ANNÉE 2015

THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention Biologie

Ecole doctorale Vie Agro Santé

présentée par

Flore Ballaud

préparée à l'unité de recherche 6553 Ecobio
Ecosystèmes, Biodiversité et Evolution
Université Rennes 1

Dynamique spatio-temporelle des communautés virales et microbiennes des tourbières à *Sphagnum*

Thèse soutenue à Rennes
le 17 décembre 2015

devant le jury composé de :

Jean-Christophe CLEMENT
Maître de Conférence-Grenoble / *rapporteur*

Gaël ERAUSO
Professeur-Aix-Marseille / *rapporteur*

Fatima LAGGOUN
Directrice de Recherche-Orléans / *examinatrice*

Daniel GILBERT
Professeur-Besançon / *examinateur*

Télesphore SIME-NGANDO
Directeur de Recherche-Clermont-Ferrand / *examinateur*

Alexis DUFRESNE
Chargé de Recherche-Rennes / *examinateur*

André-Jean FRANCEZ
Maître de Conférence-Rennes / *directeur de thèse*

Achim QUAISER
Maître de Conférence-Rennes / *directeur de thèse*

Résumé

Les tourbières couvrent 3 % des surfaces continentales et jouent un rôle important dans le cycle du carbone en stockant le tiers du carbone des sols. L'accumulation de tourbe est liée au déséquilibre production primaire/décomposition du à une activité microbienne limitée par les conditions environnementales. L'infection et la lyse virale ont un impact sur la diversité et l'activité des communautés microbiennes, et influent sur le cycle du carbone. Cependant, le fonctionnement de ce compartiment viral n'avait jamais été pris en compte dans les études de fonctionnement des tourbières. Le but de ce travail de thèse était d'analyser et comprendre la dynamique spatiale et temporelle de l'abondance et de la diversité virale des tourbières à *Sphagnum*. L'analyse de l'abondance virale et procaryote et de 12 metaviromes en lien avec la physico-chimie d'une tourbière tempérée en France montre une forte variation saisonnière des communautés virales. Cette variation semble très liée aux conditions environnementales générées par la fluctuation de la nappe d'eau. Dans cette même tourbière, l'analyse de la diversité taxonomique et fonctionnelle des communautés de microorganismes présents (métagénomes) et métaboliquement actifs (métatranscriptomes) indique que la structure taxonomique est différente entre les des deux principaux stades, le fen et le bog, mais que ces communautés présentent une diversité fonctionnelle similaire, dont l'expression est liée aux changements des conditions environnementales avec la profondeur. L'abondance des particules virales étudiées dans 5 tourbières à *Sphagnum* réparties en Finlande, au Canada, en France, et sur l'île subantarctique d'Amsterdam varie fortement avec les sites. L'analyse de la diversité virale de la matrice et de l'eau de tourbe du Canada et de Finlande montre que la diversité virale est structurée par le site, puis le stade dynamique, puis la profondeur, avec un rôle important de la saturation en eau au niveau du site. Ces résultats valident le fonctionnement proposé du compartiment viral et de la communauté d'hôtes procaryotes. Ces connaissances ont été utilisées pour analyser le fonctionnement du compartiment microbien de tourbières à *Sphagnum* soumises à des perturbations d'origine anthropique. Les 31 métaviromes produits pour cette thèse constituent l'une des plus grandes bases de données sur la diversité virale des écosystèmes alors que la diversité virale des sols n'avait presque jamais été étudiée auparavant.

Mots clés : virus, particules virales, procaryotes, abondance, diversité, activité, interactions, tourbière, *Sphagnum*, fen, bog, dynamique spatiale, dynamique temporelle, conditions environnementales, métavirome, métagénome, métatranscriptome

---

Abstract

Peatlands cover 3 % of the continental surfaces but represent up to a third of the soil carbon stock. Peat accumulation results from the imbalance between primary production and decomposition due to the limitation of the prokaryote activity caused by the environmental conditions. Viral infection and lysis impact the diversity and the activity of the microbial communities and influence the carbon cycle. However, the functioning of the viral compartment had never been taken into account in peatlands. The aim of this thesis was to gain knowledge about the spatio-temporal dynamic of viral abundance and diversity in *Sphagnum*-dominated peatlands. Spatio-temporal analysis of viral and prokaryote abundance and of 12 metaviromes (viral diversity) in relation to the physico-chemical features in a temperate *Sphagnum*-dominated peatland in France revealed the high seasonal variability of the viral communities. This dynamic appeared mainly related to the environmental conditions shaped by the fluctuation of the water-table level. In the same peatland the taxonomic diversity of the present microorganisms (metagenomes) differed between the fen and the bog, but these communities present a similar functional diversity, which expression in selected in the same way in the two dynamic stages, in relation to depth-related environmental conditions. Viral abundance analyzed in 5 *Sphagnum*-dominated peatlands from Finland, Canada, France and subantarctic Amsterdam Isle presented a high geographical variability. Investigation of the diversity of the viral communities from the peat matrix and the pore-water in Finland and Canada emphasized the structuration of the viral communities by the site, then the dynamic stage, and finally depth. These results confirm the first hypotheses about the functioning of the viral compartment depending on environmental conditions and prokaryote activity. Effects of human-derived disturbances on viral ecology in peatlands were investigated based on this knowledge. While soil viral diversity was poorly documented at the start of this thesis, the collection of 31 metaviromes from *Sphagnum*-dominated peatlands produced for this project represents the second largest dataset representing the viral diversity from environmental samples.

Key words: viruses, viral particles, prokaryotes, abundance, diversity, activity, interactions, peatland, *Sphagnum*, fen, bog, spatial dynamic, temporal dynamic, environmental conditions, metavirome, metagenome, metatranscriptome,

Remerciements

C'est après avoir rédigé ce manuscrit et soutenu ma thèse que je prends le temps de remercier, enfin, ceux qui ont contribué à cette aventure scientifique. Ce projet a duré trois ans, et il n'aurait pu aboutir sans l'aide, le soutien et la contribution d'un nombre impressionnant de personnes. Prendre conscience que l'on a été aussi entouré pendant toutes ces années est au moins aussi émouvant que d'avoir réussi à obtenir ce dernier diplôme.

Je tiens donc naturellement et en premier lieu à remercier mes directeurs de thèse, André-Jean et Achim pour le rôle important qu'ils ont joué au cours de ces trois années et demies. Merci d'avoir toujours pris en compte mes idées, de m'avoir aidé à les mûrir, les faire aboutir ou les abandonner. Merci d'avoir eu la patience de me recadrer quand j'étais déboussolée ou entêtée, merci d'avoir eu confiance et d'avoir beaucoup donné de votre temps dans la réflexion puis la rédaction de ce travail. Je souhaite à tous les doctorants de bénéficier du même soutien intellectuel et humain que celui que vous m'avez apporté. Ne changez rien.

Je remercie également Alexis Dufresne pour son implication au démarrage de ce projet : merci pour tes questions pointues, un peu déstabilisantes parfois mais qui font beaucoup avancer. Et merci aussi de continuer à poser des questions tant que tu n'es pas convaincu ! Je remercie aussi les rapporteurs, Gaël Erauso et Jean-Christophe Clément pour avoir pris le temps d'examiner le manuscrit, ainsi que les examinateurs : Fatima Laggoun, Télésphore Sime-Ngando, Daniel Gilbert et Alexis Dufresne pour être venus évaluer ma soutenance, ce fut l'occasion d'une discussion très enrichissante.

Je tiens ensuite à remercier toutes les personnes qui ont investi une partie de leur précieux temps dans ce projet, que cela soit sur le terrain ou au laboratoire.

Un grand merci tout d'abord à Jonathan Colombet et Télésphore Sime-Ngando qui ont permis de poser les premiers jalons de ce travail. Merci également à Harri Vasander, Line Rochefort, Claire Boismenu, Brigitte Vimard, Elena Lasphina, Yarouslav Solomin, Fatima Laggoun, Franck Le Moing, Daniel Gilbert, Marc Lebouvier et Nathalie Van Der Putten pour l'organisation et leur implication dans l'échantillonnage des différentes tourbières : c'était formidable de recevoir ces morceaux de tourbe de sites si éloignés, j'espère avoir la chance de tous les découvrir un jour de mes propres yeux. Merci aussi à toute l'équipe de la Yugra State University et de la station de Mukhrino (Evgueni, Nina, Ilia…) qui nous ont aidés lors de notre campagne de terrain dans la belle Sibérie.

