

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE
DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS**

19-010-2.2-C-IRDA

DURÉE DU PROJET : 02/2019 AU 02/2024

RAPPORT FINAL

Réalisé par :
Richard Hogue, Ph.D., IRDA
Joël D'Astous-Pagé, M.Sc., IRDA
Thomas Jeanne, M.Sc., RDA

Février 2024

Les résultats, opinions et recommandations exprimés dans ce rapport émanent de l'auteur ou des auteurs et n'engagent aucunement le ministère de l'Agriculture, des Pêcheries et de l'Alimentation.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

1. RÉSUMÉ DU PROJET

Le rapport d'étape précédent (1 février 2021) avait fait état de quelques difficultés liées à la récolte d'échantillons virosés. Dans le contexte de l'épidémie de Covid19, nous avons proposé de prolonger le projet jusqu'en février 2024 pour recueillir des échantillons de framboisier infectés par des virus lors de campagnes additionnelles réalisées en 2022 et 2023, en partenariat avec le LEDP et des conseillers du MAPAQ.

Les campagnes 2019-2020 nous ont permis de réaliser l'objectif 1 du projet chez le fraisier. Tous les échantillons de tissus de fraisier montrant des symptômes d'infection virale reçus par le LEDP ont été traités par séquençage à haut débit (SHD) selon les technologies Illumina NextSeq et Oxford Nanopore MinION. Les diagnostics ont été comparés avec la méthode de référence ELISA ou RT-PCR (Obj. 2) par deux professionnels de notre équipe.

Nous avons prévu un horaire de formation et de transfert technologique avec le personnel du LEDP pour débuter en 2022 la formation à l'utilisation de ces nouvelles méthodes SHD au dépistage des virus du fraisier. Des protocoles ont été rédigés à cet effet (Obj. 4) et ils ont été utilisés lors des formations. Des rencontres périodiques avec le personnel du LEDP ont été réalisées pour comparer les coûts d'utilisation et d'implantation de ces nouvelles méthodes avec ceux des méthodes actuelles. Un tableur d'évaluation des coûts d'analyses a été développé en partenariat avec le personnel du LEDP (Obj. 3).

En 2021, un étudiant en génie biotechnologie et un professionnel de recherche ont participé au développement de deux sous-programmes (partie logique) nécessaires au fonctionnement de l'outil d'aide à la décision (OAD) utilisé pour le diagnostic viral. De plus, un second professionnel a débuté le développement du visuel de l'OAD (partie interface). Les deux parties logiques permettent : 1- Le traitement automatisé et l'identification des virus à partir d'un fichier de séquençage SHD; 2- La construction automatisée d'une base de données de signatures virales de référence (SVr) utilisés pour l'examen d'échantillons de fraisiers ou de framboisiers ou de pommes de terre.

Un troisième sous-programme a été développé et complété pour identifier des signatures de virus inconnues (SV) (Obj. 2). La partie visuelle offerte par l'OAD est une interface conviviale que le phytopathologiste utilise pour effectuer ses analyses et réviser ses diagnostics. Nous avons présenté en 2022-2023 une version bêta de l'OAD au personnel du LEDP. En automne 2023, nous avons assuré la formation à l'utilisation de l'OAD pour le diagnostic des virus infectant les fraisiers, les framboisiers et les pommes de terre sur les plateformes NextSeq et MinION. En 2024, les phytopathologistes ont pu réaliser les essais de formation à l'emploi autonome de l'OAD et l'équipe de recherche du Laboratoire d'écologie microbienne de l'IRDA (LEM) et du Centre de Recherche du CHU de Québec (CRCHUL) a assuré le suivi de la formation et de la mise à niveau de l'OAD en fonction des rétroactions.

Les principaux résultats obtenus depuis le dernier rapport d'étape sont :

- 1) La comparaison des résultats diagnostiques entre les méthodes du LEDP et les deux méthodes de séquençage proposées NextSeq et MinION.
- 2) Le développement de sous-programmes des parties logique et d'interface de l'OAD.
- 3) Mise au point et validation des paramètres de séquençage MinION.

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

Table des matières

| | | |
|------|---|----|
| 1. | RÉSUMÉ DU PROJET | 1 |
| 2. | OBJECTIFS ET APERÇU DE LA MÉTHODOLOGIE | 6 |
| 2.1. | Objectifs | 6 |
| 2.2. | Méthodologie..... | 7 |
| 3. | RÉSULTATS SIGNIFICATIFS OBTENUS | 11 |
| 3.1. | Base de séquences de référence spécifiques adaptée aux données de SHD | 11 |
| 3.2. | Cas diagnostiques obtenus | 13 |
| 3.3. | Bilan des détections par approche conventionnelle et SHD..... | 14 |
| 3.4. | Proposition d'utilisation du SHD | 15 |
| 3.5. | Outils de traitement et de visualisation des résultats (PhytoSHD) | 16 |
| 3.6. | Modèle prédictif | 20 |
| 3.7. | Analyse économique | 21 |
| 3.8. | Tutoriel et protocoles..... | 22 |
| 4. | ACTIVITÉS DE DIFFUSION ET TRANSFERT AUX UTILISATEURS | 23 |
| 4.1. | Diffusions | 23 |
| 4.2. | Transfert au LEDP..... | 24 |
| 5. | PERSPECTIVES | 24 |
| 6. | POINT DE CONTACT POUR INFORMATION..... | 25 |
| 7. | REMERCIEMENTS AUX PARTENAIRES FINANCIERS | 25 |
| 8. | RÉFÉRENCES | 25 |
| 9. | ANNEXES..... | 27 |

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

Liste des figures

| | |
|--|----|
| Figure 1. Schéma du processus de traitement d'un cas diagnostique par SHD | 7 |
| Figure 2. Principales étapes de l'approche de détection de virus par SHD avec l'outil PhytoSHD | 7 |
| Figure 3. Résumé des principales étapes de traitement bio-informatique | 10 |
| Figure 4. Fréquence du pourcentage de couverture du génome viral par des signatures virales d'un virus de la base de signatures de référence selon le type d'hôte..... | 12 |
| Figure 5. Nombre moyen de signatures virales selon leurs longueurs pour les hôtes d'intérêts. | 13 |
| Figure 6. Schéma de l'architecture de l'application PhytoSHD..... | 17 |
| Figure 7. Menu « Accueil » de l'application PhytoSHD | 17 |
| Figure 8. Menu « Saisie des données » de l'application PhytoSHD | 18 |
| Figure 9. Rapport d'évaluation de la qualité d'un séquençage Illumina NextSeq produit par FastQC..... | 18 |
| Figure 10. Rapport d'évaluation de la qualité d'un séquençage Nanopore MinION produit par FastQC | 19 |
| Figure 11. Échantillons sélectionnés dans le menu « Traitement des données » de l'application PhytoSHD..... | 19 |
| Figure 12. Exemple de visualisation des résultats dans le menu « Résultats » de l'application PhytoSHD..... | 20 |
| Figure 13. Exemple de visualisation avancée des résultats dans le menu « Résultats » de l'application PhytoSHD | 20 |
| Figure 14. Application « FastqFinder-ONT » pour accélérer la saisie des noms de fichiers FASTQ | 23 |

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

Liste des tableaux

| | |
|--|----|
| Tableau 1. Virus avec une couverture du génome viral inférieure à 80%..... | 12 |
| Tableau 2. Comparatif des fonctionnalités entre les méthodes actuelles et SHD..... | 15 |
| Tableau 3. Évaluation du modèle prédictif pour le séquençage sur la plateforme Illumina ... | 21 |
| Tableau 4. Évaluation du modèle prédictif pour le séquençage sur la plateforme Nanopore | 21 |
| Tableau 5. Coût matériel pour les différentes technologies d'analyse (\$CAD) arrondi au dollar en haute saison. | 22 |
| Tableau 6. Activités de diffusion | 23 |

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

Liste des Abréviations et Acronymes

| | |
|--------|---|
| ADN | Acide désoxyribonucléique |
| ARN | Acide ribonucléique |
| BD | Bases de données |
| ELISA | <i>Enzyme-linked immunosorbent assay</i> |
| HPLC | High-performance liquid chromatography |
| LEDP | Laboratoire d'expertise et de diagnostic en phytoprotection |
| LEM | Laboratoire d'écologie microbienne |
| OAD | Outil d'aide à la décision |
| ONT | Oxford Nanopore Technologies |
| PDIV | Processus de détection et d'identification des virus |
| RT-PCR | <i>Reverse transcription polymerase chain reaction</i> |
| SHD | Séquençage à haut-débit |
| SV | Signatures virales |
| SVr | Signatures virales de référence |
| USDA | Département de l'Agriculture des États-Unis |

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

2. OBJECTIFS ET APERÇU DE LA MÉTHODOLOGIE

2.1. Objectifs

L'objectif général du projet vise à développer une méthodologie de détection moléculaire rapide, fiable et sensible qui assure l'identification des virus des framboisiers et des fraisiers. Ce processus de détection et d'identification des virus (PDIV) combine les techniques de SHD à des outils d'analyses bio-informatiques pour aider le phytopathologiste à faire un diagnostic. Les techniques SHD peuvent être basées sur la plateforme Illumina NextSeq ou sur la plateforme MinION d'Oxford Nanopore. Le choix de l'une ou l'autre repose sur des considérations techniques, économiques et d'efficacité diagnostique.

