)**, selon la méthode du maximum de vraisemblance. Cette approche vise à identifier l'arbre le plus compatible avec les différences observées, selon un modèle d'évolution moléculaire. En complément, un réseau d'haplotypes a été réalisé à l'aide de **PopART (v1.7)** selon la méthode *median-joining network* (voir l'Annexe II).

Chapitre IV : Résultats



Ressources

L'ENSEMBLE DES FIGURES RELATIVES A LA SECTION DES RESULTATS ET LE CODE R SE TROUVE ICI :

[HTTPS://DOI.ORG/10.5281/ZENODO.20689318](https://doi.org/10.5281/ZENODO.20689318)

DANS LE BUT D'OPTIMISER LA QUALITE DES ILLUSTRATIONS ET DE FACILITER LA LECTURE DE CE MEMOIRE, LE DOI SUIVANT DONNE ACCES A L'ENSEMBLE DE MES FIGURES POUR LA PARTIE RESULTATS.

CHAPITRE IV : RÉSULTATS

4.1) Séquençage Illumina iSeq100 d'*E. multilocularis*

La figure 4.1 présente les principales métriques de qualité obtenues après le séquençage Illumina iSeq100 d'échantillons d'*E. multilocularis* issus de la cohorte liégeoise (16 échantillons, ID1 à ID16). Les résultats regroupent deux séries de séquençage. Pour certains échantillons, plusieurs réplicats techniques ont été réalisés et sont indiqués par un suffixe ajouté à l'identifiant, par exemple ID8_2 et ID15_3, correspondants respectivement à un duplicata et à un triplicata. Trois contrôles sont également inclus dans l'analyse : EMG-neg, correspondant à un contrôle négatif d'extraction (issu du premier run) ; ctrl-seq, correspondant au contrôle de séquençage (blanc) ; et ctrl-em, correspondant à un contrôle négatif d'extraction (issu du deuxième run) (voir Matériel et méthode 3.10). La majorité des échantillons analysés présente une couverture élevée du mitogénome, généralement proche de 100 % (figure 4.1.A). Cependant, plusieurs échantillons affichent une couverture plus faible, notamment ID1, ID4, ID7, ID13 et ID12_2 (figure 4.1.A). Le contrôle négatif EMG-neg montre une couverture partielle, estimée à environ 48,2 %, tandis que les deux autres contrôles affichent des valeurs supérieures à 90 % (figure 4.1.A). La distribution globale montre que la plupart des échantillons se répartissent dans des classes de couverture élevées, principalement entre 80 % et 100 % (figure 4.1.B). Les échantillons présentant une couverture élevée du mitogénome affichent généralement des profondeurs de séquençage élevées, souvent supérieures à 2 000× (figure 4.1.C). La profondeur maximale est observée pour l'échantillon ID8_2. À l'inverse, les échantillons ID1, ID4, ID7 et ID13 présentent des profondeurs nettement plus faibles, tout comme les contrôles négatifs (figure 4.1.C). La distribution des profondeurs confirme cette tendance, avec une majorité d'échantillons fortement couverts et un nombre restreint présentant des profondeurs plus faibles (figure 4.1.D). Pour la majorité des échantillons, aucune position masquée n'est observée, ce qui correspond à des consensus mitochondriaux entièrement résolus (figure 4.1.E). En revanche, plusieurs échantillons présentent un nombre plus élevé de positions non résolues, notamment ID1, ID4, ID7 et ID13 (figure 4.1.E). La distribution des positions masquées montre ainsi que la majorité des échantillons présente peu ou aucune position non résolue, tandis qu'un nombre limité d'échantillons se distingue par des valeurs plus élevées (figure 4.1.F). Dans les échantillons présentant une bonne couverture et une profondeur élevée du mitogénome, le nombre de variants reste globalement stable après le filtrage bioinformatique (figures 4.1.G et 4.1.H).

4.2) Séquençage Illumina iSeq100 d'*E. granulosus*

La figure 4.2 présente les métriques de qualité obtenues après le séquençage sur l'iSeq 100 d'Illumina des échantillons d'*E. granulosus* issus de la cohorte liégeoise. Les données analysées incluent trois échantillons cliniques (ID18 à ID20) ainsi qu'un contrôle négatif d'extraction, noté Eg-neg. Comme pour les échantillons d'*E. multilocularis*, l'analyse repose sur la reconstruction du mitogénome à partir de plusieurs fragments de PCR chevauchants (voir Matériel et méthode 3.10).

Les trois échantillons cliniques présentent des valeurs de couverture proches les unes des autres, comprises entre 75,9 % et 78,7 % (figures 4.2.A et 4.2.B). Aucun des trois échantillons n'atteint donc une couverture complète du mitogénome. Cette couverture partielle doit cependant être interprétée en tenant compte de la région répétitive du mitogénome d'*E. granulosus*, difficilement résolue par le séquençage Illumina. Le contrôle négatif Eg-neg présente, quant à lui, une couverture de 47,7 %, nettement inférieure à celle des échantillons cliniques (figure 4.2.A).

Les trois échantillons cliniques présentent des profondeurs de séquençage élevées, comprises entre 2 500× et 3 256× (figures 4.2.C et 4.2.D). Ils présentent donc un profil comparable, associant une couverture partielle du mitogénome à une profondeur élevée dans les régions correctement séquencées. À l'inverse, le contrôle négatif Eg-neg montre une profondeur très faible, d'environ 3× (figure 4.2.C). Les trois échantillons cliniques présentent entre 3 947 et 6 006 positions masquées (figures 4.2.E et 4.2.F). Ces positions correspondent à des sites du mitogénome pour lesquels aucun nucléotide suffisamment étayé n'a pu être déterminé dans la séquence consensuelle finale. Le contrôle négatif Eg-neg se distingue par un nombre beaucoup plus élevé de positions masquées, soit environ 17 675, ce qui correspond à l'absence de consensus exploitable sur la quasi-totalité du mitogénome (figure 4.2.E).

Enfin, les trois échantillons cliniques présentent au moins 30 variants bruts avant filtrage, dont une partie est conservée après application des critères bioinformatiques (figures 4.2.G et 4.2.H). Mais une diminution importante du nombre de variants est observée après filtrage, ce qui indique que seuls les variants suffisamment supportés sont retenus (figures 4.2.G et 4.2.H). Le contrôle négatif Eg-neg présente également plusieurs variants détectés avant le filtrage, mais aucun variant n'est conservé après le filtrage bioinformatique (Figure 4.2.G).

Figure 4.2.

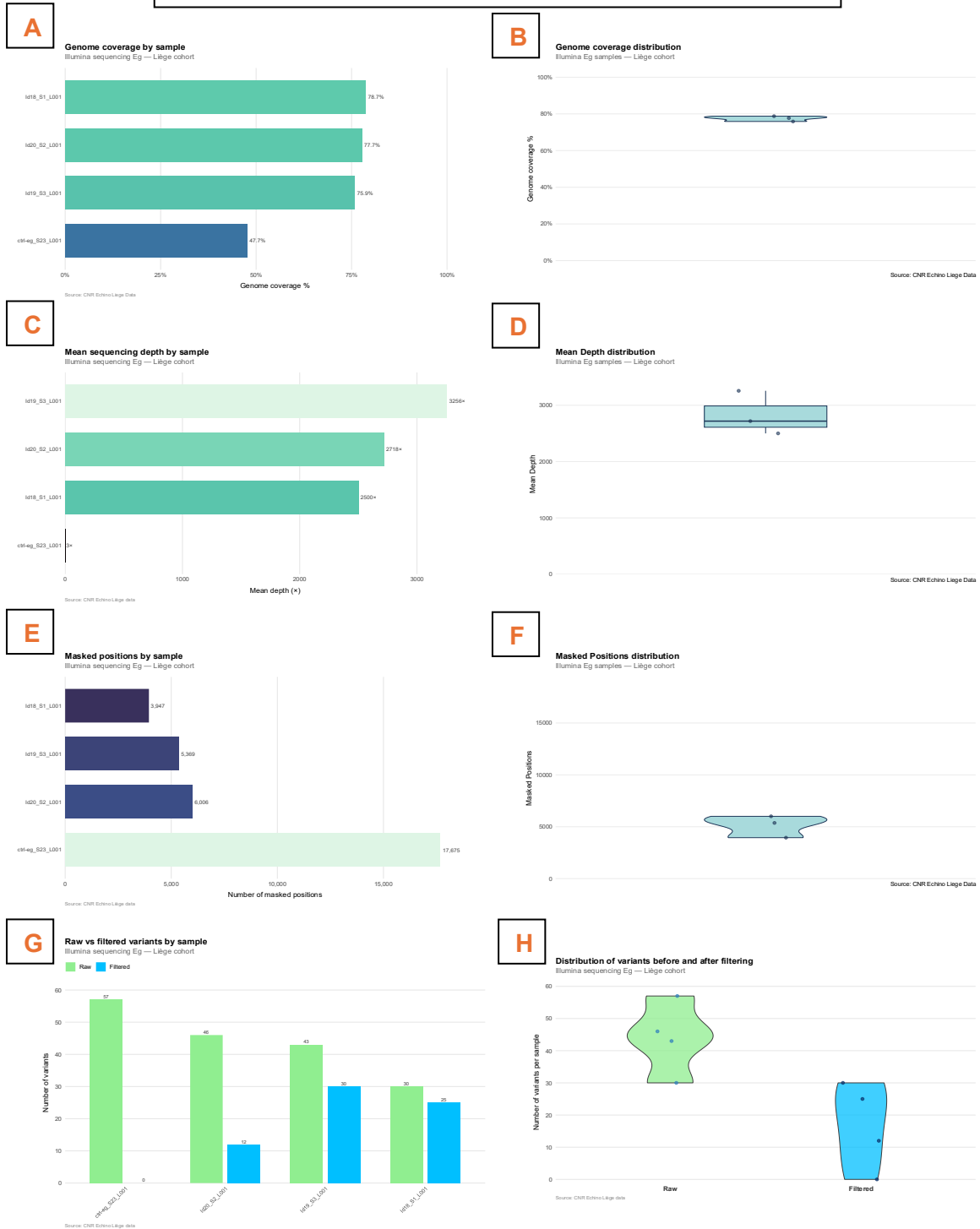


Figure 4.2. Métriques de qualité du séquençage Illumina pour la reconstruction du mitogénome d'E. granulosus dans la cohorte de Liège, à partir de 3 échantillons.

(A) couverture du génome mitochondrial par échantillon, (B) distribution globale de la couverture, (C) profondeur moyenne de séquençage par échantillon, (D) distribution de la profondeur de séquençage, (E) nombre de positions masquées dans les séquences consensus, (F) distribution des positions masquées entre les échantillons, (G) nombre brut et filtré de variants mitochondriaux par échantillon, et (H) distribution de variants mitochondriaux filtrés et bruts par échantillon.

4.3) Séquençage ONT MinION Mk1D d'*E. multilocularis* / Liège

La figure 4.3 présente les métriques de qualité obtenues après le séquençage Nanopore MinION Mk1D des échantillons d'*E. multilocularis* issus de la cohorte liégeoise. Deux stratégies de barcoding ont été employées pour l'échantillon ID8. Les identifiants ID8_1 à ID8_7 correspondent aux différents amplicons de l'échantillon ID8, chacun barcodé et analysé séparément. Ils ne constituent donc pas des réplicats techniques, mais des sous-unités, chacune correspondant à une portion distincte du mitogénome. L'échantillon ID8all correspond au regroupement bioinformatique de ces amplicons barcodés individuellement. À l'inverse, les échantillons ID8, ID5, ID6, ID9 et ID15 correspondent à une stratégie de barcode unique appliquée à l'ensemble des amplicons d'un même échantillon. Un contrôle négatif d'extraction est également inclus dans l'analyse (Neg_) (voir Matériel et méthode 3.10).

Les échantillons ID8all, ID8, ID5, ID6, ID9 et ID15 présentent une couverture complète ou quasi complète du mitogénome (figures 4.3.A et 4.3.B). Les amplicons ID8_1 à ID8_7, analysés individuellement, présentent des couvertures plus variables et, globalement, plus faibles, ce qui est attendu puisqu'ils ne couvrent chacun qu'une portion du mitogénome (figure 4.3.A). Après regroupement bioinformatique des sept amplicons, l'échantillon ID8all atteint une couverture complète, comparable à celle observée pour l'échantillon ID8 traité selon une stratégie de barcode unique (figure 4.3.A). Parmi les échantillons analysés, ID8all présente la profondeur moyenne la plus élevée, avec une valeur de 4 696× (figure 4.3.C). Cette profondeur est supérieure à celle observée pour l'échantillon ID8 séquençé selon une stratégie de barcode unique, qui atteint 801× (figure 4.3.C). Les amplicons ID8_1 à ID8_7 montrent des profondeurs variables selon le fragment considéré. Les échantillons ID5, ID15, ID6 et ID9 présentent également des profondeurs mesurables dans les régions couvertes, mais avec des valeurs inférieures à celles observées pour ID8all (figures 4.3.C et 4.3.D). Le contrôle négatif présente une profondeur très faible, d'environ 1× (figure 4.3.C). Les échantillons ID8all et ID8 présentent un faible nombre de positions masquées, avec respectivement 4 et 6 positions non résolues (figure 4.3.E). Les échantillons ID5, ID15, ID6 et ID9 présentent un nombre plus élevé de positions masquées que les échantillons ID8 et ID8all (figures 4.3.E et 4.3.F). L'échantillon présentant le plus grand nombre de positions masquées est l'ID6, avec 7 266 positions non résolues (figure 4.3.E). Les échantillons reconstruits à l'échelle du mitogénome, en particulier ID8all, ID8, ID15, ID5, ID6 et ID9, conservent des variants après filtrage bioinformatique (figures 4.3.G et 4.3.H). Les amplicons ID8 analysés séparément conservent moins de variants, en lien avec leur couverture partielle du mitogénome. Le contrôle négatif ne conserve aucun variant après le filtrage bioinformatique (figure 4.3.G).