Un immense merci à Valérie, Sandra, Isabelle, Fabienne et Tifenn pour leur inestimable efficacité à gérer la partie administrative et logistique de la recherche, et le tout avec le sourire. Merci à Badia de l'école doctorale qui leur fait concurrence en termes de rapidité et d'amabilité. Et merci à Valérie pour tous les articles qu'elle a réussi à trouver !

Merci à Nathalie Le Bris et Frédérique Pallois pour les analyses physico-chimiques, pour leur gentillesse et leur patience pour nous apprendre à utiliser correctement le centre commun, et pour les sympathiques pauses café au bout du couloir. Merci à Sophie, Alexandra et Patricia de la plateforme de génomique pour leur investissement dans le séquençage de nos petits virus. Merci enfin à Agnès Burel pour les grand moments d'émotion à la plateforme de microscopie électronique quand j'ai pu voir mes premiers phages de « jus » de tourbe !

Une thèse ne serait pas non plus la même sans regard extérieur. Merci à Claire Geslin et Télésphore Sime-Ngando pour leur participation à mes comités de thèse, ainsi qu'un immense merci à Christine Baysse, ma tutrice de thèse, qui s'est beaucoup impliquée dans le suivi de l'avancement de mon travail et qui a pris le temps de me remotiver quand je doutais.

Merci aussi Christine pour les bons moments passés dans le cadre de mes deux missions d'enseignements : j'ai beaucoup appris à tes côtés, j'espère que les étudiants aussi ! J'en profite d'ailleurs pour remercier les équipes pédagogiques d'écologie et de microbiologie avec qui j'ai beaucoup aimé travailler.

Je remercie toute l'équipe RBPE, et plus particulièrement Cécile Monard, Philippe Vandenkoornhuyse et Françoise Binet, qui ont été disponibles et attentifs tant scientifiquement qu'humainement. Merci aussi à Ben Abbott pour les supers conversations sur les microorganismes du sol et surtout pour les corrections d'anglais !

J'en profite pour remercier Aude Ernoult, Françoise Burel et Rémi Duflot qui ont encadré mon stage de M1 à Ecobio. Par extension, je tiens d'ailleurs à remercier tous ceux qui, depuis la licence, m'ont formée à la recherche et ont pris grand soin de me motiver à poursuivre jusqu'en thèse : merci à Vincent Courdavault, Grégory Guirimand, Tahina Roi-des-Copépodes, et Olivier Glippa. Merci à Denis Poinsot pour sa confiance et ses conseils.

J'espère en ce sens ne pas avoir trop découragé Nadia Larabi, ma « petite » stagiaire durant ces deux mois de lysogénie et de cytomètre sur une expérience pas toujours évidente. Merci pour ton travail !

Je remercie aussi chaleureusement celles qui passent régulièrement apporter leur sourire et la propreté dans nos locaux : Virginie, Nadine et Armelle, merci pour tous ces joyeux moments !

Vient ensuite le moment tant attendu de remercier tous les compagnons d'infortune avec qui j'ai partagé les joies, les angoisses et les questions existentielles d'apprenti chercheur.

Un immense merci à tous les gens qui sont passés plus ou moins longtemps dans le chaleureux bureau 216 du bâtiment 14B. Merci à Steph Raquette et Marie pour m'avoir notamment aidée à parfaire ma culture musicale des 80s-90s. Un immense merci aussi à Lixou et Elo pour leur investissement dans la décoration de ce bureau, leur énergie et surtout leur amitié et toutes les bonnes choses qui vont avec : vivement qu'on ouvre notre labo ! Merci à Marion G., Nathan, Kévin, Amandine, Dyung et Nico qui ont chacun à leur manière contribué à la bonne ambiance générale dans ce bureau.

Merci à tous les charmants doctorants qui étaient disséminés dans les murs d'Ecobio au cours de ces trois années: merci à Maxime Dahirel pour toutes les discussions approfondies sur l'évolution, les stats, et pour les rédactions/validations/rappels de mails durant les deux années où nous avons géré l'Institution (#JournalClub) : désolée de ne pas avoir rédigé ce manuscrit dans le même style ☺. Merci à la génération précédente : Kévin, Cécile, Sarah, Mamadou, Lisa et Xavier. Merci aussi à la très sympathique délégation corse, constituée d'Anne-Antonnella. Merci aux doctorants-bientôt-docteurs de ma génération : Nathalie (merci pour les pauses cafés, les supers diners et la dégustation de produits réunionais) et Greg. Merci aussi et bon courage aux petits poussins de la génération d'après : ma « petite » Morgane, ma « cara Diana », Tiphaine des Kerguelen, Maud…et tous ceux que j'oublie ! Merci à Benoit G. et Stéphane Mahé pour les potées auvergnates et les concerts !

Je remercie aussi les joyeux lurons du département caillou qui jouxte notre bâtiment : Roman, Marie O., Youss, Gemma, Bob, la Boch', la Deul', la Bab', Réglisse, Dani-la-malice, Jérôme, Chrichri, Thomas, Caro, Sylvia, Jeanette (désolée, c'était la catégorie la plus adéquate), Pipo, Marie (désolée aussi pour la catégorie), Justine… qui ont su me faire apprécier les déjeuners au RU, les discussions de sédimento, de sismique ou la psychologie de Sylvain Tesson autour de bons produits de chez vous ou ramenés du terrain. Désolée pour ceux que j'oublie !

Un merci tout particulier à mes deux meilleurs compagnons de route : Quenoit Besnel pour son inestimable énergie positive et son enthousiasme communicatif et débordant qui ont été un vrai moteur, tout particulièrement quand nous rédigions nos manuscrits respectifs. Merci aussi à Bastien qui a apporté en Bretagne son soleil du Sud-Ouest et son art de la délicatesse (#grasdecanard, #sticksdemozza, #rugby) ainsi que ses idées farfelues mais au fond pas toujours si mauvaises.

Merci aussi à tous les compagnons du Master qui ont poursuivi ou non en thèse, mais avec qui on a continué à se poser de vraies questions. Merci à mes chères Alexia et Marion pour tous les

thés/bières/saucissons avec vous ainsi que pour tous les conseils de bricoloage, merci à Nico et Brieuc, et merci à tous les autres : Sandra, Jérém, Solenn, Romain, Val', Nolwenn, Floriane, et Simon. Un merci plus spécial à Pierrot pour m'avoir offert un magnifique lieu de retraite dans la capitale des Alpes lors de la rédaction de ce manuscrit.

J'aimerais aussi remercier ceux qui ne travaillent pas forcément dans la science mais qui ont contribué à me faire avancer dans ce travail et à nourrir ma réflexion: merci à Camille et Jojo pour les grands débats sur la vie en général, et un immense merci à Alexis et Elise pour avoir si souvent partagé ma passion du phytoplancton et des bonnes tables. Merci à Julien, sans qui je n'aurais peut-être pas entamé ce projet. Un merci tout particulier à ma Laurette et Damien qui ont toujours soutenu ce projet sous prétexte que c'était une étape nécessaire si jamais je voulais un prix (Ig-) Nobel.

Merci à toute ma famille pour leur soutien dans cette aventure : merci plus particulièrement à papa et maman pour leur confiance, et à mes frères pour leur curiosité (même si vous préfériez les copépodes et le requin pèlerin).

Merci enfin à Maël pour la mise en forme de la biblio. Merci aussi pour tes conseils pour l'achat de BDs et les sessions pêche sur le Maroni. Et merci éventuellement pour ton infaillible soutien, ta confiance et ta présence au cours et dans la poursuite de ce projet.

Table des matières

Introduction :

Etat de l’art et objectifs de la thèse

Introduction : Etat de l'art et objectifs de la thèse

1. Fonctionnement et intérêts conceptuels du modèle « tourbière à Sphaignes »

1.1. Qu'est-ce qu'une tourbière à Sphaignes et où les trouve-t-on ?

Les tourbières sont des écosystèmes de zones humides, accumulant un matériel organique appelé tourbe. Le processus de tourbification résulte d'un déséquilibre entre la production primaire de matière organique par les végétaux via la photosynthèse, et sa décomposition et sa minéralisation par les microorganismes. Ce double bilan excédentaire (eau et matière organique) caractérise l'ensemble des tourbières du Globe. Ainsi, il en existe de plusieurs types sous des climats et des latitudes variées. On distingue selon les zones biogéographiques et les biomes associés, des tourbières tropicales et des tourbières boréales. Les tourbières couvrent environ 400 millions d'hectares (3 % de la surface des écosystèmes terrestres) dont la plupart se situe dans l'hémisphère nord (350 millions). Dans les tourbières à Sphaignes, le déséquilibre production primaire/décomposition s'explique par la combinaison d'un climat froid et humide, de l'acidité du milieu, de la pauvreté en éléments nutritifs et de la saturation en eau, très défavorables à l'activité des microorganismes.