Les objectifs spécifiques du projet sont de:

- 1) Développer une collection d'échantillons de fraisiers et de framboisiers sains et virosés ainsi que de séquences génomiques de chaque virus ciblé;
- 2) Comparer l'efficacité des méthodes diagnostiques ELISA et RT-PCR actuellement utilisées par le LEDP à celles d'un PDIV basé sur des techniques SHD. Ultimement, le PDIV conduira au développement d'un outil d'aide à la décision (OAD) pour le diagnosticien;
- 3) Comparer les coûts d'implantation et d'utilisation des méthodes diagnostiques actuellement utilisées par le LEDP à ceux du PDIV + OAD;
- 4) Former le personnel du LEDP et d'autres utilisateurs potentiels à l'emploi du PDIV, de sa base de données et de son interface conviviale, l'OAD.

L'approche de détection par SHD se divise en deux grandes étapes (Figure 1) selon le lieu d'exécution. Une première étape est réalisée au laboratoire (Observation et séquençage) et la seconde étape réalisée au bureau du phytopathologiste (Outil d'aide au diagnostic et de visualisation des résultats SHD) mène à un verdict du cas diagnostique soumis.

Au cours de la période 2020-2021, pour compenser le faible nombre d'échantillons de virus du framboisier, et pour profiter des activités du Laboratoire d'analyse biologique de l'IRDA qui analyse plusieurs milliers d'échantillons de feuilles et de tubercules infectés par des virus, nous avons réalisés les objectifs prévus pour le développement d'un PDIV en utilisant des centaines d'échantillons pour lesquels un diagnostic de RT-PCR qui confirmait la présence d'un ou de plusieurs virus dans l'échantillon soumis au dépistage SHD. Le travail réalisé avec les échantillons de pommes de terre a permis d'accélérer l'atteinte des objectifs du projet lorsque le nombre d'échantillons de tissus virosés de fraisiers et de framboisiers a été jugé suffisant.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

Étape 1: Observations et séquençage



Étape 2: Outil d'aide au diagnostic et à la visualisation

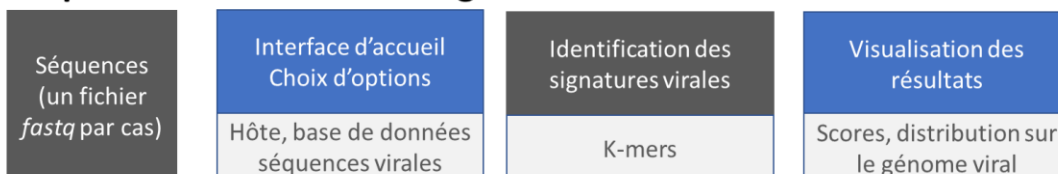


Figure 1. Schéma du processus de traitement d'un cas diagnostique par SHD

2.2. Méthodologie

Approche globale de la méthode par SHD

Globalement l'approche se divise en cinq grandes étapes : 1. Sélection du cas et prélèvement des tissus; 2. Extraction des ARN totaux; 3. Librairie et séquençage SHD; 4. Analyses des séquences; 5. Visualisation des résultats et aide à la décision (Figure 1). La stratégie adoptée vise à arrimer le plus facilement possible les nouvelles méthodes en lien avec le SHD aux méthodes actuelles du laboratoire du LEDP.

Nous avons réceptionné des échantillons provenant de cas reçus par le LEDP, de la collection de Robert Martin du Département de l'Agriculture des États-Unis (USDA) et de nos campagnes de collectes d'échantillons auprès des producteurs. L'approche a été optimisée pour la détection dans les feuilles de fraisier, de framboisier et de pomme de terre. Tous les échantillons ont été identifiés par un numéro unique et ont été conservés à -80°C en vue de l'extraction d'ARN. Nous avons aussi collecté des données connexes (région d'origine, photo des plants/cultures, type de symptômes, âge des cultures, type de cultivars, etc) lorsque possible.

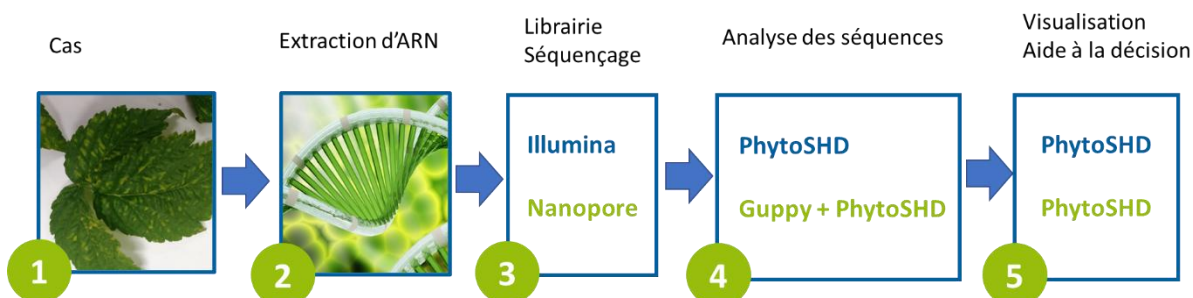


Figure 2. Principales étapes de l'approche de détection de virus par SHD avec l'outil PhytoSHD

Extraction de l'ARN

Deux protocoles sont possibles pour l'extraction des ARN totaux, dont les ARN viraux, des tissus. L'approche chimique utilisant du CTAB 3% (v/v) avec précipitation au chlorure de lithium est décrite au protocole (Pro_EXT_CTAB_ARN_2022.03) et l'approche par affinité est

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS 19-010-2.2-C-IRDA

décrite au protocole (Pro_EXT_QiagenARN_2022.03) et utilise la trousse RNeasy Plant Mini (QIAGEN, Toronto, Canada). L'approche chimique a l'avantage d'extraire un plus grand nombre d'échantillons en même temps (jusqu'à 80 échantillons), mais elle demande d'être réalisée par un personnel ayant reçu plus de formation. L'approche par affinité a l'avantage d'utiliser des trousse d'extraction (*kits*) que le personnel du LEDP est déjà familier à utiliser. Il est possible d'automatiser l'étape d'extraction avec la plateforme QIAcube déjà utilisée par le LEDP (l'automatisation de l'extraction n'a pas été utilisée dans le cadre de ce projet). Pour ses avantages, l'approche par affinité est considérée comme la méthode principale et l'approche chimique comme une méthode alternative. Les deux méthodes d'extraction n'influencent pas le diagnostic et peuvent être utilisées avec les deux méthodes de séquençage présentées dans les prochaines sous-sections.

Préparation des librairies et séquençage

Le séquençage peut être fait sur un séquenceur NextSeq de Illumina (San Diego, CA, États-Unis) avec des amorces aléatoires. Ce service est offert par la plateforme de génomique de l'Institut de recherche en immunologie et en oncologie (IRIC). Le séquençage peut aussi être réalisé au LEDP par séquençage MinION (Oxford Nanopore Technologies, Oxford, UK).

En bref, le séquençage MinION a l'avantage d'être plus rapide et de diminuer les coûts d'analyse bien que les séquences produites soient de moins bonne qualité que celles obtenues par la technologie NextSeq. Plus de détails sont disponibles sur les protocoles mentionnés en annexe et fournis au LEDP lors des formations.

Préparation des librairies pour séquençage NextSeq

Le séquençage par NextSeq demande le moins d'intervention de la part du LEDP. Les extraits d'ARN totaux peuvent être envoyés sur glace sèche directement à la plateforme de l'IRIC. Il faut prévoir entre 5 à 10 jours ouvrable pour recevoir les résultats.

Préparation des librairies pour séquençage MinION

Plus de détails sont disponibles sur les protocoles détaillés mentionnés en annexe (Pro_LIB_MinION_2024.01) et fournis au LEDP lors des formations. Il est important de noter que la technologie MinION évolue très rapidement, les trousse et cellules utilisées ont évoluées au cours du projet. La méthode la plus récente utilise la chimie V10 avec les cellules R9. Nous proposons aussi un ajustement au protocole afin de rendre la chimie compatible avec la nouvelle version de cellule (R10.4) [1]. Une nouvelle trousse (SQK-PCB114.24) a été rendue disponible en décembre 2023 et permet l'utilisation de cellules R10.4 sans l'achat d'accessoire supplémentaire. Vu les courts délais depuis sa mise en service en décembre 2023, nous n'avons pas été en mesure de tester la nouvelle trousse, mais elle semble compatible avec le protocole mentionné en annexe.

1. Réactifs nécessaires

- PCR-cDNA Barcoding (SQK-PCB111.24; ONT) [2]
- Billes magnétiques (Agencourt AMPure XP ou AxyPrep Mag PCR Clean-Up)
- dNTP (10 mM)
- LongAmp Taq
- Maxima H Minus RT ou Superscript IV (SS-4)

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

- RNaseOUT (40 U/μl)
 - Exonuclease I
 - Amorces CRTA-6N (50μM) – purifiées par HPLC [3]
 - Cellule de type Flow Cell (R9 ou R10.4)
2. Étapes à suivre (voir le protocole détaillé mentionné en annexe)
- Transcriptase inverse et synthèse de l'ADNc (PCR1)
 - Purification des produits PCR et contrôle qualité
 - Indexation (PCR2)
 - Purification des produits PCR et contrôle qualité
 - Purification supplémentaire du PCR2 – facultatif
 - Chargement de la cellule (*flow cell*) et séquençage

Amorces CRTA-6N (purifié par HPLC) [3] :

5'-CTTGCCTGTCGCTCTATCTTCAGAGGAGAGTCCGCCGCCCGCAAGNNNNNN-3'

Analyse des séquences SHD

Le traitement initial des séquences (i.e. identification des paires de bases, qualité du séquençage) est fait automatiquement par la plateforme MinION (ou NextSeq) et il produit à la fin du processus de séquençage un fichier *fastq* contenant les séquences à haut débit. Dans le cas de la plateforme MinION, il est possible de faire un traitement intermédiaire (*basecalling*) des séquences pour améliorer la définition des résultats. Il sera discuté dans la section Résultats de nos recommandations concernant le traitement intermédiaire.