Figure 4.3.

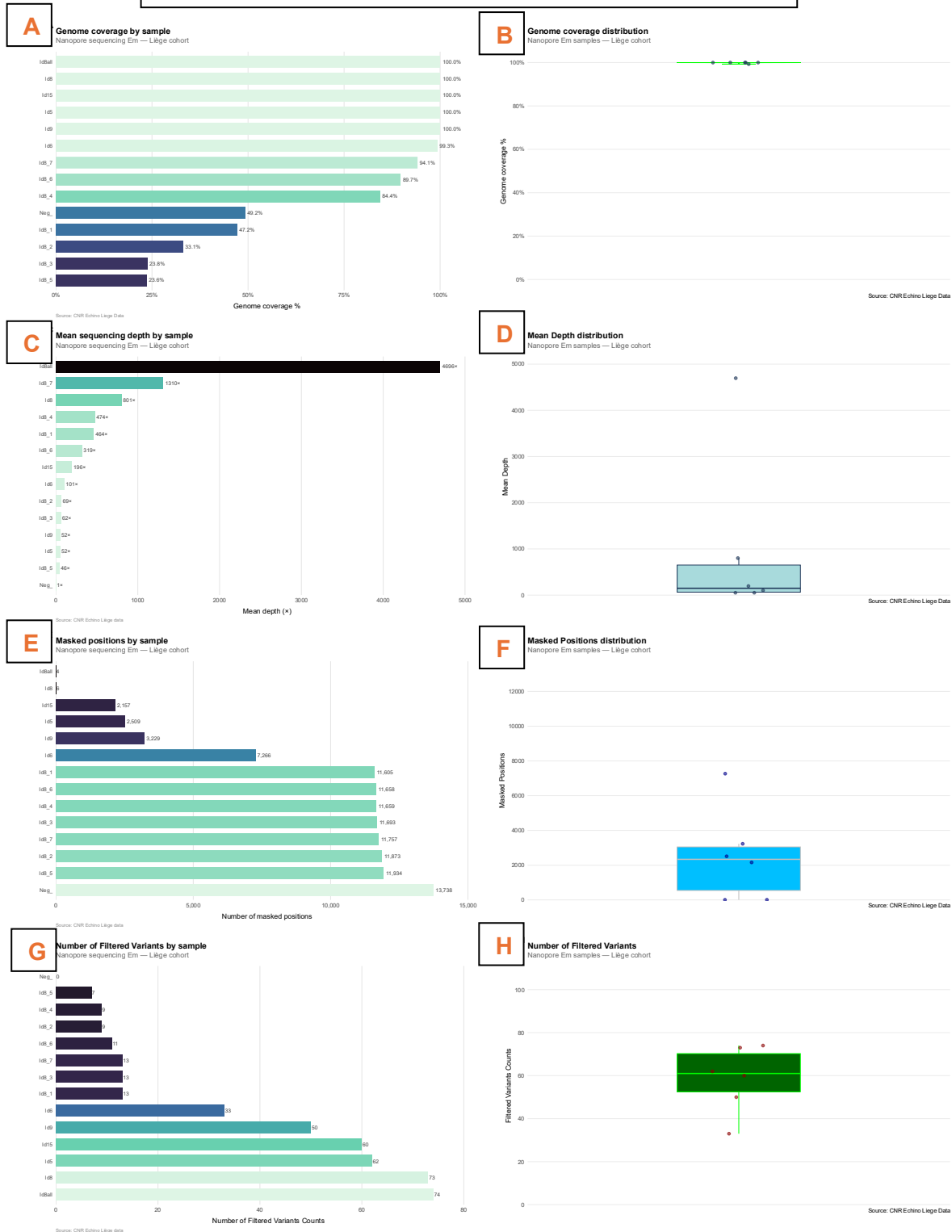


Figure 4.3. Métriques de qualité du séquençage Nanopore pour la reconstruction du mitogénome d'E. multilocularis dans la cohorte de Liège, à partir de 5 échantillons.

(A) couverture du génome mitochondrial par échantillon, (B) distribution globale de la couverture, (C) profondeur moyenne de séquençage par échantillon, (D) distribution de la profondeur de séquençage, (E) nombre de positions masquées dans les séquences consensus, (F) distribution des positions masquées entre les échantillons, (G) nombre de variants mitochondriaux filtrés par échantillon, et (H) distribution de variants mitochondriaux filtrés par échantillon.

4.4) Séquençage ONT MinION Mk1D d'*E. granulosus* / Liège

La figure 4.4 présente les métriques de qualité obtenues après le séquençage Nanopore MinION Mk1D des échantillons d'*E. granulosus* issus de la cohorte liégeoise. Les identifiants ID19_1 à ID19_8 correspondent aux différents amplicons de l'échantillon ID19, chacun barcodé et analysé séparément. L'échantillon ID19all correspond au regroupement bioinformatique de ces amplicons barcodés individuellement. Les échantillons ID17, ID18 et ID20, ainsi qu'un contrôle négatif d'extraction (Neg_), sont également inclus dans l'analyse (voir **Matériel et méthode 3.10**).

L'échantillon ID20 atteint une couverture complète du mitogénome, avec une valeur de 100 % (**figure 4.4.A**). Les échantillons ID17, ID18, ID19 et ID19all présentent des couvertures proches, variant de 74,1 % à 80,4 % (**figures 4.4.A et 4.4.B**). Les amplicons ID19_1 à ID19_8, analysés individuellement, montrent des couvertures partielles comprises entre 24,9 % et 50,6 % (**figure 4.4.A**). Le contrôle négatif présente la couverture la plus faible, avec une valeur de 5,9 % (**figure 4.4.A**). L'échantillon ID19all se distingue par la profondeur moyenne la plus élevée, avec une valeur de 21104× (**figure 4.4.C**). Cette profondeur est nettement supérieure à celle observée pour l'échantillon ID19 séquençé sous un barcode unique, qui atteint 2 556× (**figure 4.4.C**). Les échantillons ID17, ID18, ID20 et ID19 présentent des profondeurs plus faibles que celles d'ID19all (**figures 4.4.C et 4.4.D**). L'échantillon ID20 présente le nombre le plus faible de positions masquées parmi les échantillons cliniques testés, avec 1 745 positions non résolues (**figure 4.4.E**). Les échantillons ID19all, ID19 et ID18 présentent des valeurs proches, comprises entre 4 104 et 4 349 positions masquées (**figures 4.4.E et 4.4.F**). L'échantillon ID17 présente le plus grand nombre de positions masquées parmi les échantillons cliniques, avec 5 777 positions non résolues (**figure 4.4.E**). Le contrôle négatif présente le nombre le plus élevé de positions masquées, avec 17 675 positions non résolues, ce qui correspond à l'absence de consensus exploitable sur la quasi-totalité du mitogénome (**figure 4.4.E**).

Les échantillons reconstruits sur une grande proportion du mitogénome, notamment ID17, ID18, ID19, ID20 et ID19all, conservent des variants après le filtrage bioinformatique (**figures 4.4.G et 4.4.H**). Le nombre de variants retenus varie selon les échantillons. Les amplicons ID19_1 à ID19_8, analysés séparément, conservent un nombre plus faible de variants après filtrage, en lien avec leur couverture partielle du mitogénome (**figure 4.4.G**). Le contrôle négatif ne conserve aucun variant après le filtrage bioinformatique (**figure 4.4.G**).

Figure 4.4.

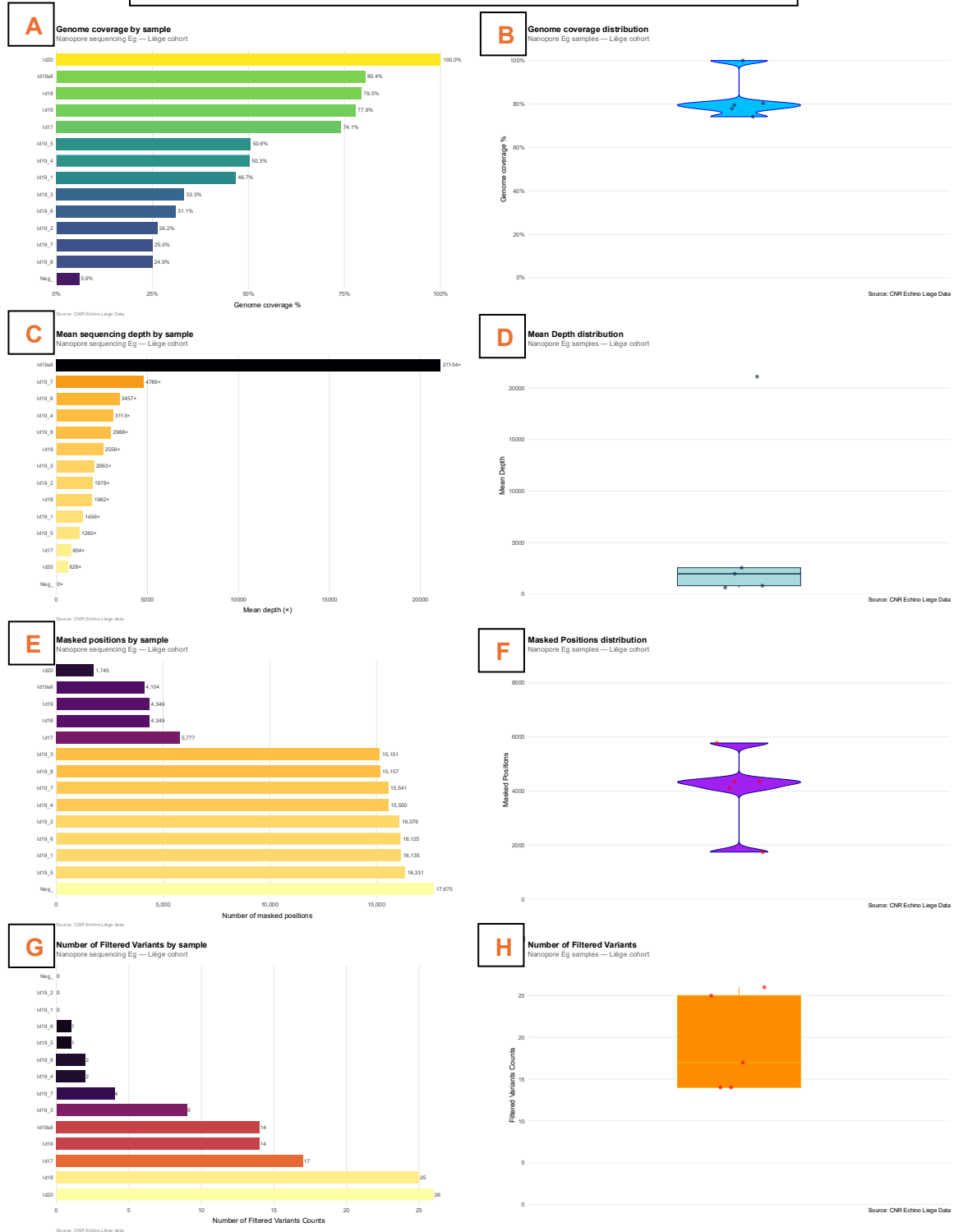


Figure 4.4. Métriques de qualité du séquençage Nanopore pour la reconstruction du mitogénome d'E. granulosus dans la cohorte de Liège, à partir de 4 échantillons.

(A) couverture du génome mitochondrial par échantillon, (B) distribution globale de la couverture, (C) profondeur moyenne de séquençage par échantillon, (D) distribution de la profondeur de séquençage, (E) nombre de positions masquées dans les séquences consensus, (F) distribution des positions masquées entre les échantillons, (G) nombre de variants mitochondriaux filtrés par échantillon, et (H) distribution de variants mitochondriaux filtrés par échantillon.

4.5) Séquençage ONT MinION Mk1D d'*E. multilocularis* / Besançon

La figure 4.5 présente les métriques de qualité obtenues après le séquençage Nanopore MinION Mk1D des échantillons d'*E. multilocularis* issus de la cohorte de Besançon, réalisé dans notre laboratoire. Trois échantillons cliniques sont représentés : ID21, ID22 et ID23. Un contrôle négatif d'extraction est également inclus dans l'analyse (Neg_Multi) (voir Matériel et méthode 3.10). Pour chaque échantillon, les amplicons ont été répartis sur deux barcodes, chacun contenant environ la moitié des amplicons, afin d'augmenter le nombre de *reads* disponibles par mitogénome et donc la profondeur de séquençage.

Les trois échantillons cliniques atteignent une couverture de 100 % du mitogénome (figures 4.5.A et 4.5.B). Le contrôle négatif présente une couverture plus faible, à 16,6 % (figure 4.5.A). Les profondeurs moyennes des échantillons cliniques varient entre 661× et 1 642× (figures 4.5.C et 4.5.D), tandis que le contrôle négatif présente une profondeur nulle ou quasi nulle (figure 4.5.C).