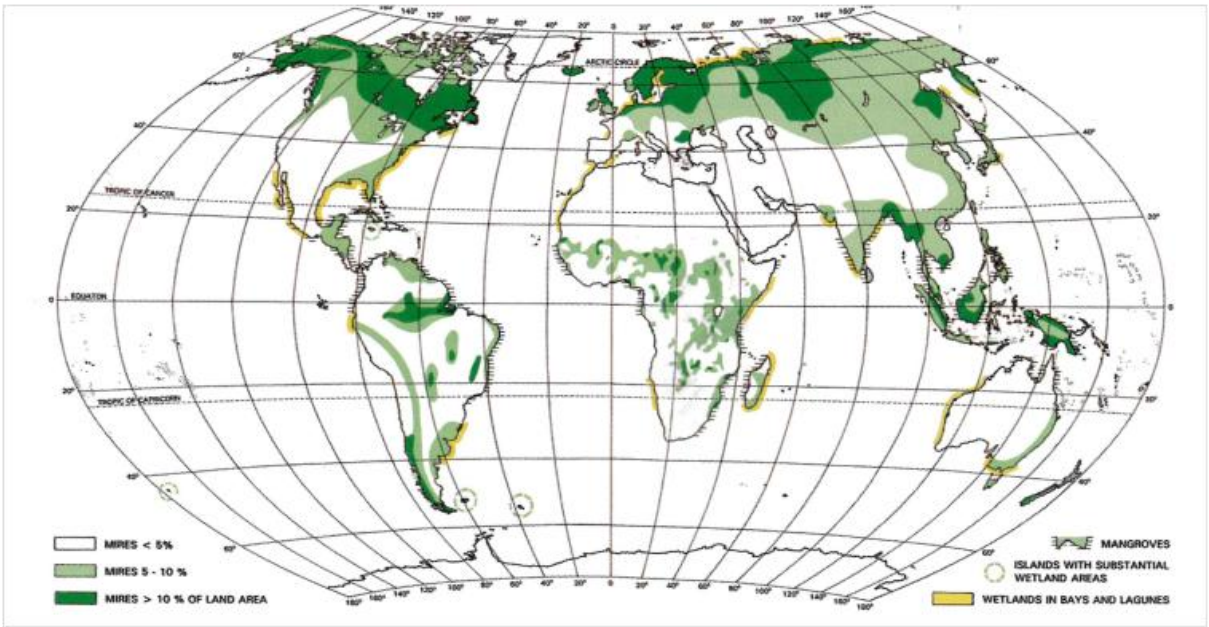


Figure 1 - Distribution mondiale des tourbières (vert foncé : surface > 10 % territoire ; vert clair : 5-10 % ; blanc : < 5% ; en jaune zones humides de baies et lagunes) (D'après Lappalainen 1996).

Géographiquement, ces tourbières sont de fait majoritairement situées dans les régions les plus septentrionales de l'hémisphère boréal, en Fenno-Scandinavie, Sibérie et Canada. Cependant, plus au Sud à la faveur de la présence de massifs montagneux et d'un climat suffisamment froid, elles ont pu se développer et se maintenir en région tempérée (Lappalainen, 1996 ; Figure 1). En France, les tourbières couvrent environ 100 000 ha et ont subi de fortes pressions anthropiques. Elles sont particulièrement bien représentées à l'étage montagnard du Jura et du Massif central. En plaine les tourbières à Hypnacées dominent (Francez et al., 1992; Manneville et al., 2006).

1.2. Comment se sont-elles formées ?

1.2.1. L'ingénierie des Sphaignes et la sphagnosphère

Les Sphaignes sont des Bryophytes à la morphologie et l'écologie bien particulières (Figure 2). D'un point de vue taxonomique, elles constituent le seul genre *(Sphagnum)* de la famille des Sphagnacées (ordre des Sphagnales). Comme toutes les Bryophytes, les Sphaignes sont dépourvues de tissus spécialisés. Elles croissent par leur apex (capitulum) et meurent par leur base. Le tissu cellulaire de la Sphaigne correspond à une mosaïque de cellules photosynthétiques, les chlorocystes, enchassées dans un réseau de cellules mortes, les hyalocystes, dont la morphologie "en outre" confère à ces cellules une très forte capacité de rétention en eau (Clymo and Hayward 1982).

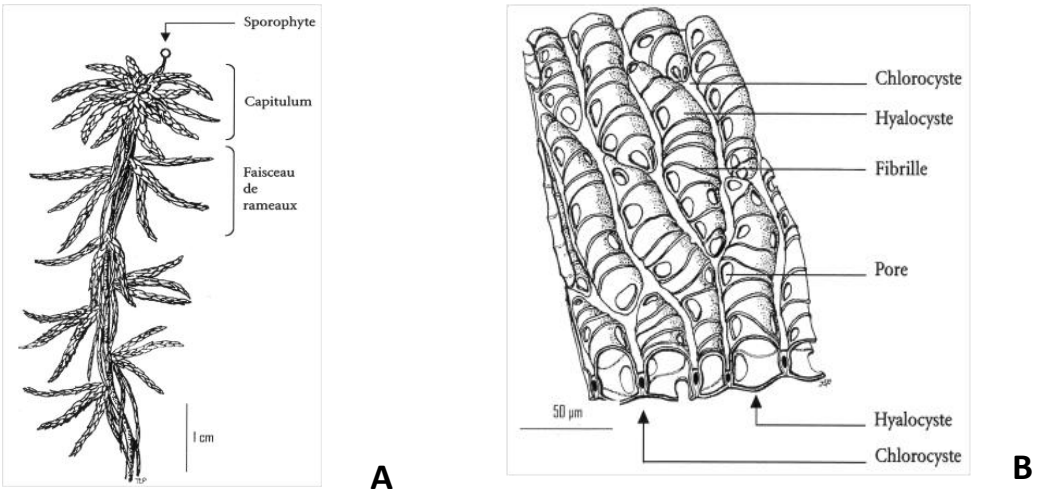


Figure 2 - Schéma d'une Sphaigne : A) rameaux et feuilles en faisceau et B) détail d'une feuille illustrant la disposition en alternance des chlorocystes et des hyalocystes (Dessins de Jean-Louis Polidori *in* Gauthier 2001).

Contrairement aux Mousses, le modèle morphologique global des Sphaignes est assez homogène et rend difficile la détermination des espèces (Gauthier, 2001). On en compterait moins de 150 espèces (Rydin and Jeglum, 2006).

Les Bryophytes constituent une composante fondamentale de nombreux écosystèmes aquatiques (mares, rivières, etc.) et terrestres (forêts, prairies humides, ...) et jouent un rôle

déterminant dans leur dynamique en tant que compartiment traversé par des flux. Ainsi, elles couvrent une surface suffisamment importante, elles contrôlent les cycles de l'eau et de chaleur, notamment en zone arctique (Blok et al., 2011). Elles abritent de nombreuses communautés de microorganismes (Bactéries, Champignons, Protozoaires) et une microfaune capable de s'enkyster lorsque les Bryophytes s'assèchent (Rotifères, Tardigrades, Nématodes, etc.) très diversifiées. Ainsi, Lindo et Gonzales (2010)(Figure 3) ont proposé le concept de 'Bryosphere', définie comme *"la combinaison des interactions complexes entre les parties vivantes et mortes et des organismes associés, constituant un système caractérisé par un couplage de processus de type 'above-belowground'*."