À l'aide de l'application web PhytoSHD, le phytopathologiste pourra importer les données produites au paragraphe précédent, démarrer une analyse via le module PhytoV et visualiser les résultats de détection des signatures virales (SV). De plus, des modules d'analyses des séquences utilisant l'apprentissage automatisé viendront aider le phytopathologiste à apprécier les résultats et poser son diagnostic. Deux modules d'analyses sont inclus avec l'OAD, l'un est utilisé pour les séquences issues d'échantillons de fraisier, l'autre est utilisé pour les échantillons de framboisier. Ses modules seront discutés dans la section Résultats des modèles.

Les différentes phases de traitement bio-informatique du module PhytoV vont ainsi :

1. Importation: Les fichiers de séquences et les métadonnées sont intégrés à l'OAD. La qualité des séquences est évaluée.
2. Filtration : Il y a une filtration des séquences selon leur qualité, et il y a un retrait des séquences non biologiques.
3. Nettoyage : Retrait des séquences de l'hôte.
4. Détection des virus : Détection des signatures de séquences virales et évaluation de la couverture du génome virale pour chaque détection.
5. Prédiction : Utilisation d'un modèle d'intelligence artificiel pour identifier les faux positifs.
6. Intégration à la BD et rapport : Intégration des résultats dans la base de données de l'OAD et préparation des visuels pour le phytopathologiste.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

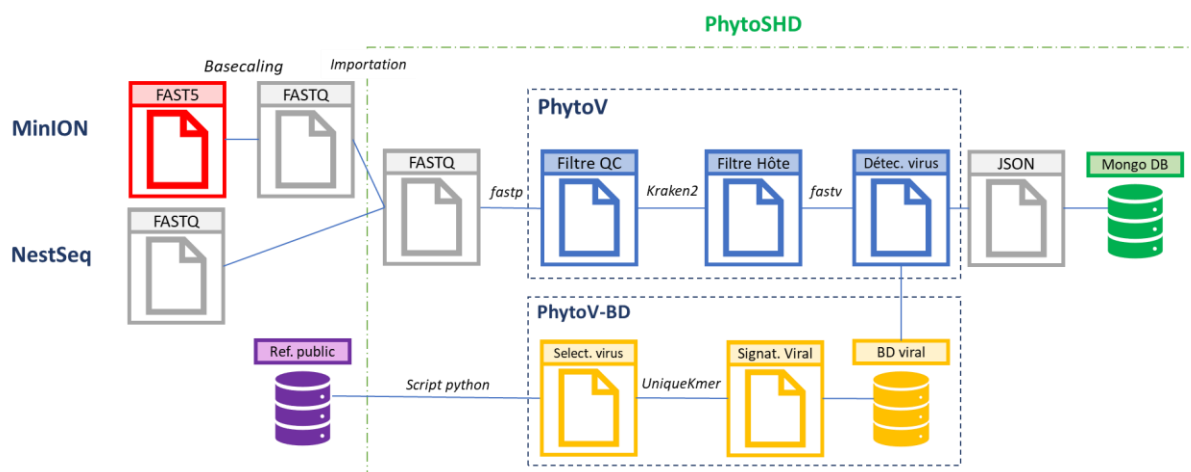


Figure 3. Résumé des principales étapes de traitement bio-informatique

Validation et contrôle qualité

Pour s'assurer du bon déroulement et de la fiabilité des résultats, des points de contrôles ont été inclus dans les protocoles. Ainsi aux étapes clés des protocoles, des dosages sont effectués pour s'assurer du bon déroulement des analyses. Il est recommandé de toujours effectuer les points de contrôles prioritaires pour les analyses routinières et d'effectuer les points de contrôles secondaires lors de mise au point ou lors de difficultés rencontrées à réaliser les protocoles. Les points de contrôles sont les suivant :

1. Dosage et quantification (Absorbance et fluorescence) de l'ARNt post-extraction d'ARN - Prioritaire
2. Dosage (Absorbance) de l'ADN post-PCR1 d'ARN - Secondaire
3. Dosage et quantification (Absorbance et fluorescence) de l'ADN post-PCR2 - Prioritaire
4. Quantification (Fluorescence) du pool d'ADN pré-chargement dans la cellule - Secondaire

De plus, nous avons développé un contrôle interne à partir d'un fragment de virus du saumon (Atlantic salmon calicivirus.; NC_024031.1; 1541-2240pb) [4]. Ce fragment de virus peut facilement être produit en laboratoire et il ne contient aucune région commune avec les autres virus de la base de signatures virales de référence du fraisier et du framboisier. La signature du fragment de virus du saumon est ainsi facile à distinguer dans les échantillons d'extraits de tissus de fraisiers et de framboisiers. Le contrôle positif est un échantillon spécifique ajouté au début de la préparation des librairies dans un extrait de plant sain (ou non). La séquence du virus contrôle interne est incluse dans la base de signatures virales de référence et sera détectée comme un virus (avec la mention **référence**) dans la visualisation des résultats par l'application web PhytoSHD. Ainsi, il sera possible au pathologiste d'utiliser le résultat de cet échantillon témoin afin de confirmer facilement et rapidement le bon déroulement des étapes de préparation des librairies, de séquençage, d'analyse des séquences et de visualisation des résultats. Ce contrôle fait office de témoin positif et il n'est pas recommandé d'ajouter le contrôle dans chacun des échantillons lors d'analyses de routine.

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

3. RÉSULTATS SIGNIFICATIFS OBTENUS

3.1. Base de séquences de référence spécifiques adaptée aux données de SHD

La détection virale se base sur l'identification de signatures virales de référence (SVr) spécifiques à une liste de virus d'intérêt. Ces SVr sont générées par l'outil UniqueKMER [5] à partir des génomes de référence des virus présents dans la base de données virales NCBI. Par défaut, le module PhytoV utilise une sélection d'environ 284 génomes viraux sélectionnés par notre équipe.

Pour faire la sélection des génomes viraux d'intérêt, notre modèle utilise la base de données *virushostdb* [6] qui associe des virus à leurs hôtes. En date du 17 décembre 2023 (dernière mise à jour de notre module PhytoV), il y avait 13 771 virus de répertoriés dans la base de données, dont 2 381 ont un hôte appartenant au clade des Euphyllophyta. Cette première sélection permet d'éliminer tous les virus ne s'attaquant pas à des plantes. Par la suite, nous avons éliminé les virus satellites car leur génome très court génère peu de k-mers spécifiques. Pour terminer, nous avons limité la collection aux virus s'attaquant à 10 genres d'hôte (Malus, Rubus, Prunus, Fragaria, Solanum, Nicotiana, Humulus, Beta, Cucumis et Rheum). Le but étant d'englober les virus d'intérêts chez la fraise, la framboise et la pomme de terre ainsi que les virus des cultures pouvant se retrouver botaniquement liée à ces plantes ou encore à proximité de ces cultures au Québec. De plus, une sélection manuelle a été faite pour s'assurer qu'il n'y ait pas de virus d'intérêt manquant ou en duplicata (méthode discutée plus bas). Le module PhytoV par défaut de l'OAD, contient 284 virus qui sont listés en annexe (VirusOAD).

Nous travaillons à rendre possible aux utilisateurs du LEDP de créer leurs propres modules de détection en important des fichiers *fasta* de génomes viraux qu'ils auront sélectionnés. Cette fonctionnalité ne constituant pas un livrable du projet, elle n'est pas encore fonctionnelle, mais elle pourrait être finalisée si le besoin du LEDP évolue.

Les SVr sont spécifiques pour un virus dans la base de référence. De plus, les SVr ciblant l'hôte ou des régions peu variables ont été retirées. Pour la grande majorité des virus, plus de 80% du génome viral est couvert par une SVr (Figure 4). Les virus ayant une signature de plus petite longueur sont présentés dans le tableau 1. Cette plus faible couverture peut avoir différentes raisons dont, des régions communes avec d'autres virus, avec l'hôte ou une faible variabilité du génome viral. Pour assister le diagnosticien, la couverture maximale est affichée dans la visualisation des résultats de l'application web PhytoSHD.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

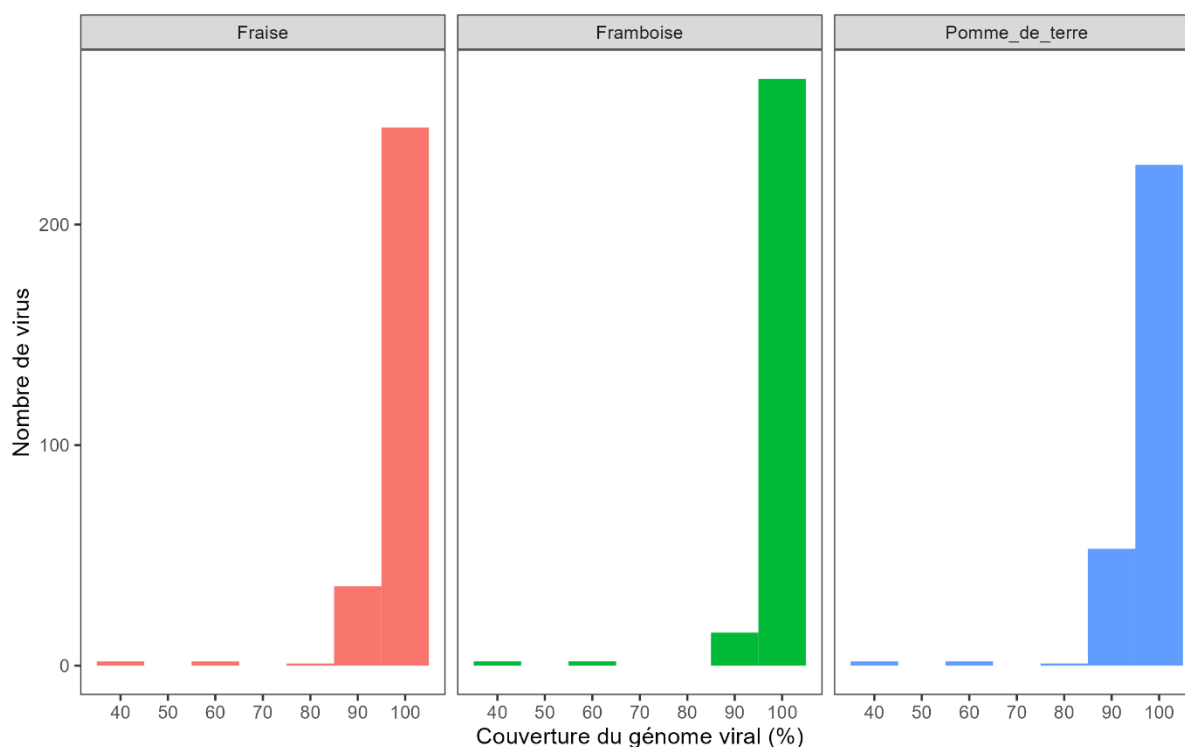


Figure 4. Fréquence du pourcentage de couverture du génome viral par des signatures virales d'un virus de la base de signatures de référence selon le type d'hôte.