Les échantillons ID21 et ID22 ne présentent aucune position masquée (figures 4.5.E et 4.5.F). Les consensus mitochondriaux correspondants ne comportent donc aucune position non résolue. L'échantillon ID23 présente un faible nombre de positions masquées, avec seulement 38 positions non déterminées, malgré une couverture globale complète du mitogénome (figure 4.5.E). À l'inverse, le contrôle négatif présente un nombre beaucoup plus élevé de positions masquées, avoisinant 13 738, ce qui correspond à l'absence de consensus exploitable sur la quasi-totalité du mitogénome (figure 4.5.E).

Enfin, les trois échantillons cliniques conservent un nombre de variants relativement proche après le filtrage bioinformatique, compris entre 65 et 70 variants (figures 4.5.G et 4.5.H). Le contrôle négatif ne conserve aucun variant après filtrage (figure 4.5.G).

Figure 4.5.



Figure 4.5. Métriques de qualité du séquençage Nanopore pour la reconstruction du mitogénome d'*E. multilocularis* dans la cohorte de Besancon, à partir de 3 échantillons.

(A) couverture du génome mitochondrial par échantillon, (B) distribution globale de la couverture, (C) profondeur moyenne de séquençage par échantillon, (D) distribution de la profondeur de séquençage, (E) nombre de positions masquées dans les séquences consensus, (F) distribution des positions masquées entre les échantillons, (G) nombre de variants mitochondriaux filtrés par échantillon, et (H) distribution de variants mitochondriaux filtrés par échantillon.

4.6) Séquençage ONT MinION Mk1D d'*E. granulosus* / Besançon

La figure 4.6 présente les métriques de qualité obtenues après le séquençage Nanopore MinION Mk1D des échantillons d'*E. granulosus* issus de la cohorte de Besançon, réalisé au sein de notre laboratoire. Trois échantillons cliniques sont représentés : ID24, ID25 et ID26, ainsi qu'un contrôle négatif d'extraction (Neg_granu) (voir Matériel et méthode 3.10). Pour chaque échantillon, les amplicons ont été répartis sur deux barcodes, chacun contenant environ la moitié des amplicons, afin d'augmenter le nombre de *reads* disponibles par mitogénome et donc la profondeur de séquençage.

L'échantillon ID26 atteint une couverture complète du mitogénome, avec une valeur de 100 % (figures 4.6.A et 4.6.B). L'échantillon ID24 présente également une couverture élevée de 97,1 % (figures 4.6.A et 4.6.B). L'échantillon ID25 présente une couverture plus faible que les deux autres échantillons cliniques, avec une valeur de 87,3 % (figures 4.6.A et 4.6.B). Le contrôle négatif présente une couverture plus limitée, d'environ 14,3 % (figure 4.6.A).

Les trois échantillons cliniques présentent des profondeurs de séquençage élevées, comprises entre 2 371× et 3 184× (figures 4.6.C et 4.6.D). Le contrôle négatif présente une profondeur nulle (figure 4.6.C).

Parmi les échantillons cliniques, ID26 présente le nombre le plus faible de positions non résolues, avec 2 680 positions masquées (figures 4.6.E et 4.6.F). L'échantillon ID24 présente 4 486 positions masquées, tandis qu'ID25 présente le nombre le plus élevé parmi les échantillons cliniques, avec 5 772 positions masquées (figures 4.6.E et 4.6.F). Le contrôle négatif présente le nombre de positions masquées le plus élevé de l'ensemble de la série, avec 17 675 positions non résolues (figure 4.6.E).

Les trois échantillons cliniques conservent des variants après filtrage bioinformatique, avec 14 pour ID24, 16 pour ID25 et 29 pour ID26, ce dernier présentant donc le plus grand nombre de variants filtrés (figures 4.6.G et 4.6.H). Le contrôle négatif ne conserve aucun variant après filtrage (figure 4.6.G).

Note : L'échantillon ID3 a été séquençé lors du même run Nanopore que celui utilisé pour les figures 4.7 et 4.8, mais il n'est pas représenté dans les six premières figures afin de conserver une organisation cohérente des résultats. Il sera présenté dans les analyses de validation ultérieures.

Figure 4.6.

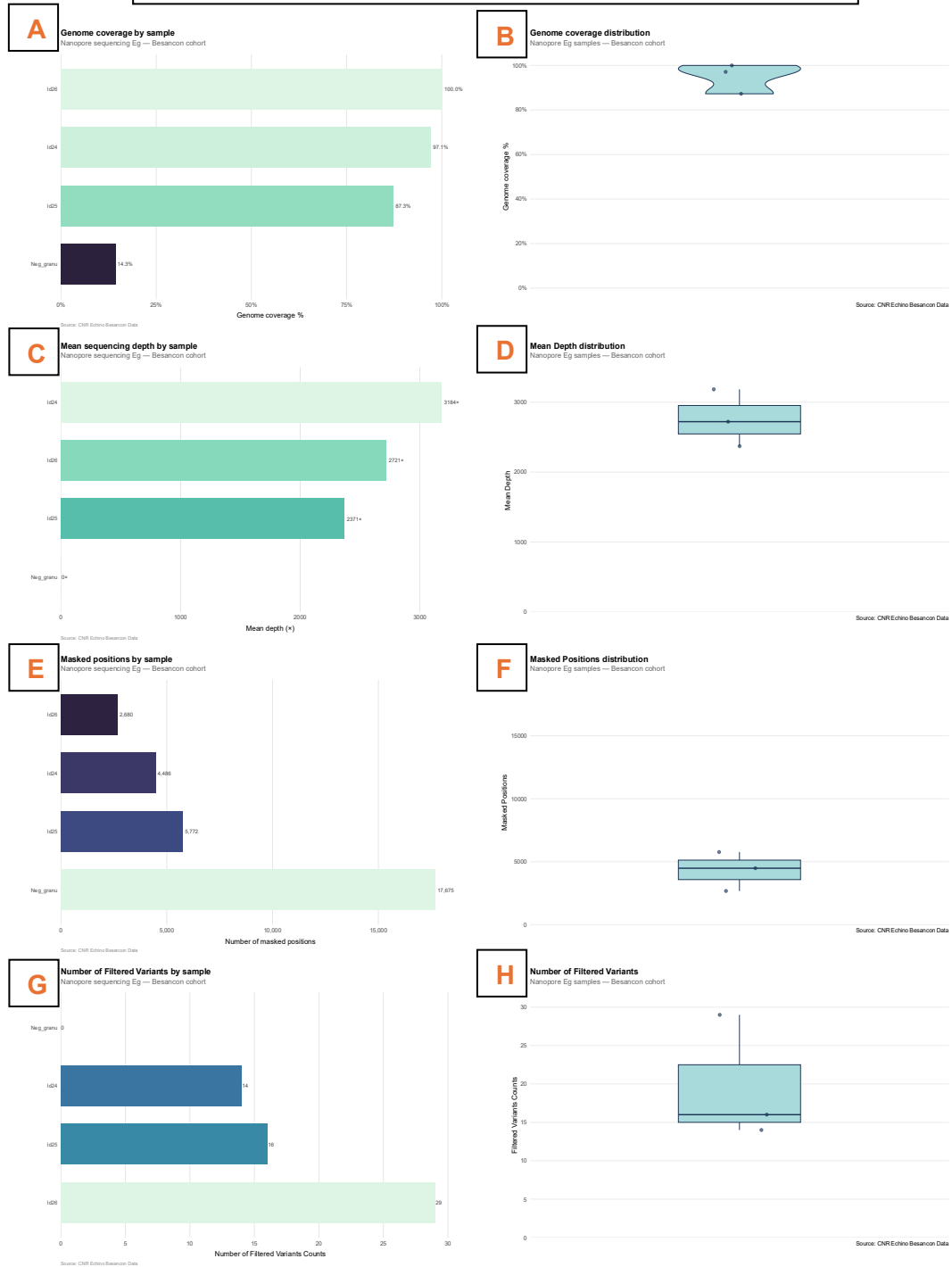


Figure 4.6. Métriques de qualité du séquençage Nanopore pour la reconstruction du mitogénome d'*E. granulosus* dans la cohorte de Besancon, à partir de 3 échantillons.

(A) couverture du génome mitochondrial par échantillon, (B) distribution globale de la couverture, (C) profondeur moyenne de séquençage par échantillon, (D) distribution de la profondeur de séquençage, (E) nombre de positions masquées dans les séquences consensus, (F) distribution des positions masquées entre les échantillons, (G) nombre de variants mitochondriaux filtrés par échantillon, et (H) distribution de variants mitochondriaux filtrés par échantillon.

4.7) Validation interne Illumina iSeq100

La figure 4.7 présente l'analyse de concordance interne (*intra-run*) du séquençage Illumina iSeq100 réalisée sur quatre échantillons en duplicata (ID8, ID9, ID15 et ID16) d'*E. multilocularis* issus de la cohorte liégeoise et présentant un mitogénome complet. Pour chaque échantillon, deux réplicats techniques indépendants, notés A et B, ont été comparés. Les analyses ont été réalisées séparément pour les polymorphismes nucléotidiques simples (SNPs), les insertions/délétions (INDELs) et l'ensemble des variants filtrés (ALL).

Les SNPs constituent la majorité des événements observés, tandis que les INDELs sont moins représentés (figures 4.7.A et 4.7.B). Parmi les substitutions, plusieurs types de changements nucléotidiques sont identifiés, avec une représentation notable des substitutions T→C, G→A, A→G, C→T et T→G (figure 4.7.A). Les variants détectés, y compris des SNPs et des INDELs, sont répartis sur plusieurs portions de la séquence mitochondriale, ce qui indique qu'ils ne sont pas limités à une seule région du mitogénome (figure 4.7.C). Plusieurs différences sont observées entre les réplicats A et B pour les échantillons ID8 et ID9, tandis qu'une seule discordance est observée pour les échantillons ID15 et ID16 (figure 4.7.D). Les variants discordants comprennent à la fois des SNPs et des INDELs.

La concordance entre les réplicats techniques a ensuite été évaluée à l'aide de l'indice de Jaccard, qui permet d'estimer la proportion de variants communs aux réplicats A et B par rapport à l'ensemble des variants détectés dans au moins l'un des deux réplicats techniques (figures 4.7.E et 4.7.F). Pour les SNPs, les valeurs de concordance sont élevées pour les quatre échantillons, comprises entre 95,4 % et 100 % (figures 4.7.E et 4.7.F). Pour les INDELs, la valeur la plus faible est observée pour ID9, avec une concordance de 50,0 %, tandis que ID8 et ID16 atteignent 100 % et ID15 atteint 83,3 % (figures 4.7.E et 4.7.F). Lorsque l'ensemble des variants est pris en compte, les valeurs de concordance restent comprises entre 91,5 % et 98,6 % (figures 4.7.E et 4.7.F). Le rappel, défini ici comme la proportion de variants détectés dans A et retrouvés par B, est compris entre 97 % et 99 % pour l'ALL (figure 4.7.G). La précision, définie comme la proportion de variants détectés dans B et confirmés par A, est comprise entre 92 % et 100 % pour l'ensemble des variants (figure 4.7.H). Les variations les plus marquées concernent principalement les INDELs, notamment pour ID9 (figure 4.7.H).

Figure 4.7.



Figure 4.7. Validation intra-run Illumina de la détection des variants mitochondriaux chez *E. multilocularis*.

A) liste et proportion des variants détectés entre les répliquats techniques Illumina. (B) distribution des types de variants détectés entre les répliquats techniques Illumina. (C) types de variants stratifiés selon leur position le long du génome mitochondrial. (D) tableau des variants discordants détectés entre le répliquat A et le répliquat B. (E) concordance de Jaccard entre les répliquats Illumina, par échantillon et par type de variant. (F) Distribution de la concordance de Jaccard. (G) rappel des variants détectés dans le répliquat A et retrouvés dans le répliquat B. (H) précision des variants détectés dans le répliquat B et confirmés dans le répliquat A.

4.8) Validation externe ONT MinION Mk1D

La figure 4.8 présente une comparaison des variants mitochondriaux détectés par séquençage sur l'Illumina NextSeq 500 (référence) et sur le Nanopore MinION MK1D. Afin de valider l'approche Nanopore, six échantillons cliniques provenant du CNR-E du CHU de Besançon ont été séquencés au laboratoire de microbiologie de Liège à l'aide de la technologie Nanopore : trois échantillons d'*E. multilocularis* (ID21, ID22 et ID23) et trois échantillons d'*E. granulosus* (ID24, ID25 et ID26).

Les résultats obtenus ont ensuite été comparés aux données de séquençage Illumina NextSeq 500 générées par l'équipe française. La comparaison porte sur les SNPs, les INDELs et les ALL. Les SNPs constituent la majorité des événements observés, tandis que les INDELs sont beaucoup moins représentés (figures 4.8.A et 4.8.B).

Les substitutions T→C, G→A, C→T et A→G figurent parmi les changements nucléotidiques les plus fréquemment observés. Les variants détectés sont répartis sur plusieurs régions du mitogénome (figure 4.8.C), et plusieurs discordances entre les deux technologies sont recensées dans la table récapitulative (figure 4.8.D).