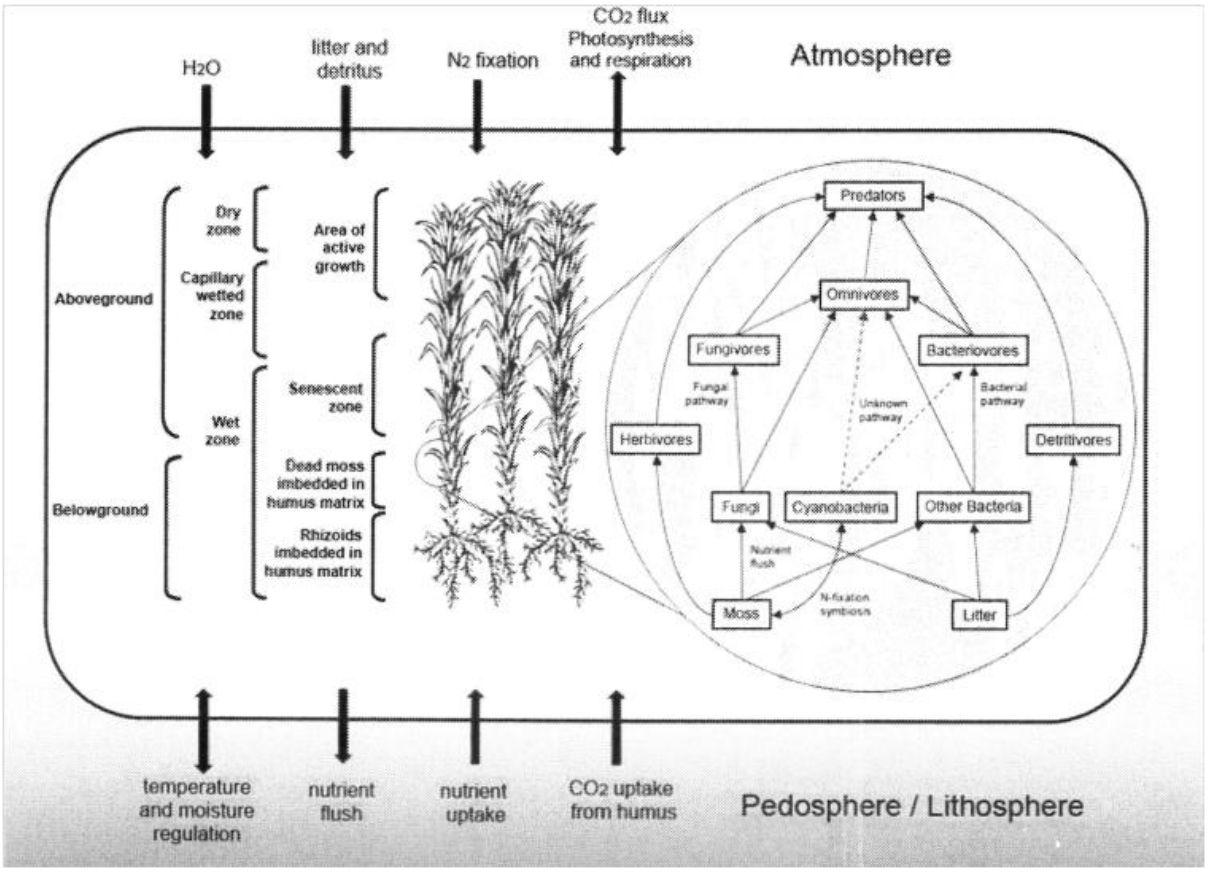


Figure 3 - Délimitation verticale (structure physique) et processus fonctionnels de la bryosphère (D'après Lindo and Gonzales 2010). Cette approche conceptuelle a été développée dans les tourbières et appliquée aux Sphaignes (concept de 'Sphagnosphère', Jassey, 2011).

Les Sphaignes contribuent à l'édification des tourbières et créent un environnement dans lequel de nombreux organismes se développent. Ce sont de véritables espèces ingénieures autogènes car elles transforment le milieu en modifiant l'environnement abiotique (pH, hydrologie, topographie) et biotique (successions végétales, microfaune), d'une structure physique à une autre (Jones et al., 1994).

Afin d'intégrer les spécificités biologiques et écologiques et les nombreuses interactions abiotiques et biotiques (Sphaignes-microorganismes en particulier via des activités enzymatiques), par rapport aux autres Bryophytes, Jassey (2011) a proposé le concept de 'sphagnosphere'. Plus particulièrement, les Sphaignes présentent un certain nombre d'attributs écologiques qui favorisent leur compétitivité par rapport aux autres espèces végétales, comme la faculté d'accroître l'acidité par échange cationique direct, la capacité de stocker de grandes quantités d'eau favorisant l'anaérobiose, l'adaptation à des milieux oligotrophes, la tolérance à la dessiccation (Clymo and Hayward 1982). Toutefois, il a été montré récemment que certaines de ces caractéristiques n'étaient pas propres aux Sphaignes et que, par exemple l'acidifiation par échanges cationiques était un mécanisme que l'on retrouvait chez d'autres Bryophytes (Soudilovskaia et al., 2010). Ce sont de plus des végétaux peu prédatés, seuls les Enchytréides *(Cognettia sphagnetorum)* et certaines larves d'Insectes (Tipules) les consomment (Smirnov, 1961; Rydin and Jeglum 2006).

1.2.2. La dynamique de la tourbière et la succession bas- / haut-marais (fen / bog)

En France, les tourbières à Sphaignes les plus anciennes ont débuté leur formation au début de l'Holocène, il y a environ 11 000 ans. Lors de la fonte des glaciers, les mouvements de moraines ont érodé le substratum géologique et permis le surcreusement de cuvettes glaciaires qui se sont colmatées au fur et à mesure de leur colonisation par les végétaux et l'accumulation de matière organique. Bien que ne correspondant pas à la seule modalité d'édification d'une tourbière à Sphaignes, ce processus d'atterrissement est bien représenté dans notre pays, illustré par l'existence de nombreux lacs-tourbières dans lesquels le rôle des Sphaignes s'accroit avec le comblement de la dépression lacustre originelle. Cette dynamique décrit une succession écologique, du stade pionnier de colonisation de la pièce d'eau par la végétation palustre au stade de haut-marais bombé. Cette succession s'illustre par la colonisation de communautés de Sphaignes spécifiques de différents stades dynamiques dont la trajectoire se caractérise notamment par de nombreux gradients hydro-écologiques (hydrique, trophique, d'acidification, de disponibilité en nutriments, etc.) en surface et subsurface ainsi qu'une accumulation de tourbe (Bridgham et al., 1996 ; Rydin and Jeglum 2006). Ce modèle dynamique de succession basé sur la végétation et les Sphaignes en particulier a été validé par des études sur les microorganismes (communautés méthanogènes : Merilä et al., 2006; communautés fongiques : Thormann et al., 2003). Toutefois, la dynamique et le développement de la tourbière ne sont pas uniformes au sein d'un même site et dépendent de facteurs physiques internes liés à la colonisation végétale, l'accroissement en épaisseur, et aux processus de transport et de subsidence (Figure 4). Il en résulte une organisation spatiale et verticale complexe, générée à la fois par des facteurs abiotiques et biotiques.

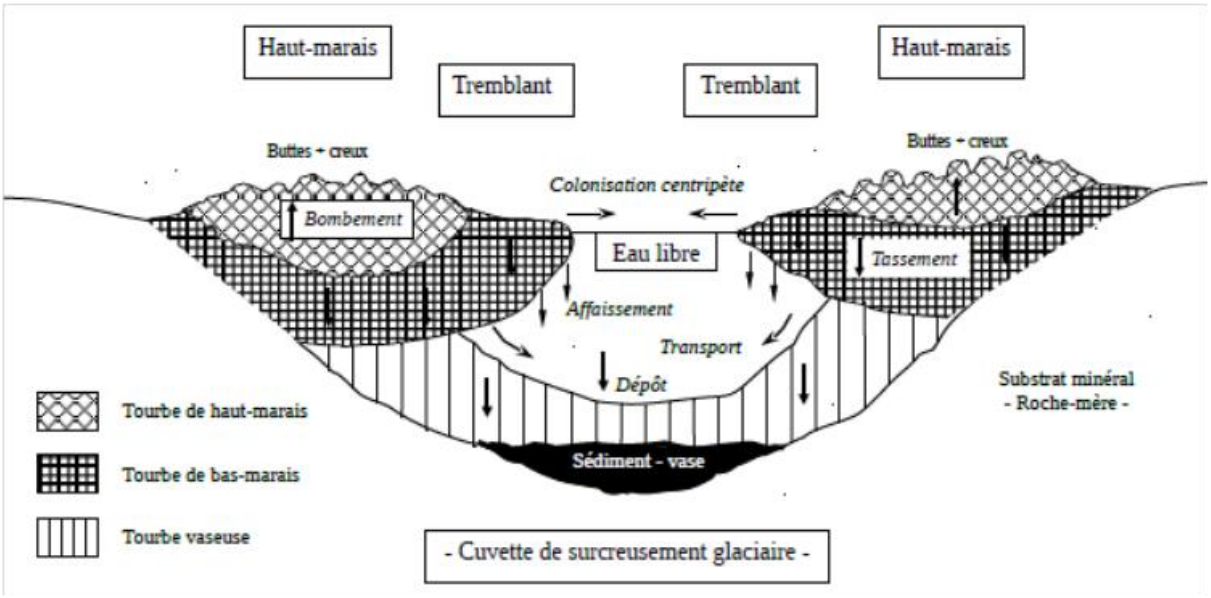


Figure 4 - Processus dynamiques mis en jeu au cours du processus d'atterrissement d'un lac-tourbière (D'après Francez 2000).