Tableau 1. Virus avec une couverture du génome viral inférieure à 80%.

| Virus | Fraise | Framboise | Pomme de terre |
|--------------------------------------|--------|-----------|----------------|
| 260378_Pepper_yellow_vein_Mali_virus | 36.89% | 38.17% | 36.37% |
| 471498_Peach_chlorotic_mottle_virus | 40.20% | 40.76% | 40.16% |
| 1195378_Tomato_dwarf_leaf_virus | 57.18% | 57.18% | 56.95% |
| 2137858_Mume_virus_A | 84.81% | 94.01% | 78.57% |

Choix de l'utilisation des signatures virales

Dans la demande de financement déposée, nous proposons la détection de virus à l'aide de la technique *e-probe*. Pour rappel, cette technologie utilisait un cocktail d'amorces spécifiques aux virus d'intérêts (chaque virus avait une liste de nombreuses amorces différentes). Lors de la première année et après des discussions avec les phytopathologistes du LEDP, nous avons conclu que l'approche utilisant des signatures virales (k-mer) était plus appropriée à la réalisation des objectifs du projet. Le choix des signatures virales comportait plusieurs avantages :

1. L'évaluation des interactions entre signatures virales est facile à mesurer.
2. Les *e-probes* étaient coûteuses à produire.
3. Il est facile et rapide de produire de nouvelles signatures virales par k-mer.

Ainsi, les SVr sont créées spécifiquement pour un hôte suivant une liste de génomes viraux de référence. Lors de la création des SVr, l'outil limite les risques d'homologies entre d'autres génomes viraux (d'une même liste) et le génome de l'hôte. Comme pour des amorces lors d'une PCR, le choix de la longueur des SVr est important. Une courte signature a peu de chance d'être spécifique et génère peu de signatures spécifiques pour un virus. Pour ce projet,

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

nous avons paramétré la base de données pour une longueur de 26 pb afin de sélectionner des SVr optimales (Figure 5).

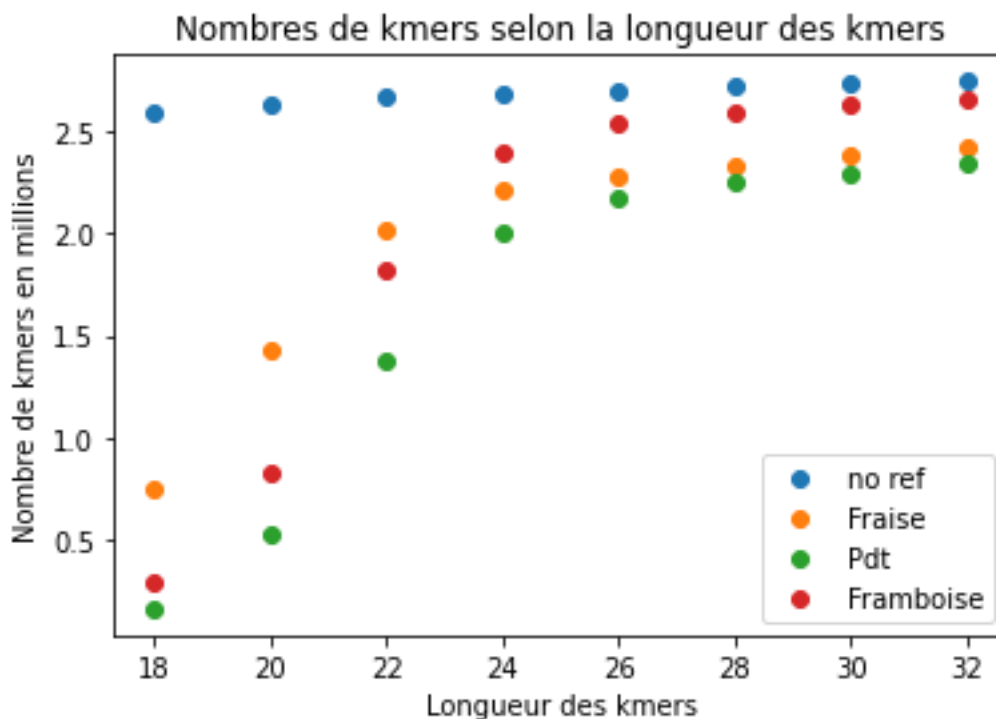


Figure 5. Nombre moyen de signatures virales selon leurs longueurs pour les hôtes d'intérêts.

Filtration supplémentaire de la BD

Tel que décrit précédemment, une succession d'étapes a permis la sélection des virus d'intérêts à inclure dans la base de données. Une filtration supplémentaire manuelle a été effectuée pour s'assurer de la qualité des SVr générées. Dans les précédents rapports d'étape nous avons discuté de l'ajout de virus d'intérêts recommandés par le LEDP. Nous avons retiré les virus de faible intérêt ou ayant trop d'homologie entre eux (ie.: plusieurs sous-espèces d'un même virus, le même virus identifié sous deux noms). De plus, nous avons effectué un alignement croisé entre les génomes viraux. Certains virus, bien que distincts, peuvent contenir des régions similaires et rendre la détection de ces virus plus difficile. Nous avons identifié neuf groupes de virus ayant plusieurs zones d'homologie. La liste des groupes est présentée en annexe (VirusHomologie). Aucun des virus d'intérêts de ce projet ne sont présents dans ces groupes.

3.2. Cas diagnostiques obtenus

Les objectifs des campagnes d'échantillonnage de plantes ou tissus infectés ont changé au cours du projet. À l'origine, nous prévoyions recevoir un volume important de cas virosés pour les framboisiers et les fraisiers. Finalement, nous avons été en mesure de collecter des ARN ou des tissus virosés pour neuf virus chez la fraise et six virus chez la framboise. Pour faciliter les analyses de comparaison d'efficacité de la méthode SHD versus les méthodes conventionnelles ELISA et RT-PCR, il a été décidé d'inclure la pomme de terre pour laquelle nous avons une collection de six virus ainsi que des cas sains. Au total, nous avons reçu du LEDP 11 cas chez le fraisier (9 positifs) et 39 cas chez le framboisier (5 positifs). De nos partenaires, nous avons reçu 46 cas chez le fraisier (25 positifs) et 165 cas chez le framboisier

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS 19-010-2.2-C-IRDA

(14 positifs). Comme expliqué dans le précédent rapport d'étape, la majorité des cas positifs chez la framboise sont des échantillons lyophilisés provenant de la collection du laboratoire de Robert Martin du Département de l'Agriculture des États-Unis (USDA). Pour maintenir une veille, nous avons analysé des échantillons de pépinières lors des campagnes des saisons 2022 et 2023. Ils se sont tous avérés négatifs. Ainsi, dans le cas du framboisier, nous sommes d'avis que l'amélioration des méthodes de reproduction dans les pépinières et les changements de pratiques agricoles aux sites de productions semblent avoir permis de limiter la propagation de virus.

3.3. Bilan des détections par approche conventionnelle et SHD

Une étape importante du processus d'évaluation des méthodes de SHD est la comparaison des diagnostics avec une méthode de référence (RT-PCR) utilisée par le LEDP.

Dans le cadre de ce projet, deux méthodes de SHD ont été proposées: la méthode Illumina et la méthode Nanopore. Les deux méthodes donnent le même type de détection pour 29 cas de Fraisier, 27 cas de Framboisier et 7 cas de Pomme de terre. Puisque les deux méthodes de SHD donnaient des résultats de détection similaire, la méthode Nanopore a été privilégiée pour la suite du projet.

Il est possible de distribuer les résultats de la comparaison des cas diagnostiques entre SHD et la méthode de référence RT-PCR en trois classes :

1. **Concordance** : Concordance entre les approches conventionnelles et SHD
La concordance entre les détections était autant pour des cas sans virus, avec un seul virus ou avec des infections multiples.
2. **Divergence** : Pas de concordance complète entre les approches conventionnelles et SHD
Ces cas concernent, Tomato ringspot virus (ToRSV; 12280) dont la détection est parfois plus faible par SHD et Strawberry crinkle virus (SCrV; 135656) dont la détection est généralement difficile par SHD.
3. **Gain** : Nouvelle détection avec l'approche SHD
Cela s'explique par le nombre limité de virus que le RT-PCR peut détecter comparativement aux génomes viraux qui se trouvent dans la base de données du SHD. Pour la majorité, les cas provenaient de la collection du laboratoire de Robert Martin (USDA). À l'exception du Strawberry polerovirus 1 (SPV1; 1564903) chez la fraise qui peut être détecté par RT-PCR, mais le test n'est pas toujours effectué en routine par le LEDP, ce qui avantage le verdict « Gain ».