La concordance entre Illumina et Nanopore a ensuite été évaluée à l'aide de l'indice de Jaccard pour les SNPs et pour l'ensemble des variants (figures 4.8.E et 4.8.F). Pour les SNPs, les valeurs de concordance varient de 80,0 % à 100,0 % (figures 4.8.E et 4.8.F). La valeur la plus faible est observée pour l'ID25, tandis que l'ID24 atteint 100,0 % de concordance. Lorsque l'ensemble des variants est considéré, les valeurs varient de 72,4 % à 100,0 %, la concordance la plus faible étant observée pour l'ID26 et la plus élevée pour l'ID24 (figure 4.8.E). Le rappel, défini ici comme la proportion de variants détectés par Illumina et retrouvés par Nanopore, varie entre 80,0 % et 100,0 % pour les SNPs et les ALL (figure 4.8.G). La précision, définie comme la proportion des variants détectés par Nanopore qui sont confirmés par Illumina, varie entre 84,0 % et 100,0 % pour les SNPs, et entre 72,4 % et 100,0 % pour l'ensemble des variants (figure 4.8.H). Les écarts les plus marqués concernent surtout les échantillons ID25 et ID26, tandis que les autres présentent globalement des concordances, des rappels et des précisions élevés.

Figure 4.8.



Figure 4.8. Validation inter-run et inter-institut de la détection des variants mitochondriaux par Nanopore et Illumina chez les cas de *E. multilocularis* et *E. granulosus* entre Liège et Besançon.

A) liste et proportion des variants détectés entre les répliquats techniques. (B) distribution des types de variants détectés entre les répliquats techniques. (C) Répartition des types de variants selon leur position le long du génome mitochondrial. (D) Tableau des variants discordants détectés entre les deux runs/instituts. (E) Concordance de Jaccard entre les runs Illumina et Nanopore, par échantillon et par type de variant. (F) Distribution de la concordance de Jaccard. (G) Rappel des variants détectés dans le run/institut A (Besançon - Illumina) et retrouvés dans le run/institut B (Liège - Nanopore). (H) Précision des variants détectés dans le run/institut B (Liège - Nanopore) et confirmés dans le run/institut A (Besançon - Illumina).

4.9) Validation interne ONT MinION Mk1D

La figure 4.9 présente une comparaison interne des variants mitochondriaux détectés par séquençage Nanopore MinION Mk1D et par séquençage Illumina iSeq100 (référence) pour trois échantillons d'*Echinococcus* appartenant à la cohorte liégeoise. Deux échantillons d'*E. multilocularis* (ID3 et ID8) et un échantillon d'*E. granulosus* (ID18). Les échantillons retenus correspondent à ceux présentant les consensus les plus complets entre les deux technologies, afin de limiter les biais liés à la présence de nombreuses positions non résolues ou masquées. L'analyse porte sur les variants conservés après filtrage bioinformatique et compare séparément les SNPs, les INDELs et l'ensemble des variants détectés (ALL) pour les différents réplicats techniques.

Les SNPs constituent la majorité des événements observés, tandis que les insertions et délétions sont moins représentées (figures 4.9.A et 4.9.B). Parmi les substitutions, plusieurs changements nucléotidiques sont observés de manière récurrente, notamment T→C, G→A, C→T et A→G. Les variants sont répartis sur plusieurs régions de la séquence mitochondriale et ne sont pas regroupés dans une seule portion du mitogénome (figure 4.9.C). Les discordances observées entre les données Illumina et Nanopore sont présentées dans la table récapitulative (figure 4.9.D).

La concordance globale entre les deux technologies a été évaluée à l'aide de l'indice de Jaccard ainsi que des métriques de rappel et de précision. L'échantillon ID18 atteint une concordance de 100 % pour les SNPs et pour l'ensemble des variants (figures 4.9.E et 4.9.F). Les échantillons ID3 et ID8 présentent des valeurs légèrement inférieures, autour de 89–90 % pour les SNPs (figures 4.9.E et 4.9.F). Lorsque l'ensemble des variants est considéré, la concordance diminue pour ces deux échantillons, avec des valeurs proches de 85 % (figures 4.9.E et 4.9.F). Le rappel, défini ici comme la proportion de variants détectés par Illumina et retrouvés dans les données de Nanopore, atteint 100 % pour ID18 (figure 4.9.G). Pour ID3 et ID8, les valeurs de rappel restent élevées, mais diminuent légèrement lorsque l'ensemble des variants est pris en compte, avec des valeurs proches de 94–95 % pour ALL (figure 4.9.G). La précision, définie comme la proportion de variants détectés par Nanopore et confirmés par Illumina, atteint également 100 % pour ID18 (figure 4.9.H). Pour ID3 et ID8, les valeurs de précision sont comprises entre 89 % et 90 % pour les SNPs et pour l'ensemble des variants (figure 4.9.H).

4.10) Haplogroupes d'*E. multilocularis*

La figure 4.10 présente l'assignation phylogénétique et haplotypique des échantillons d'*E. multilocularis* à partir des consensus mitochondriaux complets (ID2, ID3, ID5, ID6, ID8, ID9, ID10, ID11, ID12, ID14, ID15, ID16, ID21, ID22, ID23). L'analyse comprend un arbre phylogénétique, une matrice de différences nucléotidiques moyennes avec les références de haplogroupes, ainsi qu'un réseau haplotypique. L'arbre phylogénétique présenté montre une séparation entre l'échantillon d'*E. granulosus* (ID18 granu Illumina) utilisé comme outgroup et l'ensemble des séquences d'*E. multilocularis* (figure 4.10.A). Les séquences cliniques d'*E. multilocularis* se positionnent majoritairement au sein de la diversité mitochondriale associée à l'haplogroupe 3 (HG3). Lorsque des consensus issus des deux technologies sont disponibles pour un même échantillon, ceux-ci se positionnent dans les mêmes branches ou dans des branches très proches de l'arbre phylogénétique, ce qui confirme leur appartenance au même haplogroupe (figure 4.10.A). La *HeatMap*, fondée sur la matrice des différences nucléotidiques moyennes entre les différents échantillons d'*E. multilocularis* et les références de différents haplogroupes indiquent que les plus faibles différences nucléotidiques moyennes sont observées avec les références appartenant à HG3, tandis que les différences nucléotidiques avec HG1 et HG2 sont plus élevées (figure 4.10.B). Au sein de l'haplogroupe 3, les échantillons ne présentent pas tous la même proximité avec les sous-haplogroupes HG3a, HG3b et HG3c (figures 4.10.A, 4.10.B et 4.10.C). Les échantillons ID5 et ID23 présentent les différences nucléotidiques moyennes les plus faibles avec HG3a (figure 4.10.B). Les échantillons ID6, ID9, ID12 et ID21 présentent des différences nucléotidiques moyennes plus faibles avec HG3b (figure 4.10.B). Un autre groupe d'échantillons, incluant notamment ID2, ID3, ID8, ID10, ID11, ID14, ID15, ID16 et ID22, présente les différences nucléotidiques moyennes les plus faibles avec HG3c (figure 4.10.B). Lorsque des consensus Illumina et Nanopore sont disponibles pour un même échantillon, les deux consensus présentent un placement comparable dans l'arbre phylogénétique et dans la matrice de distances (figures 4.10.A et 4.10.B). Les échantillons issus des cohortes belges et françaises ne se regroupent pas en un sous-haplogroupe unique. Ils sont répartis entre HG3a, HG3b et HG3c (figures 4.10.A, 4.10.B et 4.10.C). Le réseau haplotypique montre également une distribution des échantillons cliniques d'*E. multilocularis* dans plusieurs zones du réseau. Cette organisation correspond aux regroupements observés dans la matrice de distances (figure 4.10.C). Les consensus issus des séquençages Illumina et Nanopore pour les mêmes échantillons se situent dans les mêmes régions du réseau (figure 4.10.C).

Figure 4.10.

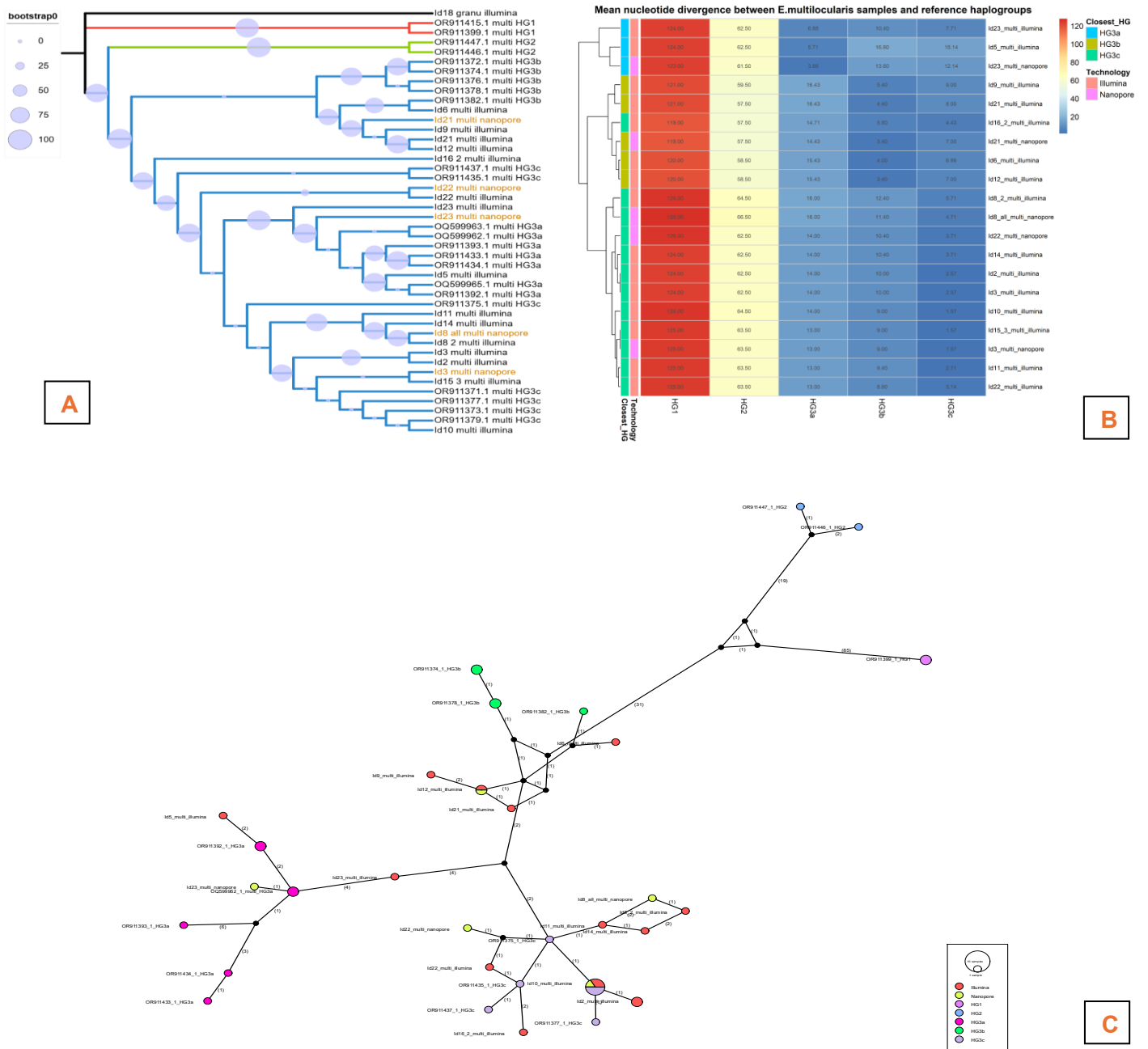


Figure 4.10) Validation de l'assignation des haplogroupes mitochondriaux d'*E. multilocularis* à partir de trois approches complémentaires.

(A) Arbre phylogénétique construit à partir des séquences mitochondriales complètes. (B) La HeatMap est fondée sur la divergence nucléotidique moyenne, permettant de visualiser les profils mutationnels partagés ou spécifiques entre les échantillons. (C) Réseau d'haplotypes illustrant les relations génétiques entre les échantillons analysés et les séquences de référence. Les analyses incluent des séquences de référence dont les accès commencent par OR ou OQ, ainsi que 15 échantillons cliniques d'*E. multilocularis* séquencés par Illumina. Parmi ceux-ci, les échantillons ID3, ID8, ID9, ID21, ID22 et ID23 ont également été séquencés par Nanopore. Un code couleur a été appliqué afin de distinguer les différents haplogroupes (arbre phylogénétique, réseau d'haplotypes) et la technologie (réseau d'haplotypes).

4.11) Génotypes d'*E. granulosus*

La figure 4.11 présente l'assignation génotypique des échantillons d'*E. granulosus* à partir des consensus mitochondriaux les plus complets possibles (ID18, ID19, ID20, ID24, ID25 et ID26).

L'analyse comprend un arbre phylogénétique, une matrice de divergence nucléotidique moyenne avec les références des génotypes 1 (G1) et 3 (G3), ainsi qu'un réseau haplotypique. Un échantillon d'*E. multilocularis* est utilisé comme groupe externe (ID2 *E. multilocularis* Illumina).

L'arbre phylogénétique montre une séparation entre l'échantillon d'*E. multilocularis* utilisé comme outgroup et l'ensemble des séquences d'*E. granulosus* (figure 4.11.A). Les échantillons se positionnent dans le clade associé au génotype G1 (figure 4.11.A).

Les séquences obtenues par Illumina et par Nanopore ne forment pas deux groupes distincts selon la technologie de séquençage. Les séquences correspondantes se placent dans des branches proches de l'arbre.