A côté du processus d'atterrissement permettant de passer d'une tourbière limnogène à une tourbière ombrogène, le haut-marais ombrogène peut se former par un processus de paludification, avec les précipitations comme seule origine de l'eau (Figure 5).

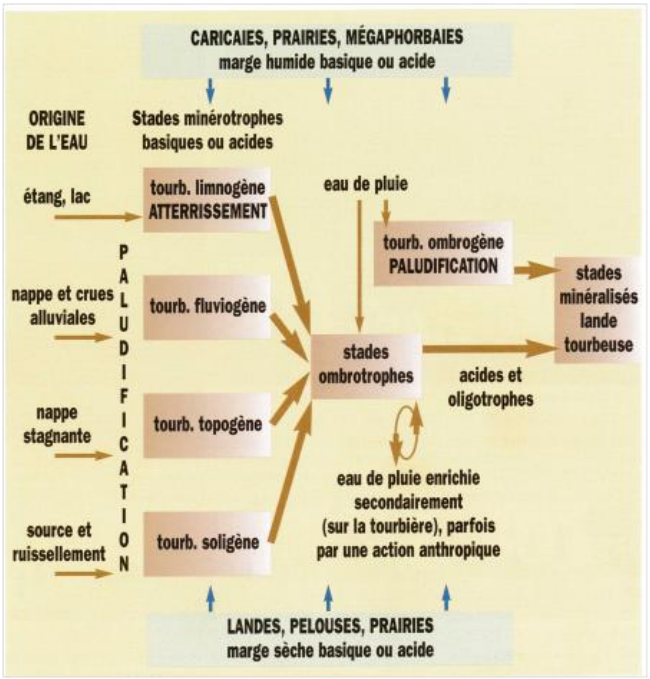


Figure 5 - Origine de l'eau, alimentation hydrique et structure formelle des tourbières (D'après Julve 1996 modifié *in* Manneville et al. 2006).

Enfin, le facteur anthropique peut contraindre la dynamique naturelle. Les tourbières à Sphaignes ont en effet été (et sont encore) utilisées pour l'agriculture ou l'extraction de tourbe, ce qui modifie la dynamique de l'écosystème. L'abandon de ces activités entraîne des dynamiques de

successions secondaires, les communautés microbiennes et leur diversité fonctionnelle évoluant de façon concomitante avec la végétation lors de la régénération du processus de tourbification, et le carbone labile contrôlant partiellement la succession microbienne (Artz et al., 2008).

1.3. L'accumulation de matière organique

1.3.1. Le concept de diplotelmie

Par le jeu de la croissance des Sphaignes, la tourbière voit sa topographie modifiée par la formation de buttes et de creux, l'ensemble vers le haut-marais bombé (bog) aux dépens du bas-marais (fen) (cf. Figure 4). L'accumulation de tourbe associée à l'évolution de l'écosystème tourbière et la saturation en eau structurent verticalement l'écosystème. Ainsi, on distingue généralement deux couches principales, (diplotelmie) aux propriétés hydro-écologiques distinctes : l'acrotelme en surface et le catotelme en profondeur (Ingram, 1978). Plus récemment, Clymo et Bryant (2008) ont proposé de considérer spécifiquement l'interface entre ces deux couches : le mésotelme. L'acrotelme, situé au dessus du niveau de fluctuation de la nappe d'eau, représente un milieu essentiellement aérobie. L'activité microbienne s'associe à une dégradation plus rapide de la matière organique, ainsi qu'à un développement plus dense du réseau racinaire des végétaux vasculaires colonisant les sphaignes. La catotelme correspond à la zone située sous le niveau de fluctuation de la nappe d'eau. La saturation en eau de cette couche génère des conditions anaérobies strictes où l'activité microbienne, plus limitée, est impliquée dans des processus de fermentation et de méthanogenèse. Le mésotelme, à l'interface, se caractérise par un fonctionnement mixte en fonction des fluctuations de la nappe d'eau, entrainant des variations de conditions d'aérobie/anaérobie et d'activité microbienne.

1.3.2. Le catotelme 'puits' de carbone

Les bilans de matière organique et d'éléments montrent que l'accumulation de tourbe s'effectue dans le catotelme (Clymo, 1984) et que les quantités de carbone stockés actuellement chaque année dans les tourbières non perturbées représentent approximativement 10% de la production primaire, soient environ 30 g C m^{-2} an^{-1} (Clymo, 1984 ; Gorham, 1991 ; Francez and Vasander, 1995 ; Roulet et al., 2007). Toutefois, de nombreuses études biostratigraphiques montrent que les taux de carbone accumulés ont largement varié au cours de l'Holocène. Ces variations résultent de changements environnementaux à différentes échelles, locales et régionales, comme par exemple des fluctuations de niveau d'eau, des périodes de réchauffement (Atlantique) ou de refroidissement (Subboréal). Il n'en reste pas moins que les tourbières ont régulé le climat de l'Holocène en fonctionnant comme des puits de carbone, la décomposition de la matière organique étant inférieure à la production

primaire. Les études actuelles soulignent la grande variabilité locale et régionale des échanges gazeux entre les tourbières et l'atmosphère (Strack, 2008).

1.4. Les services rendus par les tourbières

A côté de la fonction de puits de carbone et la régulation potentielle du climat au cours de l'Holocène, les caractéristiques des tourbières et de la tourbe en font des écosystèmes qui sont exploitées depuis l'Antiquité (Francez, 2000 ; Manneville et al., 2006). Compte-tenu de sa forte teneur en carbone (20 à 50 % de la matière sèche), la tourbe a longtemps servi de combustible au même titre que le bois ou le charbon. Les propriétés antiseptiques de la tourbe à Sphaignes ont contribué à son utilisation comme substrat en horticulture. Elles sont également drainées pour l'agroforesterie et l'agriculture. Elles sont de ce fait aujourd'hui considérées au même titre que les autres zones humides comme des écosystèmes rendant de nombreux services écosystémiques (Joosten and Clark, 2002 ; Mitsch and Gosselink, 2007) et ces services rendus à l'Homme sont maintenant largement pris en compte dans les opérations de restauration des tourbières (Kimmel and Mander, 2010). Toutefois, les catégories de services écosystémiques (MEA, 2005) sont inégalement traitées. Ce sont surtout les services de support (stockage d'éléments et biodiversité) et de régulation (climat, hydrologie) qui sont pris en compte dans les études de restauration (Kimmel and Mander, 2010). Dans cette optique, peu d'études se sont intéressées au compartiment microbien (stock et flux : Andersen et al., 2006 ; diversité fongique : Artz et al., 2007, Trinder et al., 2008).

1.5. Les tourbières et le changement global

Les tourbières subissent actuellement une combinaison de perturbations et de stress (Dise and Phoenix, 2011 ; Joosten et al., 2012). Les effets liés au changement global sont multiples. Les effets de la température sur la structure et le fonctionnement des communautés microbiennes font encore débat car ce facteur agit en synergie avec d'autres facteurs comme l'humidité ou encore la végétation (Figure 6). L'augmentation de température peut se traduire directement en favorisant l'augmentation de la fonction de respiration (auto- et hétérotrophe) de l'écosystème (Dorrepaal et al., 2009) et en augmentant potentiellement les pertes en carbone organique dissous (Fenner et al., 2007 ; Delarue et al., 2014). La température agit également indirectement, en interaction avec l'humidité, sur les émissions de gaz à effet de serre, en modifiant la végétation (sphaignes *vs* végétaux vasculaires, herbacés vs arbrisseaux) (Ward et al., 2013 ; Buttler et al., 2015), la structure des communautés microbiennes et les réseaux trophiques microbiens (Jassey, 2011 ; Gicquel, 2012 ; Kim et al., 2012). Ces changements climatiques affectent aussi les pergélisols des grandes tourbières à Sphaignes des hautes latitudes (Sibérie, Canada), entrainant la fonte de couches qui restaient

auparavant gelées, ce qui pourrait engendrer une plus forte dégradation de la tourbe, et une augmentation de l'activité méthanogène. Le méthane étant un gaz à effet de serre 20 fois plus puissant que le dioxyde de carbone, ces efflux supplémentaires pourraient amplifier l'effet de serre et le réchauffement. L'impact de l'augmentation de la température sur l'activité microbiologique pourrait donc faire basculer les tourbières d'un fonctionnement 'puits' à un fonctionnement 'source' de carbone, selon l'importance des mécanismes d'action-rétroaction environnement-végétation-microorganismes (Bardgett et al., 2008).