Pour 88% des cas de fraisier et de framboisier, les méthodes par RT-PCR et par SHD ont donné exactement les mêmes résultats (Concordance). Pour 5% des cas, le SHD détectait des virus supplémentaires que le RT-PCR ne détectait pas (Gain). Des détections ont été faites uniquement par RT-PCR dans 2% des cas (Divergence) et uniquement par SHD dans 1% des cas (Gain). Pour terminer, dans 2% des cas, le SHD et le RT-PCR avaient des détections mais pas pour les mêmes virus (Divergence et Gain). Toutes les détections par SHD avec les tissus de pomme de terre ont donné le même verdict que par RT-PCR (Concordance). En conclusion, les protocoles SHD sont efficaces et peuvent être utilisés par le LEDP pour détecter les virus infectant les fraisiers, framboisiers ou pommes de terre.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

3.4. Proposition d'utilisation du SHD

Nous avons résumé les principaux avantages et inconvénients des méthodes SHD et les méthodes actuelles dans le Tableau 2. On remarque que le SHD est généralement moins rapide et un peu plus dispendieux. Il est à noter que la technologie de SHD est plus jeune que les méthodes actuellement utilisées par le LEDP (RT-PCR et ELISA). Le coût a grandement diminué dans les dernières années et il est à prévoir que les méthodes SHD deviendront encore plus accessibles dans le futur contrairement aux méthodes actuelles qui ont atteint un niveau de maturité technologique et économique. La méthode actuelle RT-PCR établit son verdict selon le mode d'un diagnostic par présence-absence, et un test est réalisé pour un virus spécifique. Cela limite le nombre de virus à tester en simultané et augmente les coûts d'analyse pour chaque nouveau virus à tester. Contrairement aux tests SHD qui présentent un coût fixe pour le séquençage. Il est de plus possible de traiter les données de séquençage obtenues de multiples façons pour identifier de multiples virus simultanément. Dans le cas de l'application PhytoSHD, la base de données actuelle contient 284 virus d'intérêts. Le LEDP pourra conserver les fichiers de séquençage des échantillons analysés et réanalyser les séquences dans le futur s'il développe de nouvelles méthodes de traitement des séquences ou de nouvelles bases de données de référence. Dans la même optique, lors de la découverte d'un nouveau virus d'intérêt, il sera possible de l'ajouter à la base de données de référence et de traiter à nouveau les cas soumis dans la base de données pour vérifier si le nouveau virus était présent dans les échantillons déjà analysés. À la demande du LEDP et avec un nouveau financement, l'équipe du LEM pourrait finaliser le développement de cette fonction et l'intégrer dans PhytoSHD.

Tableau 2. Comparatif des fonctionnalités entre les méthodes actuelles et SHD.

| Culture | Actuelles (LEDP) | SHD-NextSeq | SHD-MinION |
|---------------------------------------|------------------|-------------|------------|
| Temps d'un diagnostic | Très rapide | Moyen | Rapide |
| Coût par analyse | \$ | \$\$\$ | \$\$ |
| Nombre de virus ciblé simultanément | + à ++ | +++ | +++ |
| Évolution technologique | Non | Oui | Oui |
| Nouvelle analyse possible (novovirus) | Non | Oui | Oui |

Nous avons identifié trois scénarios d'utilisation du SHD dans un cadre de la détection de virus par le LEDP :

1. Remplacement de la méthode actuelle. Le SHD pourrait devenir la méthode de détection principale en remplacement des tests actuels (ie. RT-PCR et ELISA).
2. Confirmation des cas négatifs. Les approches SHD pourraient être utilisées en complément des méthodes actuelles dans les cas où le phytodiagnosticien juge qu'il y a une forte présomption d'infection virale mais que les méthodes actuelles se sont avérées inconcluantes à détecter un virus. Le SHD permettrait d'éliminer un plus grand nombre de virus à tester par les tests actuels. Il serait possible de séquencer un composite de plusieurs échantillons négatif pour diminuer les coûts.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS 19-010-2.2-C-IRDA

3. Dépistage des cultures aux Québec. Il serait possible d'utiliser annuellement le SHD sur une sélection de cas d'intérêts pour construire une banque de séquences virales spécifiques au Québec. À long terme, il serait possible de faire un suivi de l'évolution des virus ou permettre la découverte de novovirus.

3.5. Outils de traitement et de visualisation des résultats (PhytoSHD)

Pour faciliter le traitement des données de SHD et en visualiser les résultats, ces processus ont été intégrés à une application web appelée PhytoSHD. Le choix du développement d'une application web visait à répondre aux besoins du LEDP d'être en mesure de traiter les données de SHD par des utilisateurs non bio-informaticiens. Pour le moment, l'application PhytoSHD est hébergée sur un serveur externe sécurisé (Google Cloud). Il n'a pas encore été décidé si le personnel du LEDP préfère le transfert de l'application sur un serveur local. À la suite de consultations avec le personnel du LEDP, nous avons élaboré les critères suivants pour le développement de l'application web PhytoSHD:

1. L'application devra héberger les données SHD analysées (pas les données brutes).
2. Les données brutes seront hébergées temporairement, le temps de choisir le type d'analyse à effectuer.
3. L'application sera en mesure de visualiser efficacement les détections virales des cas analysés par SHD.
4. L'application web sera développée de façon modulaire pour lui permettre d'intégrer de futurs processus de traitement bio-informatique (module).

Architecture

Comme mentionné précédemment, l'application est basée sur une architecture modulaire. Chaque module est indépendant et suit des règles précises pour interagir avec les autres modules. Cela permet une plus grande flexibilité pour l'évolution de l'application web. De plus, cette architecture rend possible l'intégration de nouveaux modules de traitements bio-informatique (Figure 6, block mauve).

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

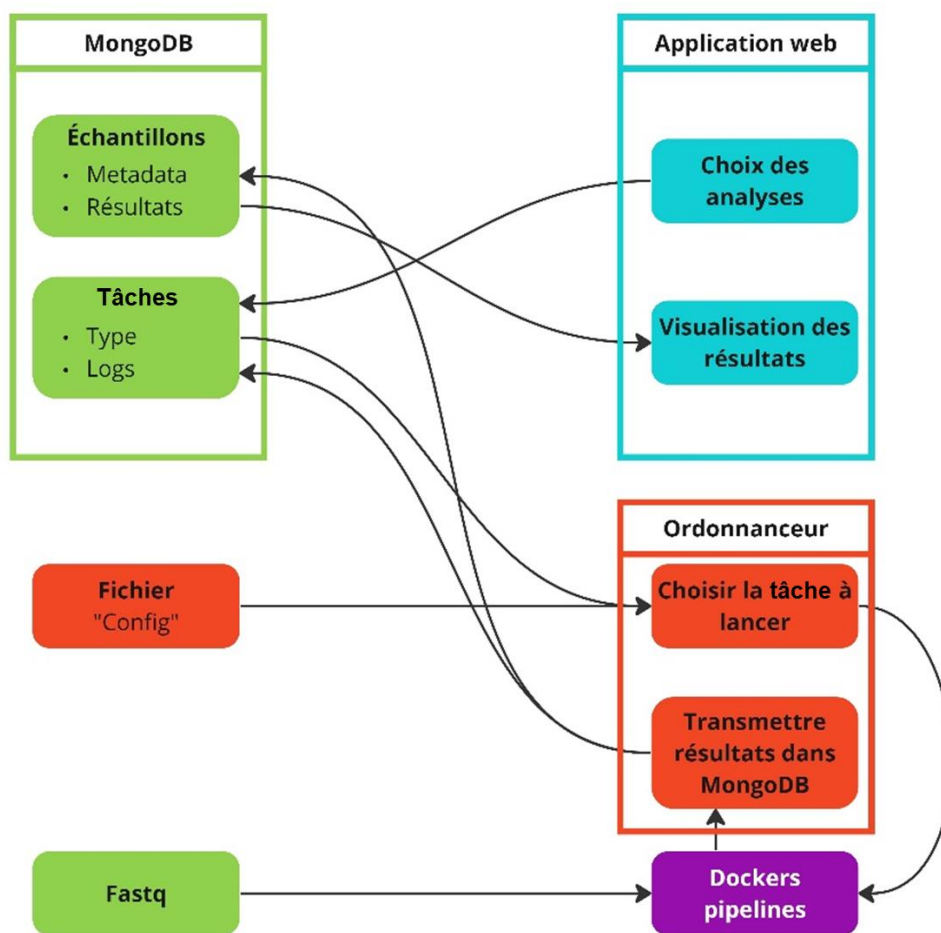


Figure 6. Schéma de l'architecture de l'application PhytoSHD

Fonctionnement de l'application PhytoSHD

L'application PhytoSHD présente quatre menus. La page principale est le menu « Accueil » qui présente un aperçu de la tâche en cours, qui permet la recherche rapide avec des mots clés pour sélectionner des cas d'intérêts et qui permet la modification d'information liée à un cas ou qui permet la suppression d'une entrée de donnée non nécessaire.