La *HeatMap* illustrant la divergence nucléotidique moyenne entre les échantillons belges d'*E. granulosus*, et les références des génotypes G1 et G3 présentent des valeurs de divergence plus faibles avec G1 qu'avec G3 (figure 4.11.B). Les valeurs les plus faibles vis-à-vis du génotype G1 sont observées pour ID20 Nanopore, ID24 Nanopore et ID19all Nanopore, avec des divergences comprises entre 14,5 et 16,5 bases (figure 4.11.B). Les échantillons ID18, ID25 et ID26 présentent des valeurs de divergence plus élevées avec le génotype G1, comprises entre 22,5 et 26,5 bases, mais restent plus proches du génotype G1 que du génotype G3 (figure 4.11.B).

Le réseau haplotypique montre également une séparation entre les références des génotypes G1 et G3 (figure 4.11.C). Les échantillons cliniques d'*E. granulosus* se situent dans la zone du réseau associée au génotype G1 (figure 4.11.C). Les consensus obtenus par Illumina et Nanopore occupent des positions situées dans la même région du réseau (figure 4.11.C).

Figure 4.11.

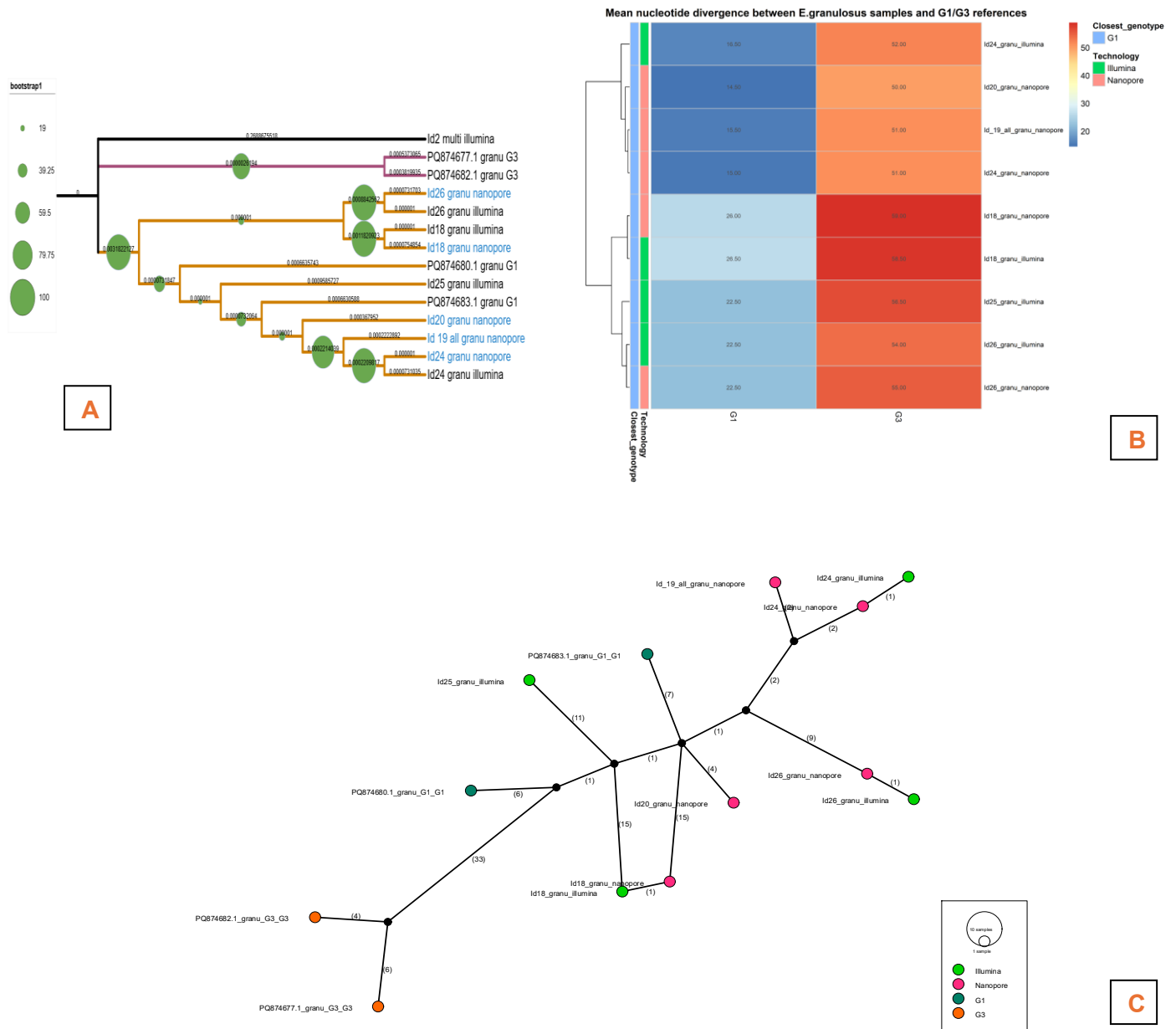


Figure 4.11) Validation de l'assignation des génotypes mitochondriaux d'*E. granulosus* à partir de trois approches complémentaires.

(A) Arbre phylogénétique construit à partir des séquences mitochondriales complètes ou quasi-complètes d'*E. granulosus*. (B) La HeatMap est fondée sur la divergence nucléotidique moyenne, ce qui permet de visualiser les profils mutationnels partagés ou spécifiques entre les échantillons. (C) Réseau d'haplotypes illustrant les relations génétiques entre les échantillons analysés et les séquences de référence. Les analyses incluent des séquences de référence dont les accès commencent par PQ, ainsi que 4 échantillons cliniques (ID18, ID24, ID25 et ID26) d'*E. granulosus* a été séquencé par Illumina et 5 échantillons (ID18, ID19, ID20, ID24 et ID26) ont également été séquencés par Nanopore. Un code couleur a été appliqué afin de distinguer les différents génotypes (arbre phylogénétique, réseau d'haplotypes) et la technologie (réseau d'haplotypes).

Chapitre V :

Discussion



CHAPITRE V : DISCUSSION

5.1) Optimisation de la méthode d'extraction d'ADN

L'extraction de l'ADN constitue une étape préliminaire essentielle au séquençage, tant avec la technologie Illumina qu'avec la technologie ONT. La qualité et la quantité de l'ADN extrait conditionnent en effet le succès des étapes ultérieures, notamment la préparation des librairies et l'obtention de données de séquençage exploitables. Dans cette optique, plusieurs méthodes d'extraction ont été évaluées afin d'obtenir un ADN présentant le meilleur rendement et la meilleure qualité possible. Une phase d'optimisation a ainsi été menée, incluant l'évaluation et la comparaison de différents kits d'extraction avant la sélection du protocole le plus adapté. Les échantillons analysés étaient constitués de tissus fortement calcifiés et nécrosés, conservés à -80 °C depuis plusieurs années, ce qui représentait une contrainte importante pour l'extraction d'ADN parasitaire. Dans ces conditions, l'utilisation du High Pure PCR Template Preparation Kit (Roche) et du QIAamp Fast DNA Tissue Kit (Qiagen, 501404), selon les protocoles standards des fabricants, n'a pas permis d'obtenir des résultats satisfaisants, malgré des essais réalisés indépendamment par trois opérateurs différents. L'extraction nous a donné des rendements proches de 0 ng/ul d'ADN selon les dosages au Qubit et au Nanodrop.

Une hypothèse pouvant expliquer ces échecs est la dégradation rapide de l'ADN après décongélation, possiblement favorisée par la réactivation de DNases dans les tissus nécrosés^{58,59}. Les DNases sont des enzymes nucléases capables de dégrader l'ADN en hydrolysant les liaisons phosphodiester entre les nucléotides. Pour cette raison, les échantillons ont dû être traités le plus rapidement possible après leur sortie du congélateur, afin de réduire au maximum le temps d'exposition de l'ADN aux DNases potentiellement actives^{58,59}.

Face à ces difficultés, le protocole du kit Roche a été adapté. Les principales modifications ont consisté en l'ajout d'une étape de *bead beating* d'une minute à $6 \text{ m} \cdot \text{s}^{-1}$, ainsi qu'en une réduction du temps d'exposition de l'échantillon à la protéinase K à 10 ou 15 minutes au lieu de 2 heures. La méthode d'extraction d'ADN ayant permis d'obtenir des résultats exploitables est celle décrite au chapitre III. Cette adaptation a permis d'obtenir des concentrations d'ADN quantifiables, variant généralement de 0 à environ 20 ng/ μL selon les mesures effectuées avec le Qubit et le NanoDrop. Cependant, malgré cette optimisation, certains échantillons sont restés difficilement exploitables. Bien que des concentrations d'ADN faibles aient parfois permis une

amplification par PCR et un séquençage satisfaisant, d'autres échantillons présentant des concentrations similaires n'ont pas généré de signal exploitable lors des étapes ultérieures. Ces observations suggèrent que la réussite de l'analyse ne dépend pas uniquement de la quantité d'ADN extraite, mais également d'autres facteurs tels que son intégrité, sa pureté ou la présence éventuelle d'inhibiteurs.

5.2) Validation des contrôles de l'Illumina iSeq100 et du Nanopore MinION Mk1D

Les résultats obtenus par séquençage sur les plateformes Illumina iSeq100 et Nanopore MinION Mk1D peuvent être considérés comme exploitables, notamment au regard du comportement conforme des contrôles négatifs. Ces derniers étaient constitués d'eau ayant suivi l'intégralité du protocole, et seul le deuxième *run* Illumina contenait un contrôle d'eau qui n'avait pas suivi l'intégralité du protocole (**voir le chapitre 3 : Matériel et méthode**). Dans les runs Illumina et Nanopore réalisés, les contrôles négatifs n'ont pas permis de générer un consensus mitochondrial fiable. Ils présentaient une profondeur de séquençage moyenne très faible, inférieure à 10 *reads* par position validée, ainsi qu'un consensus contenant plus de 99 % de bases non résolues.

Ces contrôles ne sont toutefois pas totalement exempts de signal, puisque quelques *reads* ont été détectés lors du séquençage. Cependant, leur nombre reste très limité et insuffisant pour produire un consensus exploitable ou des variants retenus après filtrage bioinformatique. Ce profil est compatible avec un bruit de fond très faible ou une contamination résiduelle mineure, mais ne remet pas en cause l'interprétation des échantillons cliniques, qui présentent des profondeurs, des couvertures et des consensus nettement plus robustes.

5.3) Analyse des métriques de séquençage : Illumina vs Nanopore

Les FASTQ bruts obtenus dans ce travail sont cohérents avec les différences généralement décrites dans la littérature entre les technologies Illumina et Nanopore^{42,43}. L'annexe 2, qui contient les fichiers MultiQC, montre que les *reads* générés par le séquenceur Illumina iSeq100 présentent un score de qualité moyen nettement supérieur à celui des *reads* obtenus avec le Nanopore MinION Mk1D. Cette différence de qualité a été prise en compte lors du traitement bioinformatique des données, en appliquant des seuils de filtrage adaptés à chaque technologie. En effet, un seuil minimal de qualité Q20 a été appliqué aux *reads* Illumina, afin de ne conserver que les bases présentant une forte fiabilité. À l'inverse, des seuils de Q10 ou Q15 ont été appliqués aux données Nanopore afin d'atteindre un compromis entre la conservation d'un nombre suffisant de séquences et l'élimination des *reads* de qualité insuffisante.

En revanche, la longueur des séquences d'ADN diffère selon les deux approches. Les *reads* Illumina présentent une distribution centrée autour de 150 pb, ce qui correspond au format attendu pour cette technologie de séquençage. À l'inverse, les *reads* Nanopore présentent des longueurs moyennes supérieures, souvent au-delà de 1000 pb dans nos données, ce qui constitue l'un des principaux avantages de cette technologie. Cette capacité à produire des *reads* plus longs est particulièrement intéressante pour l'analyse de fragments mitochondriaux amplifiés, car elle peut faciliter la couverture des régions répétitives et améliorer la continuité de l'information obtenue.

Il faut toutefois tenir compte du type de préparation de librairie utilisé. Dans ce travail, les librairies Nanopore ont été préparées à l'aide du *Rapid Barcoding Kit*. Ce kit repose sur l'utilisation d'un complexe transposase qui fragmente l'ADN tout en permettant l'ajout rapide des adaptateurs et des barcodes (codes-barres) nécessaires à l'identification des échantillons. Cette approche présente l'avantage d'être rapide, facilement adaptable et compatible avec des quantités d'ADN plus faibles, ce qui est particulièrement utile dans un contexte de validation méthodologique en laboratoire, avec des extraits d'ADN en faible quantité. En contrepartie, la fragmentation induite par la transposase peut limiter la longueur maximale des *reads* obtenus par rapport à des protocoles plus conservateurs, comme le Native Barcoding Kit, souvent utilisé lorsque l'objectif est de préserver des fragments d'ADN plus longs, mais avec un coût par kit plus élevé, une préparation des librairies plus longue et fastidieuse, et des quantités d'ADN d'entrée dans la librairie de séquençage beaucoup plus grandes.

Enfin, le séquençage sur Nanopore MinION Mk1D présente un avantage important par rapport au séquençage sur Illumina iSeq100 en termes de flexibilité et de volume de données générées. En effet, le séquenceur iSeq100 est une plateforme de séquençage à petite échelle, avec une capacité maximale d'environ 1,2 Gb (gigabases) par run et une durée maximale de run fixée à 22 h dans les conditions utilisées. Cette configuration est bien adaptée à des analyses ciblées ou à des génomes de petite taille, mais elle limite la quantité totale de données produites lors d'un même séquençage. Elle ne permet donc pas d'obtenir une couverture suffisante pour l'analyse de plus grands génomes nucléaires.