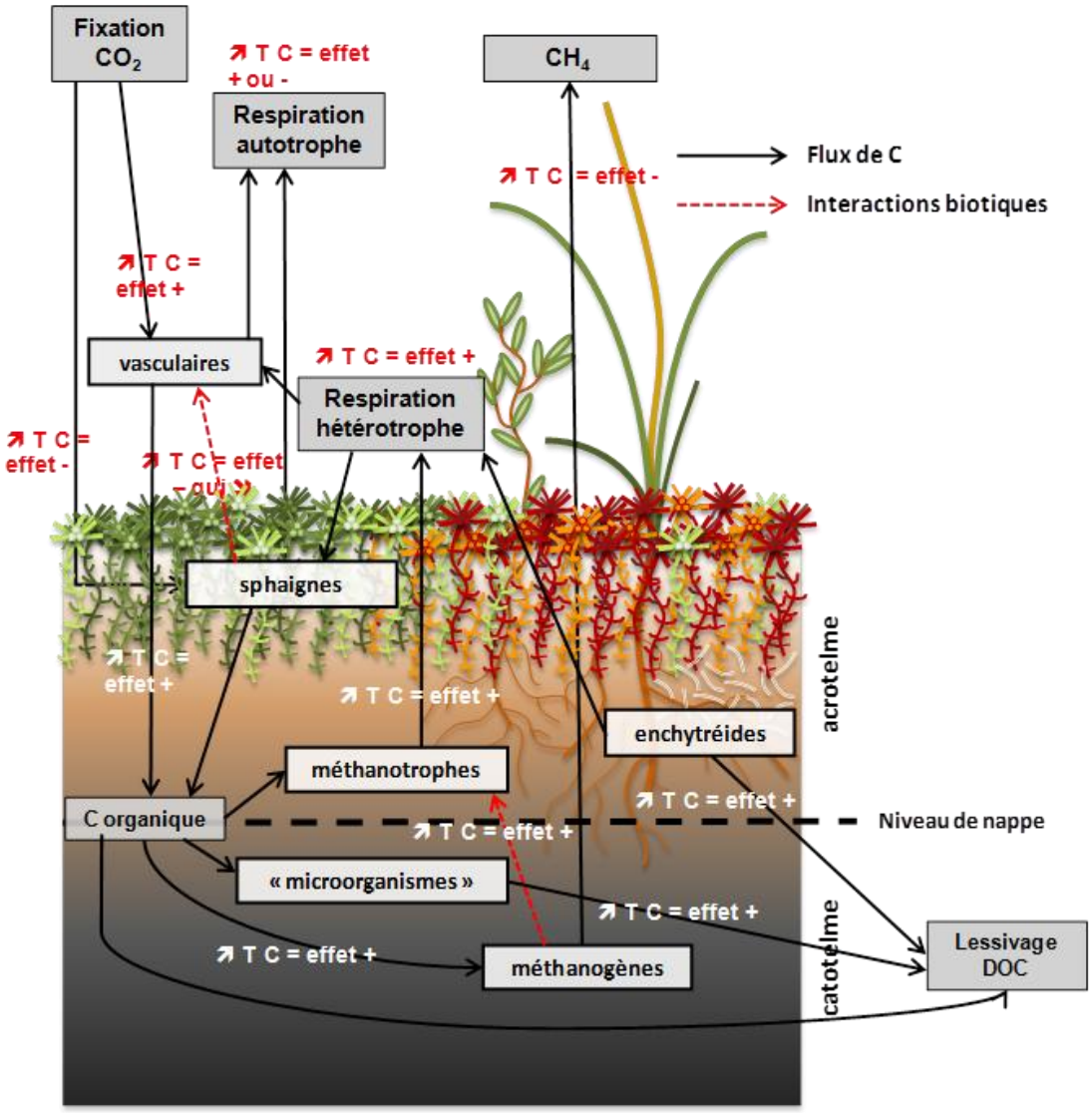


Figure 6 - Représentation schématique des conséquences d'un réchauffement climatique modéré sur les flux de carbone et les interactions biotiques dans les tourbières à Sphaignes (expérimentation Open Top Chamber programme ANR 'Peatwarm', d'après Gicquel 2012).

Il en est de même de l'augmentation des dépôts atmosphériques, essentiellement azotés, qui enrichissent les tourbières et entraînent des changements de composition et de structure des communautés végétales et microbiennes (Gilbert et al., 1998 ; Limpens et al. 2011). Les changements d'usage des terres (agriculture, foresterie, extraction de tourbe) affectent également les communautés microbiennes (Andersen et al., 2006 ; Basiliko et al., 2013).

L'exploitation des tourbières a un effet négatif sur l'activité microbienne. Croft (2001) a ainsi montré que, sur les surfaces décapées, la biomasse microbienne diminuait alors que l'extraction de tourbe engendrait dans le même temps une augmentation de l'ammonification. Ces surfaces fonctionnent comme des systèmes 'source' de carbone même si les émissions de CO_2 de la tourbe sans végétation restent inférieures à celles des tourbières non perturbées (Waddington et al., 2001). Les couches profondes de la tourbe, habituellement anaérobies, deviennent aérobies en raison du drainage, ce qui stimule l'activité microbienne (Glatzel et al., 2003).

La régénération de la tourbière est généralement évaluée par le retour des végétations typiques de l'écosystème, mais n'assure pas forcément le ré-établissement de la fonction de 'puits' de carbone (Andersen et al., 2013). Le fonctionnement du cycle du carbone et de la diversité métabolique bactérienne n'est donc généralement restaurée qu'en surface, là où la production d'un nouvel acrotelme est assuré par la ré-introduction et la croissance des tapis de Sphaignes (Andersen et al., 2006).

2. Diversité et fonctions du compartiment microbien

2.1. Un peu d'histoire

Les premières études bactério-écologiques et les questions relatives à la décomposition/accumulation de la tourbe remontent au début du XXème siècle avec les travaux pionniers de Waksman et ses collaborateurs, dont le but était de comprendre le rôle des microorganismes dans les processus de décomposition de la tourbe (cf. par exemple Waksman and Stevens, 1929). D'autres ont suivi, toujours basées sur des techniques de mise en cultures sur milieux solides et/ou liquides afin d'approcher l'abondance et l'activité des microorganismes impliquées dans les grands cycles biogéochimiques, en estimant par exemple les activités cellulolytiques, amylolytiques (cycle du C), les activités ammonifiante et nitrifiante (cycle de l'N), ou encore de dénombrer les bactéries cultivables en aérobie et anaérobie (cf. par exemple de Barjac, 1955). C'est avec l'émergence du concept de biomasse microbienne et les recherches du Programme Biologique International que les études microbiologiques se sont peu à peu affranchies de l'approche par groupes physiologiques. Différents auteurs ont cherché à estimer la biomasse bactérienne par des

méthodes de comptage direct au microscope et de conversion à partir de mesures de biovolumes (Given and Dickinson, 1975 ; Collins et al., 1978 ; Clarhom and Rosswall, 1980 ; Gilbert et al., 1998). Parallèlement aux premières estimations de la biomasse microbienne par fumigation-extraction (Williams and Sparling, 1984 ; Sparling et al., 1990) ou par dosage des adénosines (Maire, 1983 ; Francez, 1991), l'activité globale des microorganismes a été quantifiée par respirométrie en mesurant la production de CO_2 *in situ* au moyen de chambres statiques ou en conditions contrôlées de laboratoire (Yavitt et al., 1987 ; Stewart and Wheatley, 1990 ; Bridgham et al., 1992). Plus récemment la structure des communautés microbiennes (archées, bactéries, champignons, protozoaires) de tourbières a été caractérisée en analysant les profils de phospholipides, couplés à des mesures d'activités enzymatiques (Andersen et al., 2010 ; Bragazza et al., 2015 ; Delarue et al., 2015) ou à des mesures d'activité spécifiques comme la nitrification ou la dénitrification via la mesure de flux de gaz (Björk et al., 2010). Enfin, au cours des deux dernières décennies, grâce aux apports de l'écologie moléculaire, les études sur les communautés microbiennes de tourbières ont fait apparaitre une très grande diversité taxonomique et fonctionnelle, que ce soit chez les bactéries (Dedysh et al., 2006 ; Morales et al., 2006 ; Pankratov et al., 2011 ; Kip et al., 2012), les champignons (Artz et al., 2008 ; Peltoniemi et al., 2009), les archées (Basilliko et al., 2003 ; Sizova et al., 2003 ; Kotsyurbenko et al., 2004 ; Putkinen et al., 2009). Ces travaux ont été complétés plus récemment par les approches de métagénomique et métatranscriptomique (Tveit et al., 2013 ; Bragina et al., 2014 ; Lin et al. 2014 ; Tveit et al., 2014).