Accueil

Saisie de données

Traitement des données

Visualisation des résultats

Cas :

Rechercher un cas

Type d'analyse :

Rechercher un type d'analyse

Culture :

Rechercher une culture

Nom de run :

Rechercher une run

«

<

1

2

>

»

| Status | QC | Analyse |
|-----------------------------|----|---------|
| En attente d'action (⚡) | 1 | 3 |
| En cours de traitement (🔄) | 1 | 0 |
| Traitement validé (✅) | 12 | 7 |
| Une erreur est survenue (❌) | 0 | 3 |

| Cas | Échantillon | Type d'analyse | Date | Culture | Tissu | Run | QC | Résultats | logs | Éditer | Supprimer | Notes |
|-------|-------------|----------------|------------|---------|----------|------------|--------------|-------------------|------|--------|-----------|-------------|
| 30260 | 1 | Bact_NanoMiSeq | 2023-07-21 | Brocoli | Feuilles | MS-10-2023 | QC : Terminé | Analyse : Erreur | logs | Editer | Supprimer | Commentaire |
| 30260 | 2 | Bact_MiSeq | 2023-07-21 | Brocoli | Feuilles | MS-10-2023 | QC : Terminé | Analyse : Erreur | logs | Editer | Supprimer | Commentaire |
| 30485 | 1 | Champ ITS | 2023-06-04 | Chou | Racines | MS-10-2023 | QC : Terminé | Analyse : Terminé | logs | Editer | Supprimer | Commentaire |
| 30485 | 2 | Champ ITS | 2023-06-04 | Chou | Racines | MS-10-2023 | QC : Terminé | Analyse : Terminé | logs | Editer | Supprimer | Commentaire |

Figure 7. Menu « Accueil » de l'application PhytoSHD

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

Le second menu « Saisie des données » (Figure 8) est l'interface d'importation des fichiers bruts provenant du SHD (fichiers FASTQ). Il est possible d'importer individuellement ou en groupe les fichiers et les métadonnées relatives à un cas. Ces métadonnées ou champs d'intérêt ont été sélectionnés à la suite de rencontres avec l'équipe du LEDP et ils sont intégrés à la base de données de l'application PhytoSHD. Pour simplifier l'importation de groupe de fichiers, un fichier Excel avec des instructions décrivant les étapes à réaliser permet d'importer les informations en une étape dans l'application.

Figure 8. Menu « Saisie des données » de l'application PhytoSHD

Dès leurs importations, les fichiers *fastq* sont soumis à une évaluation de la qualité en utilisant l'outil « FastQC ». À la fin de l'analyse, un rapport est produit pour chaque cas et l'utilisateur de l'application web sera en mesure d'évaluer leurs qualités. Comme les critères d'évaluation de la qualité diffèrent entre le NextSeq (Figure 9) et le MinION (Figure 10) des balises d'interprétation seront présentées aux utilisateurs.

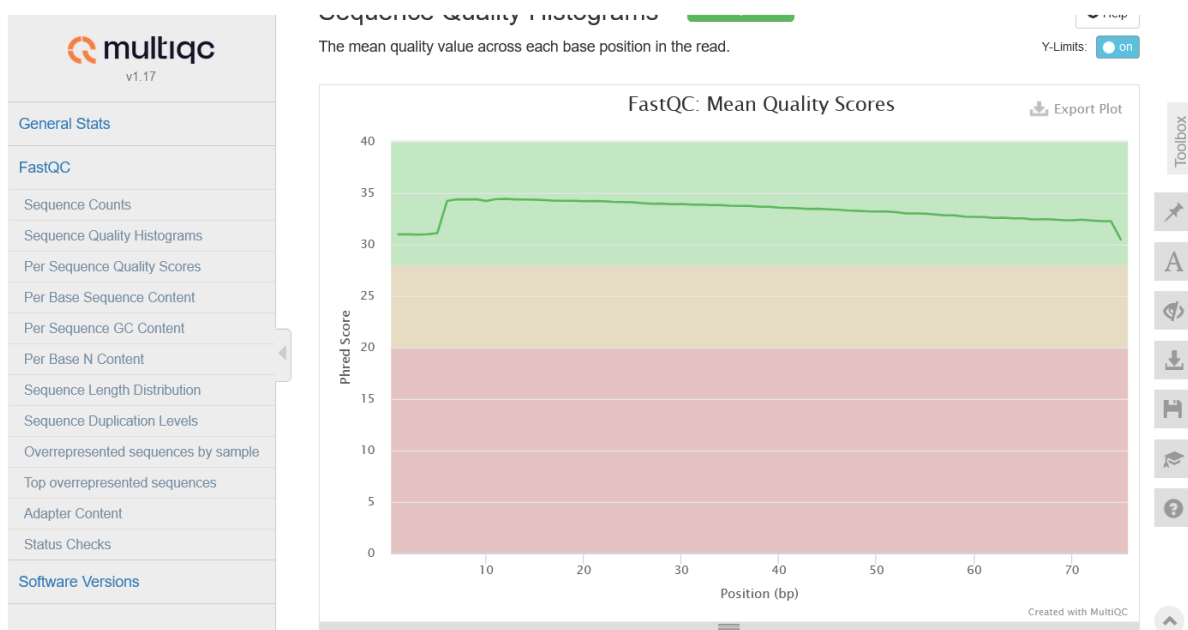


Figure 9. Rapport d'évaluation de la qualité d'un séquençage Illumina NextSeq produit par FastQC

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

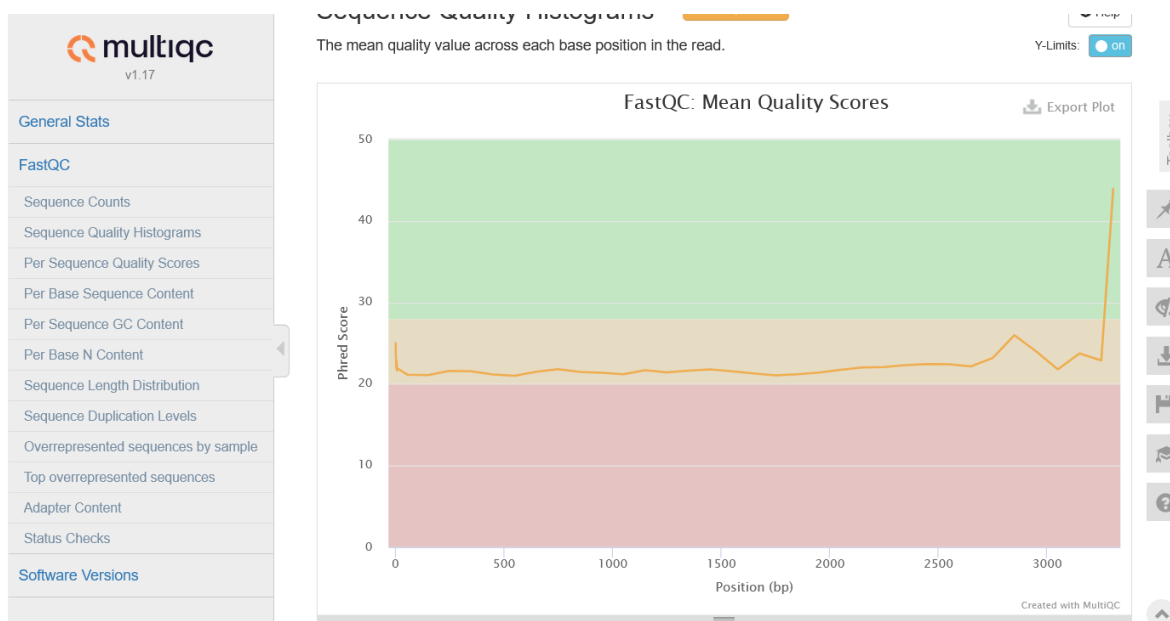


Figure 10. Rapport d'évaluation de la qualité d'un séquençage Nanopore MinION produit par FastQC

Une fois les fichiers *fastq* importés dans l'application web et leur qualité validée, l'utilisateur peut se rendre dans le menu « Traitement des données » (Figure 11). Dans ce menu, il peut sélectionner les cas à démarrer en analyse. Un récapitulatif en bas du menu permet de facilement réviser les tâches qui ont été lancées. Cela permet à l'utilisateur de prioriser certaines analyses. En effet, l'ordonnanceur traitera dans l'ordre les tâches qui lui sont soumises. Ainsi un dossier urgent peut être démarré en premier.

2

Récapitulatif de l'analyse

| cas | echantillon | date | methodeseq | runname | culture | tissus | typeanalyse | version | status | |
|---|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|----------|
| <input type="checkbox"/> filter data... | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | <input type="checkbox"/> | |
| ✕ | ONT2_test | 1 | 2023-07-21 | MinION | ONT2 | Fraise | Feuilles | Virus_Fraise_ONT | V1 | En cours |
| ✕ | NS2_test | 1 | 2023-06-21 | NextSeq | NS2 | Framboise | Feuilles | Virus_Framboise_Illumina | V1 | En cours |

Lancer les analyses

Figure 11. Échantillons sélectionnés dans le menu « Traitement des données » de l'application PhytoSHD.

Le dernier menu, « Visualisation des résultats », permet une visualisation claire et rapide des résultats de l'analyse pour un cas. En haut de la page, les métadonnées pertinentes au cas sont présentées (Figure 12). En dessous, la liste des virus détectés dans l'échantillon est présentée. La liste est triée selon le nombre de fragments (*reads*) viraux détectés. Rapidement l'usager peut voir les virus détectés, le nombre de *reads* positifs pour ce virus dans l'échantillon et le résultat du **modèle prédictif** (discuté à la sous-section 3.5). Il est possible de cliquer sur le nom d'un virus pour visualiser plus d'informations sur les paramètres de détection du virus, telles que la taille du génome et la couverture virale (Figure 13).