À l'inverse, le séquençage ONT permet une production de données plus flexible, dépendante de la qualité de la *Flowcell*, de la librairie chargée et de la durée effective du run⁴⁶. Dans de bonnes conditions, un run Nanopore peut générer plusieurs Gb de données, voire 10 à 15 Gb selon les performances de la *Flowcell*⁴⁶. La génération des données s'effectue en temps réel, ce qui permet de suivre le nombre de séquences produites au cours du run et d'adapter la durée du séquençage en fonction des besoins expérimentaux. Cette possibilité d'interrompre ou de prolonger le run constitue un avantage pratique important, notamment lorsqu'une profondeur suffisante est atteinte rapidement ou, au contraire, lorsqu'un temps de séquençage

plus long est nécessaire pour augmenter la quantité de données disponibles. Cette flexibilité rend également la technologie ONT plus adaptée à l'analyse de cibles génomiques plus longues (par exemple des petits génomes parasites comme le *Plasmodium*)⁴⁷.

5.4) Validation méthodologique du séquençage Nanopore et Illumina

Afin d'évaluer la reproductibilité du séquençage Illumina sur le séquenceur iSeq100, plusieurs réplicats techniques ont été réalisés pour vérifier si les mêmes variants mitochondriaux étaient retrouvés entre les différents réplicats et si le positionnement phylogénétique des échantillons restait stable au sein des haplogroupes *d'E. multilocularis* (figures 4.7 et 4.10). L'ensemble des échantillons comparés présente un score de concordance de Jaccard supérieur à 95 % pour les SNPs (figure 4.7). De plus, les réplicats techniques des échantillons ID16, ID8, ID15 et ID9 conservent un placement identique dans l'arbre phylogénétique, sur la HeatMap construite à partir de la matrice de différences nucléotidiques moyennes, ainsi que dans le réseau haplotypique (figure 4.10). Ces résultats soutiennent donc la reproductibilité technique de la méthodologie Illumina utilisée dans ce travail.

L'évaluation du séquençage Nanopore a été réalisée selon deux approches complémentaires. Une validation interne a d'abord été menée sur des échantillons issus de la cohorte belge, en comparant les résultats obtenus avec la plateforme Nanopore MinION MK1D à ceux du séquençage Illumina iSeq100. Une validation externe a ensuite été effectuée à partir d'échantillons de la cohorte française fournis par le CNR-E du CHU de Besançon, les données obtenues par Nanopore étant comparées à des données de référence acquises sur un séquenceur Illumina NextSeq 500 et générées par l'équipe française (figures 4.8 et 4.9). Afin de réaliser la comparaison interne dans les meilleures conditions, seuls les échantillons présentant les consensus les plus complets et le plus grand nombre de bases résolues ont été retenus. Cette validation interne montre une concordance des SNPs comprise entre 88 % et 100 % (figure 4.9), ainsi qu'un positionnement identique des réplicats techniques des échantillons ID8 et ID3 au sein des haplo-groupes *d'E. multilocularis* et l'échantillon ID18 dans les génotypes *d'E. granulosus* (figures 4.10 et 4.11). Ces observations indiquent que les consensus Nanopore obtenus au sein du service sont globalement cohérents avec ceux générés par la plateforme Illumina iSeq100.

Enfin, la validation externe réalisée à partir des échantillons français montre que la concordance de Jaccard des SNPs est comprise entre 84 % et 100 % pour les échantillons ID21 à ID26 (figure 4.8). Certains échantillons, notamment ID25 et ID26, présentent des valeurs de concordance plus faibles, en lien avec des différences de complétude entre les consensus comparés. Pour ID25, le consensus Nanopore est moins complet que celui des autres échantillons, avec 5 772 bases non résolues, dont environ 1 500 en dehors de la région répétitive. À

l'inverse, ID26 présente une meilleure récupération de la région répétitive sur Nanopore, avec environ 2 000 bases supplémentaires détectées dans cette région par rapport aux données d'Illumina. Ces différences dans les portions génomiques accessibles entre les deux technologies peuvent effectivement diminuer mécaniquement la concordance de Jaccard, sans nécessairement traduire une discordance majeure concernant les variants présents dans les régions comparables (**figure 4.8**).

L'analyse des graphiques de rappel et de précision permet ainsi de nuancer l'interprétation des valeurs de Jaccard. Les différences observées entre ID25 et ID26 semblent principalement liées à des régions manquantes ou, au contraire, mieux couvertes par la technologie Nanopore, plutôt qu'à une divergence globale des variants détectés entre les deux technologies (**figure 4.8**). En parallèle, ces échantillons conservent un positionnement cohérent dans l'arbre phylogénétique, la matrice de divergence nucléotidique et le réseau haplotypique. Les haplogroupes ou génotypes attribués restent donc concordants entre les technologies Illumina et Nanopore, ce qui soutient la robustesse globale de l'approche Nanopore pour l'assignation phylogénétique, malgré certaines limites liées à la variabilité de la complétude des consensus (**figures 4.10 et 4.11**).

5.5) Diversité génétique d'*E. multilocularis* et d'*E. granulosus* en Belgique

Dans la littérature sur le mitogénome d'*E. multilocularis*, l'haplogroupe majoritairement rapporté en Europe centrale et en Europe de l'Ouest est l'haplogroupe 3, comprenant plusieurs sous-haplogroupes, notamment HG3a, HG3b et HG3c^{29,30}. Dans notre étude, l'ensemble des échantillons belges analysés appartient à l'haplogroupe 3 (**figure 4.10**). Toutefois, ces échantillons ne se concentrent pas dans un seul sous-haplogroupe : ils se répartissent entre les trois sous-groupes identifiés : HG3a, HG3b et HG3c. HG3a est un sous-haplogroupe transatlantique qui contient l'échantillon ID5. HG3b est un sous-haplogroupe majoritairement présent en Europe centrale et regroupe trois échantillons (ID6, ID9 et ID12). HG3c est un sous-haplogroupe qui se disperse sur toute l'Europe et regroupe huit échantillons (ID2, ID3, ID8, ID10, ID11, ID14, ID15 et ID16) (**figure 4.10**). Les échantillons français analysés se répartissent également dans ces trois sous-haplogroupes : l'ID21 dans le sous-haplogroupe HG3b, l'ID22 dans le sous-haplogroupe HG3c et l'ID23 dans le sous-haplogroupe HG3a (**figure 4.10**).

Enfin, l'ensemble des échantillons d'*E. granulosus* analysé se place dans le génotype G1. Cette assignation est cohérente avec les trois approches utilisées : le placement des échantillons dans l'arbre phylogénétique, leur divergence nucléotidique la plus faible par rapport aux références du génotype G1 et leur regroupement avec ces mêmes références dans le réseau haplotypique (**figure 4.11**).

5.6) Comparaison du typage EmsB et du séquençage du mitogénome complet

La comparaison entre les deux marqueurs a été réalisée sur trois échantillons d'*E. multilocularis*. Ces échantillons ont été typés à la fois par EmsB et par séquençage du mitogénome complet. Il s'agit des échantillons ID2, ID5 et ID6⁶⁰. L'échantillon ID5 associe le profil EmsB P4 au sous-haplogroupe HG3a. L'échantillon ID6 associe le profil P1 au sous-haplogroupe HG3b. L'échantillon ID2 associe le profil P8 au sous-haplogroupe HG3c⁶⁰. Les résultats ne montrent pas de correspondance directe entre les profils nucléaires EmsB et les haplogroupes mitochondriaux car les deux marqueurs n'analysent pas le même compartiment génomique. L'EmsB est un microsatellite nucléaire multilocus. Il est transmis par les deux parents et peut subir une recombinaison. À l'inverse, le mitogénome est haploïde et transmis par la lignée maternelle. Il ne retrace donc que l'histoire maternelle de l'isolat. Ainsi, les marqueurs nucléaires et mitochondriaux peuvent fournir des signaux distincts. Ce type de discordance a déjà été décrit chez *E. multilocularis*. Il peut refléter des événements de fécondation croisée entre des génotypes d'origines différentes⁶¹. Pour finir, le mitogénome complet fournit une séquence nucléotidique. Il permet donc une analyse phylogénétique et fournit un signal phylogéographique permettant l'assignation à des clades continentaux. L'EmsB repose sur des profils de tailles de pics ; il est donc très utile pour regrouper les isolats selon leur similarité. Il permet également d'étudier la transmission à l'échelle régionale. En revanche, il ne fournit aucune information sur la séquence. Il ne permet donc ni la reconstruction phylogénétique globale de l'espèce ni l'assignation d'haplogroupes géographiques continentaux.

Chapitre VI :
Conclusions
et Perspectives



CHAPITRE VI : CONCLUSIONS ET PERSPECTIVES

Ce travail constitue la première mise en place du séquençage du mitogénome complet dans notre service de Microbiologie du CHU de Liège.

Le premier objectif était d'identifier les haplogroupes *d'E. multilocularis* présents en Belgique, en particulier les sous-haplogroupes HG3a, HG3b et HG3c. Le second objectif était de comparer deux technologies de séquençage : l'Illumina iSeq100 et le Nanopore MinION Mk1D. Pour les réplicats techniques *d'E. granulosus* et *d'E. multilocularis*, les deux plateformes ont donné les mêmes résultats. Les mêmes haplogroupes, sous-haplogroupes et génotypes ont été observés. Cela montre que les résultats sont concordants, quelle que soit la technologie utilisée. Cette concordance a permis une première validation croisée des deux approches au sein du laboratoire. Le séquençage Nanopore MinION Mk1D a montré une robustesse satisfaisante. Les métriques de concordance des variants étaient élevées par rapport aux résultats obtenus avec notre séquenceur Illumina iSeq100 et avec un autre séquenceur externe, l'Illumina NextSeq 500. De son côté, le séquençage Illumina iSeq100 a également montré une bonne fiabilité, avec des métriques de concordance solides entre les différents réplicats techniques. Cependant, notre service de Microbiologie du CHU de Liège a tranché en faveur de la technologie ONT comme technique de séquençage pour les futures études, en raison de son faible coût par rapport à la technologie d'Illumina et de son adaptabilité à un panel plus large de génomes.

Au-delà de cette validation méthodologique, ce travail rapporte, à notre connaissance, le premier séquençage du mitogénome complet *d'E. granulosus* et *d'E. multilocularis*, réalisés sur une plateforme ONT.

Les principales limites de l'étude ont été le nombre restreint d'échantillons et le faible rendement en ADN après extraction.

Pour conclure ce travail, la prochaine étape du projet consistera donc à séquencer le mitogénome complet *d'E. multilocularis* chez ses principaux hôtes définitifs en Belgique, notamment le renard roux (*Vulpes vulpes*). Cela permettra de mieux structurer les sous-haplogroupes de l'haplogroupe HG3. Cela aidera également à mieux comprendre la dispersion du parasite sur le territoire belge.



Chapitre VII :

Bibliographie



CHAPITRE VII : BIBLIOGRAPHIE

1. Vuitton, D. A. *et al.* A historical view of alveolar echinococcosis, 160 years after the discovery of the first case in humans: part 1. What have we learnt on the distribution of the disease and on its parasitic agent? *Chinese Medical Journal* **124**, 2943 (2011).
2. Tappe, D. & Frosch, M. Rudolf Virchow and the Recognition of Alveolar Echinococcosis, 1850s. *Emerg Infect Dis* **13**, 732–735 (2007).
3. Romig, T. *et al.* Rediscovery of a name-bearing type of *Echinococcus multilocularis* (Leuckart, 1863) by museum forensics: a cold case revisited. *Int J Parasitol* **55**, 649–655 (2025).
4. Auer, H. & Aspöck, H. Incidence, prevalence and geographic distribution of human alveolar echinococcosis in Austria from 1854 to 1990. *Parasitol Res* **77**, 430–436 (1991).
5. Yao BL, Huang NQ, Fu LM, Fang KH. Hepatic alveolar echinococcosis. *Chin Surg J* (Chin) 1965; 13: 460-461
6. Schantz PM, Chai J, Craig PS, Eckert J, Jenkins DJ, Macpherson CNL, *et al.* Epidemiology and control of hydatid disease. In: Thompson RCA, Lymbery AJ, eds. *Echinococcus and hydatid disease*. Wallingford: Cab International; 1995: 233-302.
7. Knapp J, Bart JM, Maillard S, Gottstein B, Piarroux R. The genomic *Echinococcus* microsatellite EmsB sequences: from a molecular marker to the epidemiological tool. *Parasitology* 2010; 137: 439-449.
8. Rostami, A. *et al.* Human Alveolar Echinococcosis—A Neglected Zoonotic Disease Requiring Urgent Attention. *International Journal of Molecular Sciences* **26**, 2784 (2025).
9. *Echinococcus multilocularis* - Learn About Parasites - Western College of Veterinary Medicine. [wcvm-learnaboutparasites https://wcvm.usask.ca/learnaboutparasites/parasites/echinococcus-multilocularis.php](https://wcvm.usask.ca/learnaboutparasites/parasites/echinococcus-multilocularis.php) (2021). Consulté en 2026.
10. CDC - DPDx - Echinococcosis. <https://www.cdc.gov/dpdx/echinococcosis/index.html> (2019). Consulté en 2026.
11. Simoncini, A. & Massolo, A. Multiscale ecological drivers of *Echinococcus multilocularis* spatial distribution in wild hosts: A systematic review. *Food Waterborne Parasitol* **34**, e00216 (2024).
12. Moro, P. & Schantz, P. M. Echinococcosis: a review. *International Journal of Infectious Diseases* **13**, 125–133 (2009).
13. Lallemand, S. Caractérisation phylogénomique de la diversité génétique d'*Echinococcus multilocularis*. : application à l'étude de patients français atteints d'échinococcose alvéolaire, <https://theses.hal.science/tel-05050886v1> (2025).
14. Biology and Systematics of *Echinococcus*. in *Advances in Parasitology* vol. 95 65–109 (Academic Press, 2017).
15. Federer, K., Armua-Fernandez, M. T., Hoby, S., Wenker, C. & Deplazes, P. In vivo viability of *Echinococcus multilocularis* eggs in a rodent model after different thermo-treatments. *Exp Parasitol* **154**, 14–19 (2015).
16. Kern, P. *et al.* WHO classification of alveolar echinococcosis: Principles and application. *Parasitology International* **55**, S283–S287 (2006).
17. Agarwal, D., Gadwal, S. K., Aswani, Y. & Das, C. J. Role of interventional radiology in the management of hepatic hydatid disease. *World J Radiol* **18**, 119319 (2026).
18. Frey, H. & Barth, T. F. E. Macroscopic and Microscopic Aspects of *Echinococcus multilocularis* and *Echinococcus granulosus* in Humans: A Guideline for Pathologists. *Visc Med* <https://doi.org/10.1159/000548706> (2025) doi:[10.1159/000548706](https://doi.org/10.1159/000548706).
19. Bi, X. *et al.* Host angiogenic reprogramming by *Echinococcus multilocularis* protoscoleces protein via PDGFR/PI3K/AKT cascade. *Front Microbiol* **16**, 1686956 (2025).