2.2. Les communautés microbiennes et les processus de décomposition

Par rapport à l'étude de la végétation des tourbières, la compréhension du fonctionnement du compartiment microbien s'est faite plus tardivement, et par étapes, conjointement aux développements des technologies adéquates permettant d'appréhender la diversité spécifique et fonctionnelle de ces organismes "invisibles à l'œil nu". En raison des différences de végétation, de dégradation de la tourbe et de physico-chimie entre les stades dynamiques (fen-bog) et la profondeur, beaucoup de travaux se sont focalisés sur la comparaison des structures de communautés microbiennes, associées à l'une ou l'autre dimension spatiale ou à la combinaison de ces deux gradients (cf. par exemple Myers et al., 2012 ; Jassey et al., 2012).

Dans les écosystèmes terrestres les apports de litières (aérienne et souterraine) constituent la principale ressource de matière et d'énergie des communautés microbiennes entretenant de nombreuses interactions complexes (Hättenschwiler et al., 2005) et la diversité microbienne aurait un impact positif sur l'efficacité de recyclage des nutriments (Loreau, 2001). Dans la tourbière, la décomposition de la matière organique et son recyclage contrôlent la disponibilité en nutriments et la production de l'écosystème et de ce fait sa capacité à fonctionner comme un puits ou une source.

L'essentiel des processus de biodégradation s'effectue dans l'acrotelme et le mésotelme (Clymo, 1984 ; Artz, 2009). Les fluctuations de la nappe d'eau conditionnent l'alternance de conditions aérobies et anaérobies et donc l'importance respective des différentes voies de décomposition (Figure 7). Dans le catotelme, saturé en eau de façon permanente, les processus de fermentation et de méthanogenèse dominent (Figure 7).

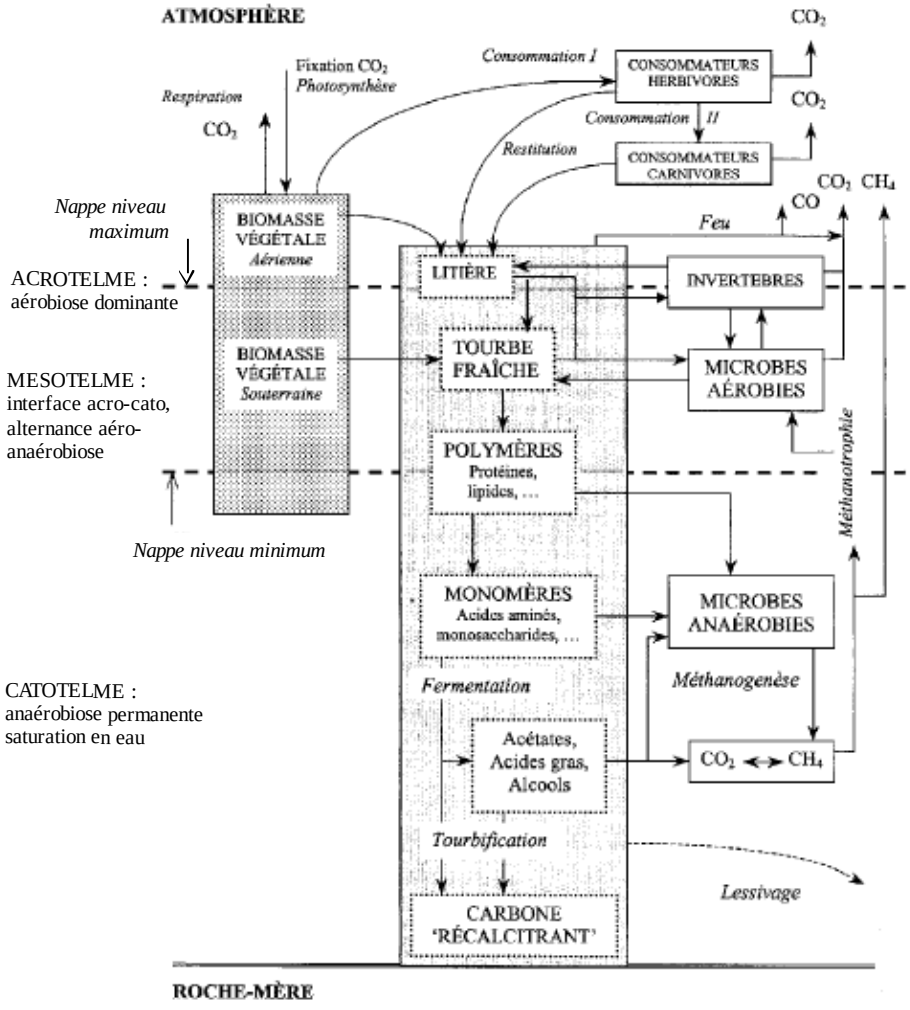


Figure 7 - Modèle descriptif du cycle du carbone illustrant les principaux processus aérobies et anaérobies impliqués dans l'activité microbienne et la décomposition de la tourbe (Francez 2000 modifié).

Les processus de décomposition dans les tourbières ont été modélisés (Frolking et al., 2001 ; Bauer, 2004) et le modèle exponentiel général de décomposition (Wieder and Lang, 1982) est applicable en tourbières au moins pour les 2 premières années de décomposition (Francez, 1995 ; Thormann and Bayley, 1997 ; Moore et al., 2005), les taux de décomposition variant avec la nature du végétal, l'âge de la tourbe et son degré d'humification dont la méthode de mesure est ancienne (indice von Post, 1924). Ainsi, le taux de décomposition dans l'acrotelme est plus élevé dans le bas-

marais que dans le haut-marais (0.4 et 0,25 an^{-1} respectivement, *cf.* synthèse in Francez 2000). Les litières de Sphaignes contiennent des composés phénoliques dont l'acide *p*-hydroxy-β[carboxyméthyl]-cinnamique ou *"Sphagnum* acid" (van der Heijden et al., 1997) qui, lié aux biopolymères des parois, confèrerait aux tissus de ces Bryophytes, des propriétés de récalcitrance à la biodégradation, proches de celles de la lignine (van Breemen, 1995). La décomposition des litières de plantes de tourbières varie également sous le contrôle de facteurs locaux et régionaux, en particulier la T°C. Ainsi il a été montré, que la température (climat local) avait une influence positive sur le coefficient k de décomposition des plantes vasculaires de tourbières (Thormann, 2004 ; Breeuwer et al., 2008).

2.3. La complexité des réseaux trophiques et la boucle microbienne

Le flux d'accumulation entrant dans le catotelme représente environ 10 % du flux de production primaire. Ainsi, 90 % de la biomasse produite annuellement sont biodégradés et minéralisés, cette transformation de la matière organique se réalisant au sein de réseaux trophiques complexes et diversifiés, des bactéries aux protozoaires prédateurs (Figure 8), incluant de nombreuses relations de types syntrophiques et 'above-belowground'. Par exemple, les bactéries méthanotrophes *(Methylocystis)* endophytes des sphaignes, oxydent le CH_4 en CO_2, celui-ci étant rapidement photosynthétisé par les Sphaignes (Raghoebarsing et al., 2005), le processus étant contrôlé par le facteur hydrique (Larmola et al., 2010). Les bactéries méthanotrophes de la sphagnosphère contribuent également au stockage d'azote, via la fixation de N_2 atmosphérique, les Cyanobactéries ayant cependant un avantage compétitif dans les tourbières les plus oligotrophes (Vile et al., 2014).