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

| Cas | Echantillon | Run | Tissus | Analyse | Base de données | Date de prélèvement |
|---|-------------|-------------------|---------|--------------------------|-----------------|----------------------|
| ONT2_test | 1 | ONT2 | Feuille | Virus_Fraise_ONT | CIMDEC19010v2 | 2023-07-21 |
| Séquences soumises | | Séquences post-Qc | | Séquences sans-hôte | | Séquences virales |
| 261667 | | 255692 | | 10966 | | 2261 |
| 12280_Tomato_ringspot_virus | | | | Séquences uniques : 2201 | | Prédiction : POSITIF |
| 12187_Strawberry_mild_yellow_edge_virus | | | | Séquences uniques : 60 | | Prédiction : NEGATIF |

Figure 12. Exemple de visualisation des résultats dans le menu « Résultats » de l'application PhytoSHD

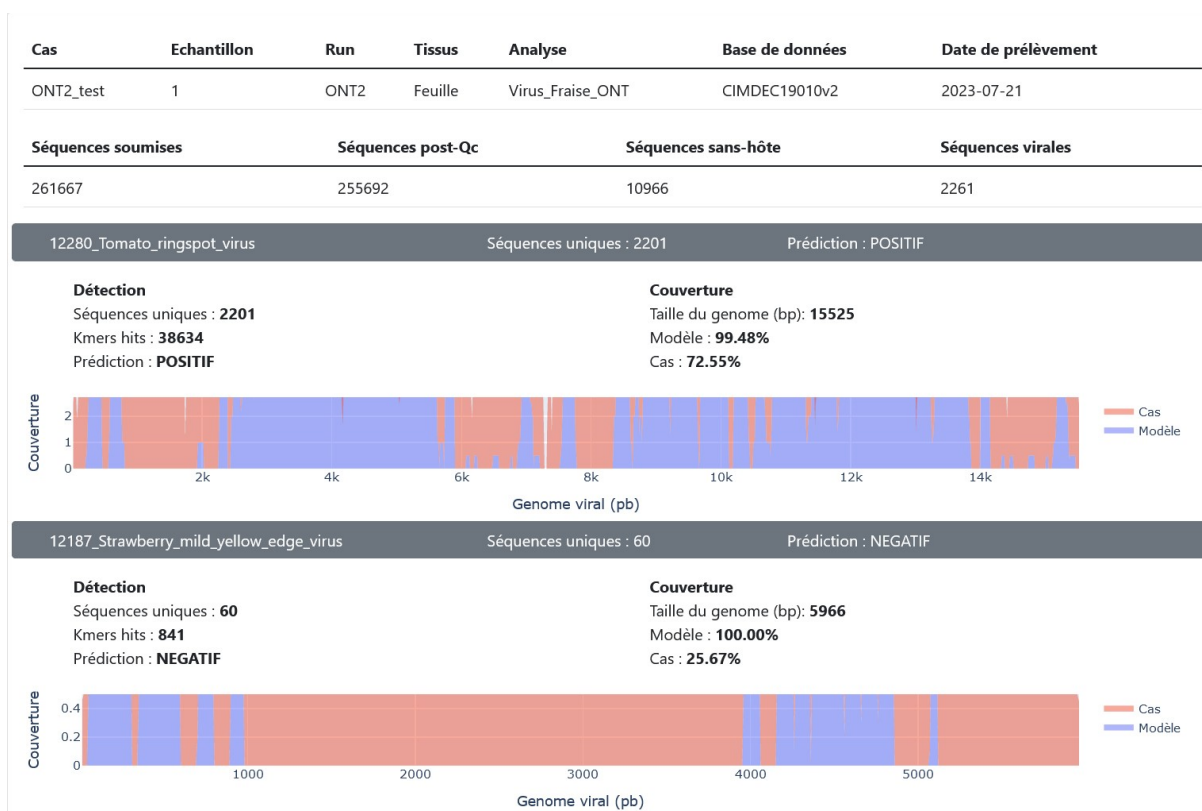


Figure 13. Exemple de visualisation avancée des résultats dans le menu « Résultats » de l'application PhytoSHD

3.6. Modèle prédictif

Pour assister l'utilisateur dans le choix du diagnostic à rendre, un modèle prédictif utilisant l'intelligence artificielle a été intégré à l'application PhytoSHD pour estimer si un cas est positif, en tenant en compte de différents paramètres (ex. le type de séquençage, le nombre de reads détectés, la couverture de la signature virale, le nombre de signatures virales ou le nombre total de reads). Le modèle prédictif pour le séquençage Illumina offre un taux de prédiction de 90.2% (Tableau 3). Le taux de prédiction est calculé par la somme des vrais négatifs et des vrais positifs.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

Tableau 3. Évaluation du modèle prédictif pour le séquençage sur la plateforme Illumina

| Diagnostic | Prédiction | |
|------------|------------|---------|
| | Négatif | Positif |
| Négatif | 47.1% | 2.0 % |
| Positif | 7.8% | 43.1% |

Pour le séquençage sur la plateforme Nanopore, le taux de prédiction est de 68.2% (Tableau 4). Ce taux plus faible est causé par une plus grande proportion de faux négatifs. Une explication possible est que le changement de chimie à la fin du projet (V9 pour V10) a rendu plus difficile la modélisation des détections. Il nous sera possible de fournir une mise à jour du modèle prédictif basé pour MinION une fois que le LEDP aura effectué un nombre plus conséquent d'analyses utilisant la nouvelle chimie Nanopore.

Tableau 4. Évaluation du modèle prédictif pour le séquençage sur la plateforme Nanopore

| Diagnostic | Prédiction | |
|------------|------------|---------|
| | Négatif | Positif |
| Négatif | 27.3% | 4.6% |
| Positif | 27.3% | 40.9% |

3.7. Analyse économique

L'analyse économique a été réalisée en deux phases en collaboration avec le LEDP. La première étape a été de construire une grille tarifaire permettant d'estimer les coûts en consommables et en temps humain pour la réalisation des analyses (coût de base). Une certaine économie de volume peut être faite selon le nombre d'échantillons analysés en même temps. Ainsi, nous avons déterminé deux catégories de coût selon le volume d'échantillons à analyser. La base Faible s'adapte à un contexte de faible réception d'échantillons (basse saison) et la base Moyenne s'adapte à un contexte de plus forte réception d'échantillon (haute saison) ou d'un dépistage annuel. Un descriptif des bases de coût est présenté en annexe (analyse_eco_CIMDEC). Les coûts ont été déterminés en janvier 2023.

La seconde phase a été d'établir un comparatif entre les chemins diagnostiques des 50 cas reçus par le LEDP (voir section 3.2) dans le cadre du projet. Pour le SHD nous avons choisi le chemin diagnostique suivant : PREP-Mol; EXT-ARN-Plant; NextSeq-Virus ou MinION-Virus. Les coûts des différents modules sont présentés à l'annexe 8. Le Tableau 5 présente les coûts en matériel arrondis au dollar (\$CAD) en haute saison (base Moyenne) qui ont servi au comparatif des différentes technologies d'analyse. On remarque que la détection de virus chez la fraise est généralement moins chère que chez la framboise vu le plus petit nombre de détections à réaliser. Le coût de la détection par SHD est plus élevé que la moyenne des coûts de détection par les méthodes actuelles du LEDP. Bien que plus cher, ces approches ont des avantages face aux méthodes actuelles en permettant la détection d'un plus large éventail de virus en même temps et la possibilité de retravailler d'anciens cas lors de la découverte de nouveaux virus. De plus, cette approche de comparaison des coûts d'utilisation des méthodes conventionnelles par rapport aux méthodes SHD exclut les coûts associés à la gestion des témoins qui peuvent être importants (selon pour les approches actuelles du LEDP (RT-PCR ou ELISA)).

Les coûts mentionnés sont adaptés aux contraintes du LEDP et peuvent différer des coûts de mise au point ou d'opération au laboratoire d'écologie microbienne de l'IRDA (LEM). Comme

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

toutes les détections de virus de pomme de terre ont été faites par le LEM, les coûts n'ont pas été comptabilisés dans l'analyse économique.

Tableau 5. Coût matériel pour les différentes technologies d'analyse (\$CAD) arrondi au dollar en haute saison (base Moyenne).

| Culture | Actuelles (LEDP) | SHD-NextSeq | SHD-MinION |
|-----------|------------------|-------------|------------|
| Fraise | 57\$ | 117\$ | 99\$ |
| Framboise | 84\$ | 117\$ | 99\$ |

Il est plus difficile d'évaluer le coût humain car les différentes technologies ont des contraintes particulières. Ainsi les méthodes actuelles du LEDP sont généralement réalisées en quelques jours mais les multiples protocoles (RT-PCR, ELISA) demandent l'utilisation constante de ressources humaines. Le SHD-MinION est un peu moins rapide que les méthodes actuelles du LEDP mais demande des ressources humaines par intermittence. Cela s'explique par un protocole composé de phases de purification suivi de phases d'automatisation (QIAcube, PCR, Séquençage) ne demandant pas de supervision humaine. À l'opposé, le SHD-NextSeq demande peu de manipulations de la part du personnel du LEDP mais a des délais plus longs pour la réception des résultats dû à la sous-traitance du séquençage à une plateforme externe (IRIC).

3.8. Tutoriel et protocoles

Dans la version initiale de l'application PhytoSHD, il avait été prévu d'inclure un tutoriel à même l'application pour expliquer les étapes d'importation, de traitement et de visualisation des données. Dans la version finale, qui sera livrée au LEDP, les différentes étapes ont été simplifiées pour rendre l'utilisation de l'application plus intuitive. Pour ce faire, il a été décidé d'inclure les instructions d'importation dans le modèle de fichier Excel qui sera utilisé pour l'importation des données. Ce fichier est préformaté pour être utilisé en partenariat avec l'application PhytoSHD. De plus, nous avons développé un petit programme (FastqFinder-ONT.exe) pour accélérer la saisie des noms de fichiers FASTQ dans le fichier Excel (Figure 14). Un guide de démarrage rapide sous forme de capsule vidéo décrivant les procédures de préparation, d'importation et de visualisation des données sera transmis au personnel du LEDP (Annexe 7).