20. Rachel, F., Conraths, F. J. & Maksimov, P. Genetic diversity and genotyping of *Echinococcus multilocularis*: a minireview. *Front. Parasitol.* **4**, (2025).
21. Vogel, M. *et al.* *Echinococcus multilocularis*: Characterization of a DNA probe. *Acta Tropica* **48**, 109–116 (1990).
22. Bowles, J., Blair, D. & McManus, D. P. Genetic variants within the genus *Echinococcus* identified by mitochondrial DNA sequencing. *Molecular and Biochemical Parasitology* **54**, 165–173 (1992).
23. Bowles, J. & McManus, D. P. Molecular variation in *Echinococcus*. *Acta Tropica* **53**, 291–305 (1993).
24. Bart, J. M. *et al.* EmsB, a tandem repeated multi-loci microsatellite, new tool to investigate the genetic diversity of *Echinococcus multilocularis*. *Infection, Genetics and Evolution* **6**, 390–400 (2006).
25. Valot, B., Knapp, J., Umhang, G., Grenouillet, F. & Millon, L. Genomic characterization of EmsB microsatellite loci in *Echinococcus multilocularis*. *Infection, Genetics and Evolution* **32**, 338–341 (2015).
26. Knapp, J. *et al.* Genotyping *Echinococcus multilocularis* in Human Alveolar Echinococcosis Patients: An EmsB Microsatellite Analysis. *Pathogens* **9**, (2020).
27. Sacheli, R. *et al.* Genetic diversity of *Echinococcus multilocularis* specimens isolated from Belgian patients with alveolar echinococcosis using EmsB microsatellites analysis. *Infection, Genetics and Evolution* **116**, 105531 (2023).
28. Rachel, F. *et al.* Typing of *Echinococcus multilocularis* by Region-Specific Extraction and Next-Generation Sequencing of the mitogenome. *Front. Microbiol.* **16**, (2025).
29. Lallemand, S. *et al.* Challenging the phylogenetic relationships among *Echinococcus multilocularis* isolates from main endemic areas. *International Journal for Parasitology* **54**, 569–582 (2024).
30. Bohard, L. *et al.* Complete mitochondrial exploration of *Echinococcus multilocularis* from French alveolar echinococcosis patients. *Int J Parasitol* **53**, 555–564 (2023).
31. Guo, B. *et al.* The genetic variation of mitochondrial sequences and pathological differences of *Echinococcus multilocularis* strains from different continents. *Microbiology Spectrum* <https://doi.org/10.1128/spectrum.01318-24> (2025) doi:[10.1128/spectrum.01318-24](https://doi.org/10.1128/spectrum.01318-24).
32. Wen, H. *et al.* Echinococcosis: Advances in the 21st Century. *Clinical Microbiology Reviews* **32**, 10.1128/cmr.00075-18 (2019).
33. Echinococcosis – ECDC Annual Zoonoses Report for 2023. [The-European-Union-One-Health-2023-Zoonoses-report.PDF](https://ecdc.europa.eu/en/health-topics/zoonoses/zoonoses-report-2023)(2023). Consulté en 2026.
34. Bresson-HADNi, S. *et al.* Échinococcose alvéolaire. EMC - Hépatogastroentérologie **2**, 86–104 (2005).
35. Pal, M., Alemu, H., Megersa, L., Garedo, D. & Bodena, E. Cystic Echinococcosis: A Comprehensive Review on Life Cycle, Epidemiology, Pathogenesis, Clinical Spectrum, Diagnosis, Public Health and Economic Implications, Treatment, and Control. *International Journal of Clinical and Experimental Medicine Research* **6**, (2022).
36. Issa, A., Al-berfkani, M., Swar, S., Mohammed, W. & Arif, S. *Animal Health Perspectives Unique Scientific Publishers CYSTIC ECHINOCOCCOSIS.* (2022). doi:[10.47278/book.ahp/2022.17](https://doi.org/10.47278/book.ahp/2022.17).
37. Bhalla, V. P., Paul, S. & Klar, E. Hydatid Disease of the Liver. *Visc Med* **39**, 112–120 (2023).
38. Liu, H. *et al.* Advances in Novel Diagnostic Techniques for Alveolar Echinococcosis. *Diagnostics (Basel)* **15**, 585 (2025).
39. Li, R. *et al.* Circulating microRNA signatures for echinococcosis. *Parasit Vectors* <https://doi.org/10.1186/s13071-025-07161-8> (2025) doi:[10.1186/s13071-025-07161-8](https://doi.org/10.1186/s13071-025-07161-8).
40. EGREK, S. & HAYETTE, M.-P. *Rapport d'activité 2024.* Laboratoire National de Référence des échinococcoses (LNRE), Centre Hospitalier Universitaire (CHU) de Liège, Service de Microbiologie Clinique (2025).
41. Eckert, J., Gemmell, M. A., Meslin, F.-X. & Pawłowski, Z. S. (eds) *WHO/OIE Manual on Echinococcosis in Humans and Animals: a Public Health Problem of Global Concern* (World Organisation for Animal Health & World Health Organization, 2001).
42. Akintunde, O., Tucker, T. & Carabetta, V. J. The Evolution of Next-Generation Sequencing Technologies. in *High Throughput Gene Screening: Methods and Protocols* (eds Carabetta, V. J. & Akintunde, O.) 3–29 (Springer US, New York, NY, 2025). doi:[10.1007/978-1-0716-4192-7_1](https://doi.org/10.1007/978-1-0716-4192-7_1).
43. Satam, H. *et al.* Next-Generation Sequencing Technology: Current Trends and Advancements. *Biology* **12**, 997 (2023).

44. 9300 *Echinococcus multilocularis*, Bordier Affinity Products. [http://www.bordier.ch/9300%20Echinococcus%20multilocularis%20\(Em2-Em18\)/index.htm](http://www.bordier.ch/9300%20Echinococcus%20multilocularis%20(Em2-Em18)/index.htm) (2026). Consulté en 2026
45. Trachsel, D., Deplazes, P. & Mathis, A. Identification of taeniid eggs in the faeces from carnivores based on multiplex PCR using targets in mitochondrial DNA. *Parasitology* **134**, 911–920 (2007).
46. Ni, Y., Liu, X., Simeneh, Z. M., Yang, M. & Li, R. Benchmarking of Nanopore R10.4 and R9.4.1 Flowcells in single-cell whole-genome amplification and whole-genome shotgun sequencing. *Comput Struct Biotechnol J* **21**, 2352–2364 (2023).
47. Tan, M. K. I. et al. Enhancing whole genome DNA amplification of *Plasmodium falciparum* for advanced molecular surveillance in malaria control. *Malar J* **24**, 445 (2025).
48. Umhang, G. Surveillance et épidémiologie d'*Echinococcus multilocularis* et d'*Echinococcus granulosus sensu lato*. (Université Paris-Est, 2017).
49. Guo, B. et al. High endemicity of alveolar echinococcosis in Yili Prefecture, Xinjiang Autonomous Region, the People's Republic of China: Infection status in different ethnic communities and in small mammals. *PLoS Negl Trop Dis* **15**, e0008891 (2021).
50. Bresson-HADNi, S. et al. Échinococcose alvéolaire. *EMC - Hépatogastroentérologie* **2**, 86–104 (2005).
51. Parija, S. C. & Pramodhini, S. Echinococcosis. in *Textbook of Parasitic Zoonoses* (eds Parija, S. C. & Chaudhury, A.) 353–368 (Springer Nature, Singapore, 2022). doi:10.1007/978-981-16-7204-0_33.
52. Testard, Q. Industrialisation des procédures d'analyses de données de séquençage pan-génomiques constitutionnelles. HAL Id : tel-03664086 , version 1.
53. Nanopore store: Rapid Barcoding Kit 24 V14. <https://store.Nanoporetech.com/us/rapid-barcoding-sequencing-kit-24-v14.html>. Consulté en 2026.
54. DeepChek® NGS Library Preparation Kit - ABL Diagnostics S.A. <https://www.abldiagnostics.com/products/deepchek-ngs-library-preparation-kit/>. Consulté en 2026.
55. Bowles, J., Blair, D. & McManus, D. P. A molecular phylogeny of the genus *Echinococcus*. *Parasitology* **110 (Pt 3)**, 317–328 (1995).
56. Kantarci M, Bayraktutan U, Karabulut N, et al. (2012). Alveolar Echinococcosis: Spectrum of Findings at Cross-sectional Imaging. *Radiographics*, **32**, 2053–2070. Disponible sur : <http://pubs.rsna.org/doi/10.1148/rq.327125708>
57. Pavlidis, E. T., Galanis, I. N. & Pavlidis, T. E. Current considerations for the management of liver echinococcosis. *World J Gastroenterol* **31**, 103973 (2025).
58. Mizuta, R. et al. DNase γ Is the Effector Endonuclease for Internucleosomal DNA Fragmentation in Necrosis. *PLOS ONE* **8**, e80223 (2013).
59. Messner, E. et al. Perish the thawed? EDTA reduces DNA degradation during extraction from frozen tissue. *PLoS One* **20**, e0321872 (2025).
60. Egrek, S. et al. EmsB Microsatellite Analysis of *Echinococcus multilocularis* Specimens Isolated from Belgian Patients with Alveolar Echinococcosis and from Animal Hosts. *Pathogens* **14**, 584 (2025).
61. Umhang, G. et al. Asian Admixture in European *Echinococcus multilocularis* Populations: New Data From Poland Comparing EmsB Microsatellite Analyses and Mitochondrial Sequencing. *Front. Vet. Sci.* **7**, (2021).



ANNEXE

A) Annexe 1 :

ID Patient	ID sample	Sexe	Age	Code postal	Localisation primaire	Localisation(s) secondaire(s) ?	Symptomatique	Type de prélèvement	Pays de contamination
E. multilocularis									
1	1	F	64	6852	Foie	NON	OUI	Biopsie foie segment IV-V	Belgique
2	2	F	69	6831	Foie	NON	OUI	Biopsie hépatique	Belgique
3	3	M	32	4750	Foie	Poumon	OUI	Biopsie hépatique	Belgique
4	4	M	69	6600	Foie	NON	OUI	Nodule hépatique	Belgique
5	5	F	42	1640	Foie	NON	NON	Biopsie foie	Belgique
6	6	F	51	6640	Foie	NON	OUI	Biopsie foie segment VII et VIII	Belgique
6	7							Biopsie foie segment VI	Belgique
7	8	M	13	6740	Foie	NON	OUI	Biopsie foie segment IV	Belgique
7	9							Biopsie foie segment VIII	Belgique
8	10	M	62	4141	Foie	Poumon	OUI	Biopsie hépatique	Belgique
9	11	M	33	4560	Foie	Poumon	OUI	Biopsie pulmonaire	Belgique
9	12							Kyste poumon	Belgique
10	13	M	68	6860	Foie	NON	OUI	Biopsie hépatique segment VII et VIII	Belgique
10	14							Biopsie hépatique	Belgique
11	15	M	33	6600	Foie	NON	NON	Biopsie hépatique segment IV	Belgique
12	16	F	72	4020	Foie	NON	NON	Biopsie foie segment IV	Belgique
E. granulosus									
13	17	M	43	4020	Foie	NON	OUI	Biopsie hépatique	Maroc
14	18	M	52	5000	Foie	NON	OUI	Kyste rein	Maroc
15	19	F	24	4800	Foie	Poumon	OUI	Liquide (origine non spécifiée)	Kosovo
16	20	F	36	1020	Foie	NON	OUI	Biopsie hépatique	Turquie
E. multilocularis Besançon									
17	21	x	x	x	x	x	x	x	x
18	22	x	x	x	x	x	x	x	x
19	23	x	x	x	x	x	x	x	x
E. granulosus Besançon									
20	24	x	x	x	x	x	x	x	x
21	25	x	x	x	x	x	x	x	x
22	26	x	x	x	x	x	x	x	x

Annexe 1 : Tableau reprenant les données cliniques des patients belges infectés par l'*E. multilocularis* (ID 1 à ID12) et *E. granulosus* (ID13 à ID16). Les échantillons de la cohorte française (CNR-E, CHU de Besançon) sont également inclus mais les données cliniques sont manquantes (ID17 à ID22).