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

Figure 14. Application « FastqFinder-ONT » pour accélérer la saisie des noms de fichiers FASTQ

4. ACTIVITÉS DE DIFFUSION ET TRANSFERT AUX UTILISATEURS

4.1. Diffusions

La liste des diffusions effectuées dans le projet est présentée dans le tableau suivant.

Tableau 6. Activités de diffusion

| Nom de l'activité | Type de diffusion | Événement | Date | Public | Réf. |
|------------------------|------------------------|-----------|---------|----------------------------|------|
| Les nouvelles fraîches | Présentation du projet | Journal | 10/2019 | Agriculteur, agronomes | A9-1 |
| RAP-Petits fruits | Présentation du projet | RAP 2019 | 11/2019 | Professionnels, agronomes | A9-2 |
| SPPQ | Présentation orale | SPPQ 2021 | 09/2021 | Chercheurs, professionnels | A9-3 |
| RAP-Petits fruits | Présentation orale | RAP 2021 | 11/2021 | Professionnels, agronomes | A9-4 |
| SPPQ | Présentation orale | SPPQ 2024 | 06/2024 | Chercheurs, professionnels | A9-5 |

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

4.2. Transfert au LEDP

Durant le projet, plusieurs activités de formations ont été réalisées pour :

1. Nous avons organisé des formations en 2022 et 2023 pour présenter les protocoles en détail et expliquer le flux de travail pour réaliser un séquençage.
2. Nous avons formé le personnel et produit des capsules vidéo du LEDP pour l'utilisation de la plateforme web PhytoSHD pour importer, traiter et consulter les résultats obtenus à partir des données de séquençage à haut débit.
3. Nous avons accompagné le personnel du LEDP lors de l'utilisation des protocoles pour la préparation de librairies et le séquençage sur la plateforme MinION.
4. Lors de la dernière phase du projet, un suivi serré a été réalisé pour identifier et corriger les irritants dans les protocoles d'extraction, de séquençage et dans l'utilisation de l'application web PhytoSHD.

5. PERSPECTIVES

Séquençage MinION sur plateforme Nanopore

Le séquençage Nanopore est une technologie en constante évolution avec un rythme itératif rapide. Le projet a couvert trois générations de cellule de séquençage. Le but de ces innovations étant de donner un meilleur rendement et des séquences de meilleure qualité. Il est possible qu'à moyen terme de nouvelles cellules soient développées, offrant de meilleurs résultats. Il y a une période de chevauchement entre l'arrivée d'une nouvelle cellule et la fin d'une ancienne pour permettre de faire la transition. De plus, les clients sont avertis d'avance par ONT (Oxford Nanopore Technologies, UK). Le module de détection PhytoV est assez robuste pour s'adapter à un changement de cellule. Il est possible, comme expliqué plus haut, que les modèles prédictifs doivent être réentraînés suivant un changement de cellule.

Enrichissement des ARN viraux

Le principal avantage de la détection par SHD est le nombre de virus qu'il est possible de tester en même temps et la possibilité de retester d'anciens échantillons advenant la découverte d'un nouveau virus. La limite est que cette approche « aléatoire » demande une profondeur de séquençage (nombre de fragment séquencé) importante pour être certain de détecter les fragments viraux dans la soupe de fragments d'ARN de l'hôte. Ainsi, trois principales méthodes de concentration des extraits sont généralement utilisées en biologie moléculaire : 1. L'amplification de patrons viraux; 2. La déplétion de l'ARN ribosomale; 3. L'enrichissement de l'ARN double brins.

L'amplification de patrons viraux est surtout utilisée lorsqu'il y a seulement un virus d'intérêt et qu'on souhaite suivre l'évolution de ses mutations (ie. SARS-CoV-2). La déplétion de l'ARN ribosomale est prometteuse car selon nos analyses, l'ARN ribosomale peut représenter entre 30 et 70% de l'ARN total extrait. Malheureusement, la majorité des méthodes de déplétion de l'ARN ribosomale sont complexes et dispendieuses. Dans les dernières années, certaines méthodes plus accessibles ont été développées et il serait intéressant d'évaluer si ces méthodes peuvent être adaptées et si elles influencent la détection des virus d'intérêts. Pour finir, l'enrichissement de l'ARN double brins est très spécifique pour la détection des virus. Il existe des méthodes rapides et peu coûteuses qui pourraient améliorer la détection des virus par séquençage.

NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS

19-010-2.2-C-IRDA

L'ajout d'une étape d'enrichissement des ARN viraux rapide et peu coûteuse serait sûrement un gain au protocole actuel. Le défi étant de transférer des méthodes principalement utilisées dans une logique de recherche à une logique de service.

Base de référence pour d'autres hôtes

Pour le moment, l'application a été développée pour trois hôtes (ie. Fraise, Framboise et Pomme de terre) mais ce ne sont pas les seules cultures qui présentent des défis de détection des virus au Québec. Nous n'avons qu'à penser à la culture de la Tomate ou du Bleuets. Le bleuets est particulièrement intéressant vu la nature pérenne des plants et le temps nécessaire pour remplacer des plants détruits par un virus. Notre outil est flexible et il serait possible d'élaborer et de tester sa sensibilité avec de nouvelles cultures.

6. POINT DE CONTACT POUR INFORMATION

Richard Hogue, Ph.D., responsable du Laboratoire d'écologie microbienne (LEM) et responsable du projet PV19-010-2.2-C-IRDA (référence IRDA projet #400128)
richard.hogue@irda.qc.ca

7. REMERCIEMENTS AUX PARTENAIRES FINANCIERS

Ce projet a été réalisé en vertu du sous-volet 2.2 du programme Prime-Vert 2018-2023 et il a bénéficié d'une aide financière du ministère de l'Agriculture, des Pêcheries et de l'Alimentation du Québec (MAPAQ).

La réalisation du projet a également été rendu possible par la contribution nature de plusieurs partenaires impliqués principalement à la collecte des échantillons par les partenaires du milieu horticole (Ferme Onésime Pouliot inc., Services Agriexpert, Ferme horticole Gagnon, Agriconseils, Fertior, Club conseil Profit-eau-sol et des agronomes du MAPAQ) et au développement de l'OAD par le personnel des équipes du Dr Arnaud Droit du CRCHQ – U. Laval et du Dr. Richard Hogue du Laboratoire d'analyse biologique de l'IRDA.

8. RÉFÉRENCES

- [1] Eccles, D. (2023). *PCR cDNA Barcoding Kit Protocol*. Hugh Green Cytometry Centre.
www.gringene.org/nanopore/protocols/DAE_PCR_cDNA_Barcoding_Kit_Protocol_2023-Jun-23.pdf
- [2] ONT (2023). *PCR-cDNA Barcoding Kit (SQK-PCB111.24)*. Library Prep Protocols.
community.nanoporetech.com/attachments/11541
- [3] ONT (2023). *Chemistry Technical Document. Methodology best practice*.
<https://community.nanoporetech.com/attachments/11631/download>
- [4] Mikalsen, A. B., Nilsen, P., Frøystad-Saugen, M., Lindmo, K., Eliassen, T. M., Rode, M., & Evensen, Ø. (2014). *Characterization of a novel calicivirus causing systemic infection in*

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

Atlantic salmon (Salmo salar L.): proposal for a new genus of Caliciviridae. PloS one, 9(9), e107132.

[5] Shifu Chen, Changshou He, Yingqiang Li, Zhicheng Li, Charles E Melancon III. *A Computational Toolset for Rapid Identification of SARS-CoV-2, other Viruses, and Microorganisms from Sequencing Data.* bioRxiv 2020.05.12.092163; doi: <https://doi.org/10.1101/2020.05.12.092163>

[6] Mihara, T., Nishimura, Y., Shimizu, Y., Nishiyama, H., Yoshikawa, G., Uehara, H., Hingamp, P., Goto, S., and Ogata, H.; *Linking virus genomes with host taxonomy.* Viruses 8, 66 doi:10.3390/v8030066 (2016).

**NOUVELLE APPROCHE MOLÉCULAIRE POUR LA DÉTECTION SIMULTANÉE DES
VIRUS PATHOGÈNES AUX FRAMBOISIERS ET AUX FRAISIERS
19-010-2.2-C-IRDA**

9. ANNEXES

Liste des protocoles, guide et tableaux transmis au LEDP :

- 1- Protocole d'extraction d'ARN:
Pro_EXT_CTAB_ARN_2024.01.pdf (chimique)
Pro_EXT_QiagenARN_2024.01.pdf (affinité)
- 2- Protocole de préparation de librairie pour MinION :
Pro_LIB_MinION_2024.01.pdf (Cellule R9/R10; chimie v10)
- 3- Protocole de préparation de séquençage pour MinION :
Pro_SEQ_MinION_2024.01.pdf
- 4- Contrôle interne :
Virus_Saumon.fasta
- 5- Protocole de détection de virus par RT-PCR :
RT-PCR ONE STEP QIAGEN_virus fraise.xls (fraise)
RT-PCR ONE STEP QIAGEN_virus framboise.xls (framboise)
RT-PCR ONE STEP QIAGEN_virus pdt.xls (pomme de terre)
ListeAmorces.xlsx
- 6- Liste des virus et base de données :
VirusOAD.xlsx
VirusHomologie.xlsx
VirusProjet.xlsx
- 7- Capsules vidéo de démarrage rapide PhytoSHD :
1_PréparationSéquence.mkv
2_ImportationSéquence.mkv
3_VisualisationRésultat.mkv
- 8- Analyse économique :
analyse_eco_CIMDEC.xlsx
- 9- Activités de diffusion :
ActivitésDiffusion.pdf
- 10-Virus détectés pour chaque culture
VirusCulture.docx

L'ensemble de ces fichiers électroniques a été partagé avec le LEDP.