B) Annexe 2 :

L'ensemble des sources et pipelines de code se trouve dans ce répertoire : <https://doi.org/10.5281/zenodo.20689406>

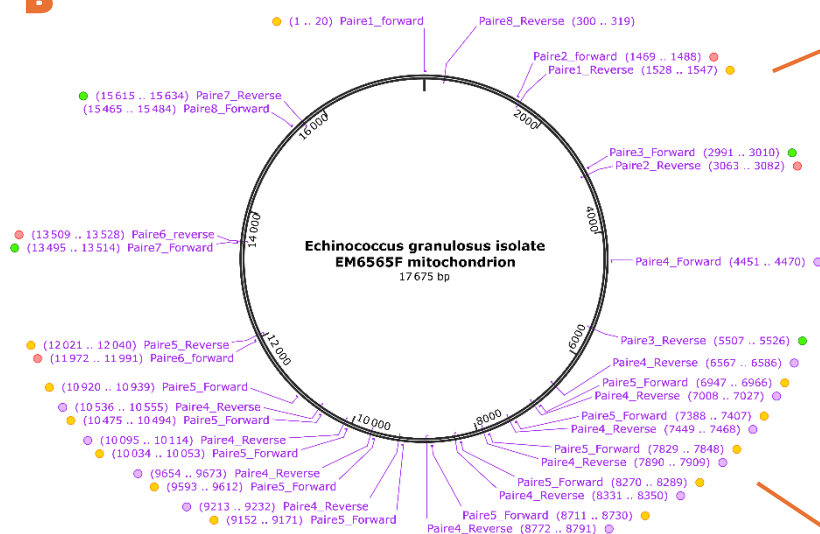
Celui-ci contient les fichiers FASTQC et MultiQC de l'ensemble des séquençages réalisés et analysés, un exemple de pipeline dédié au séquençage Illumina et Nanopore, ainsi que le code pour la construction d'arbres phylogénétiques.

C) Annexe 3 :

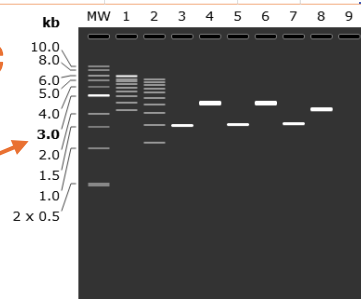
A

Description	Sequence	Nucleotide_Number	GC_content	Weight(Da)	A	T	C	G	Extinction coefficient(l/(mol-cm))	Molecular weight(g/mol)	nmol	µg/OD260
Paire1_forward	ATGAGAGTGGTGTGATTAGG	20-mer	45% GC	6292 5.0	6.0	0.0	9.0	209800.0	6292.2	4.8	30.0	
Paire1_Reverse	GTATAATCCACGCTATACGG	20-mer	45% GC	6101 6.0	5.0	5.0	4.0	198300.0	6101.0	5.0	30.8	
Paire2_forward	CTGGTTTAGTGGATTGTTTG	20-mer	40% GC	6200 2.0	10.0	1.0	7.0	188000.0	6200.1	5.3	33.0	
Paire2_Reverse	CTAATAGAATCGCTCACTG	20-mer	40% GC	6085 7.0	5.0	5.0	3.0	195900.0	6085.0	5.1	31.1	
Paire3_Forward	GATGACAGGATTAGATACC	20-mer	45% GC	6190 7.0	4.0	3.0	6.0	208300.0	6190.1	4.8	29.7	
Paire3_Reverse	CACCACTATCAACATAGAAG	20-mer	40% GC	6112 10.0	2.0	5.0	3.0	210200.0	6112.1	4.8	29.1	
Paire4_Forward	GAGGGTTGGTGTATTATG	20-mer	40% GC	6249 3.0	9.0	0.0	8.0	198100.0	6249.1	5.0	31.5	
Paire4_Reverse	ACCCTAACACCAATAATCC	20-mer	40% GC	5983 9.0	3.0	8.0	0.0	196700.0	5983.0	5.1	30.4	
Paire5_Forward	TCTGTGATTCCTGGCTTTAG	20-mer	45% GC	6105 2.0	9.0	4.0	5.0	180800.0	6105.0	5.5	33.8	
Paire5_Reverse	AACAGTAGCAGCCCAATAAG	20-mer	45% GC	6128 9.0	2.0	5.0	4.0	208900.0	6128.1	4.8	29.3	
Paire6_forward	GTTATGGGTGAGCGGTTTAC	20-mer	50% GC	6219 3.0	7.0	2.0	8.0	194800.0	6219.1	5.1	31.9	
Paire6_reverse	GGCTCAGCATGAACATTG	20-mer	50% GC	6117 5.0	5.0	5.0	5.0	191000.0	6117.0	5.2	32.0	
Paire7_Forward	TCACACGTGATTACCAATAG	20-mer	40% GC	6085 7.0	5.0	5.0	3.0	199200.0	6085.0	5.0	30.5	
Paire7_Reverse	GGCATAGACAATGGAGTAAC	20-mer	45% GC	6199 8.0	3.0	3.0	6.0	210600.0	6199.1	4.7	29.4	
Paire8_Forward	GCGTGTTTCTATAGTGAGAC	20-mer	45% GC	6163 4.0	7.0	3.0	6.0	195000.0	6163.1	5.1	31.6	
Paire8_Reverse	GTCACGACTCAAGCATTC	20-mer	45% GC	6070 7.0	4.0	6.0	3.0	194200.0	6070.0	5.1	31.3	
Paire_Test_Forward	TATTGTGTACGTGGTGTCGT											
Paire_Test_Reverse	GTAAGAACGATTAACAGCCCA											

B



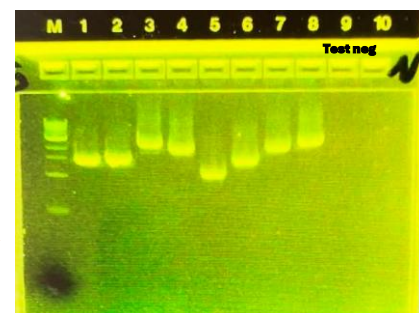
C



1.5 % agarose

1 → 1 PCR simplex : Paire 4
2 → 1 PCR simplex : Paire 5
3 – 4 → 1 PCR multiplex : Paires 1 et 3
5 – 6 → 1 PCR multiplex : Paires 6 et 8
7 - 8 → 1 PCR multiplex : Paires 2 et 7

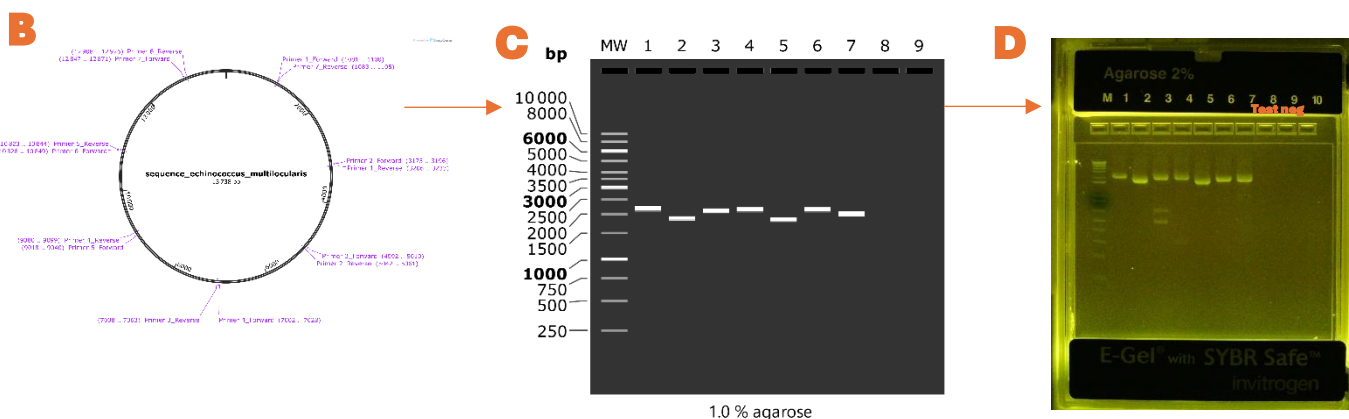
D



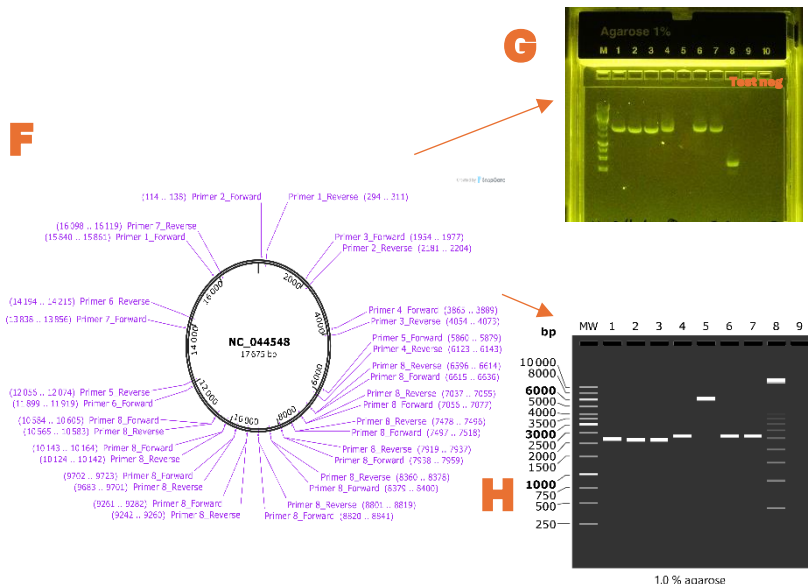
Annexe 3 : Le tableau A liste l'ensemble des amorces PCR conçues par le service de microbiologie de Liège afin d'amplifier le mitogénome de l'*E. granulosus*, la figure B représente leur distribution sur le mitogénome, la figure C nous montre le résultat attendu par PCR et sa légende en orange. Pour finir la figure D représente le test non multiplexé sur un gel d'agarose 2%.

D) Annexe 4 :

Primer name	Primer sequence (5' 3')	Primer position	PCR product length (bp)	Fragment sizes a (bp)	Multiplex/ uniplex PCR
E1f	TTCAAAGAGTGCTGGATTTC	1081...1100	2153	-	1
E1r	CAATACTCCAGTAAAAATTGAAGAATC	3206...3233			
E2f	GAAGTAGTGGGTTTGATTGTC	3175...3196	1887	1373; 514b	2
E2r	CACCCAACAACCTAAGCTTC	5042...5061			
E3f	CATGCTGAGGCTACTAGAG	4992...5010 7038...	2072	1120; 545;	3
E3r	GTATCCCTTTAAATCTGTTGAATCTG			288; 115; 5c	
E4f	GGGGGTTATGTAGAATTATTTATTTG	7002...7028	2098	2098b	2
E4r	TACACCAACCAAGAGAAAG	9080...9099			
E5f	ATAGGTGTGGTTAGAGAAAATTG	9018...9040	1827	1487; 340c	3
E5r	CAGCTATACAACAGGTTAATG	10823...10844			
E6f	ACCTGGTTGTATAGCTGTGAAG	10828...10849	2099	1272; 827b	2
E6r	CACCAACAGGAAACGAATC	12908...12926			
E7f	GACTATAAGATTGTTGGTCATCAG	12847...12871	1997	1704; 248; 45c	3
E7r	GAAAGGAATCCAGCACTCTTTG	1083...1105			



nom	sequences
F1	TCGTTTTACACGCGATTGAAC
R1	CTCAAAGCATTCAAACGC
F2	GTTGATTCGTGTTAATTTTTGGAG
R2	GCTACCTTTGCACAGTCAATATAC
F3	GGTACCTAGTTTTGTTATATTGT
R3	ACAACCATCTACAGCAGCAA
F4	GATGCTGTTCTCTGCTCTTAATC
R4	TGCTTAGTAAAAACACCCCA
F5	AAGAAGATTGTGGCTTTGTC
R5	CAACCACAGGCAACCTATC
F6	GGGCATGGCTTTGTATTATGG
R6	CAAAAACACCTCATTAAACCAC
F7	TTGTGGTGTTTTATGATG
R7	TGAGAATAGCCAAAACCTTAC
F8	CTCTGTCTATGGTTGGGTAATG
R8	CTACTAATGCGACCGACAC



Annexe 4 : Les tableaux A et E listent l'ensemble des amorces PCR conçues par les chercheurs de Besançon afin d'amplifier le mitogénome de l'*E. multilocularis* (A-D) et de l'*E. granulosus* (E-H), les figures B et F représentent leur distribution sur le mitogénome, les figures C et H montrent le résultat attendu par PCR. Pour finir, les figures D et G présentent le test non multiplexé sur un gel d'agarose à 2 %.