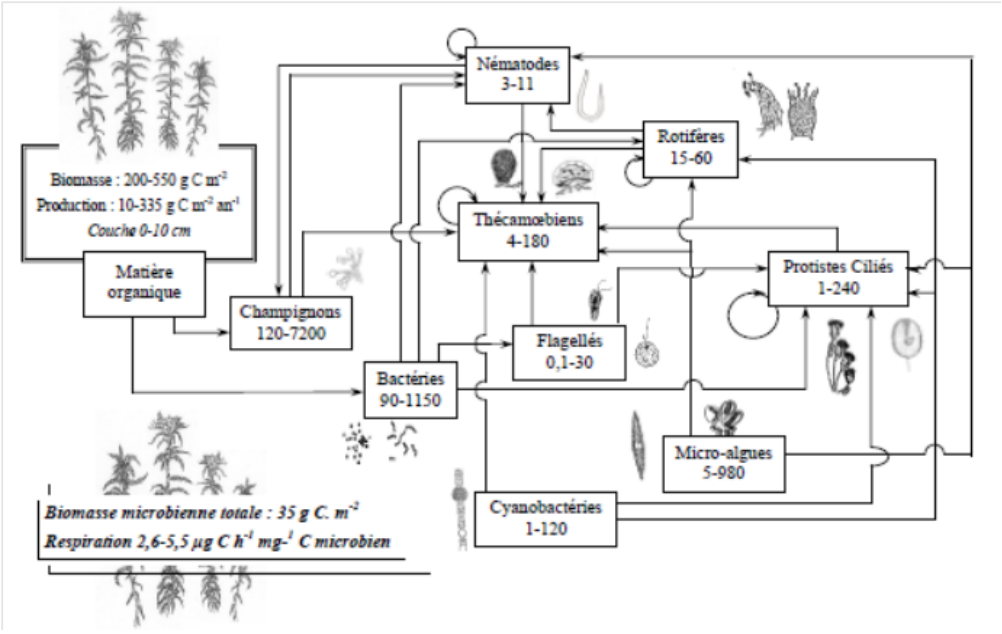


Figure 8 - Réseaux trophiques microbiens. Les flèches indiquent les relations prédateur-proie et les estimations de biomasse (synthèse divers auteurs) sont précisées pour chacun des compartiments (D'après Francez 2000).

Dans le fonctionnement des réseaux trophiques microbiens, les Protozoaires prédateurs (Ciliés, Amibes, Flagellés) assurent un lien entre les décomposeurs bactériens et les niveaux trophiques supérieurs (Nématodes, Rotifères, larves d'Insectes). Ils stimulent également dans les écosystèmes terrestres la croissance des plantes, bien que les mécanismes sous-jacents ne soient pas encore complètement élucidés (Ekelund et al., 2009) et stimulent l'activité des microorganismes impliqués, par exemple, dans la minéralisation de l'azote (Griffiths, 1989). Le concept de boucle microbienne ('microbial loop') a été proposé par Azam et al., (1983) en milieu marin. Le carbone organique libéré par le phytoplancton sous forme dissoute, est assimilé par les bactéries et réintroduit dans la chaine des brouteurs par la prédation des Flagellés. Le rôle des Protozoaires prédateurs dans la régulation des populations de bactéries édaphiques et le recyclage du carbone et de l'azote dans les sols est connu depuis longtemps (Coûteaux et al., 1988). L'équivalent terrestre du concept de boucle microbienne correspond au concept de 'fast flora' (Coleman et al., 1976, *in* Clarholm, 1994). Le carbone libéré par les racines est incorporé par les bactéries qui dans le même temps interceptent les sels nutritifs transportés vers les racines par diffusion et qui, de ce fait, deviennent inaccessibles aux plantes. Le broutage des bactéries par les Protozoaires permet la reminéralisation d'une grande partie de ces nutriments pour les producteurs primaires. Les Protozoaires, essentiellement des amibes, agissent ainsi comme des régénérateurs de nutriments (Clarholm, 1994). Plus récemment, il a été montré que des molécules excrétées par certains Protozoaires telles que le nitrate pouvaient servir de signal à la croissance des racines (Bonkowski, 2003). Dans les tourbières, Gilbert et al. (1998), observant l'abondance de nombreuses espèces de Rhizopodes Testacés, ont proposé d'étendre le concept de boucle microbienne au fonctionnement des réseaux trophiques microbiens de la sphagnosphère, ces prédateurs jouant un rôle clé dans les transferts de carbone et d'azote, ingérant selon les espèces, des bactéries et des mycéliums de champignons ou encore des Ciliés et des microalgues (Jassey et al., 2012).

3. Les nouveaux champs de recherche en écologie microbienne ouverts par les techniques de séquençage

Malgré l'attention portée au compartiment microbien des tourbières, Gilbert et Mitchell (2006), dressaient le bilan suivant : bien qu'avérée, la diversité microbienne des tourbières n'avait reçu que peu d'attention, et n'était toujours approchée que de manière dichotomique : soit par l'approche diversité soit par l'approche fonctionnelle. Cependant, depuis une vingtaine d'années, de nouvelles technologies permettent désormais d'ouvrir 'la boîte noire' et d'appréhender à la fois la diversité taxonomique et fonctionnelle des communautés procaryotes par des approches métagénomiques et permettent de s'affranchir des contraintes et limites liées aux méthodes de culture. Ces approches

dites « culture-indépendentes» ou métagénomiques sont basées sur l'extraction du pool d'ADN environnemental d'un échantillon.

3.1 Sélectionner, amplifier et séquencer un gène d'intérêt

Dans la plupart des cas, pour identifier la diversité des organismes présents, on cherche dans ce pool d'ADN à sélectionner et à amplifier par PCR (polymerase chain reaction) un gène présent chez tous les organismes cellulaires : le gène de l'ARN ribosomal (Woese, 1987). Cette première étape nécessite de pouvoir développer des amorces spécifiques du gène que l'on veut amplifier. Pour qualifier la diversité, les fragments d'ADN amplifiés sont visualisés sous forme de bandes sur un gel d'agarose après migration par électrophorèse. Le front de migration d'une bande dépend de la longueur et de la composition de chaque fragments amplifié, qu'ils soient juste dénaturés (DGGE) (Artz et al., 2007 ; Høj et al., 2005 ; Høj et al., 2007) ou digérés par des enzymes de restrictions (T-RFLP) (Morales et al., 2006 ; Pankratov et al., 2011 ; Preston et al., 2012). La comparaison des communautés est donc basée sur les profils de migration. L'inconvénient de cette méthode est qu'elle n'est pas quantitative, et qu'il reste difficile d'identifier formellement les organismes dont sont issus les gènes amplifiés. Pour avoir une identification de l'individu, il faut séquencer la séquence de gène d'ARNr ou du gène d'intérêt amplifié, et le comparer aux bases de données (Dedysh et al., 1998 ; Dedysh et al., 2006 ; Opelt et al., 2007 ; Artz et al., 2007 ; Pankratov et al., 2011). Cette méthode présente des limites, puisque pour amplifier la séquence, il faut des amorces spécifiques du gène que l'on veut sélectionner, donc une connaissance de la cible. Or, le gène d'ARNr diffère entre le règne eucaryote et procaryote (taille des sous unités plus élevée dans le règne eucaryote). Il faut donc amplifier différemment ce gène pour avoir un aperçu total de la diversité microbienne d'un échantillon (Preston et al., 2012).

3.2 Identifier les séquences issues d'organismes non-cultivés : exemple avec les séquences de gènes d'ARNr

Les banques de données d'ARNr les plus utilisées sont SILVA (www.arb-silva.de) ou RDP (Ribosomal Database Project). Elles contiennent les séquences d'ARNr des organismes cultivés et surtout non-cultivés. En fonction des variations présentes sur les séquences du gène d'ARN ribosomique 16S des échantillons par rapport aux séquences des banques des données, il est possible de déterminer une affiliation phylogénétique de l'organisme (exemples d'arbres phylogénétiques dans Dedysh et al., 1998 ; Pankratov et al., 2011 ; Serkebaeva et al., 2013).

3.3 Diversité taxonomique et fonctionelle

Ces différentes techniques basées sur les étapes de sélection, d'amplification, et de séquençage ont été utilisées pour analyser les diversités archéenne, bactérienne et/ou fongique de tourbières naturelles, exploitées, abandonnées, ou restaurées, par DGGE ou T-RLFP. Elles ont souvent permis d'explorer un ou deux groupes de microorganismes comme les archées (Høj et al ., 2005 ; Putkinen et al., 2009), les bactéries (Dedysh et al., 2006 ; Morales et al. 2006 ; Pankratov et al., 2011) ou les champignons (Artz et al., 2007 ; Peltoniemi et al., 2009).

Les études fonctionnelles sur les communautés microbiennes des tourbières ont beaucoup privilégié les communautés de microorganismes impliqués dans le cycle du méthane, bactéries méthanotrophes et archées méthanogènes, ou les gènes impliqués dans la méthanogénèse ou la méthanotrophie (par exemple le gène d'oxydation du méthane via la methyl monooxygénase codée par des gènes de type *mmo*-) (Dedysh et al., 1998 ; Basiliko et al. 2003, Sizova et al. 2003, Kotsyurbenko et al., 2004, Juottonen et al., 2005 ; Raeghoebarsing et al. 2005, Kip et al. 2012 ; Bragina et al., 2013). Il a été ainsi montré que le fen caractérisé par de fortes émissions de méthane comparé au bog, présente une plus grande diversité du gène Methyl-coenzyme reductase gene (*mcrA*) impliqué dans la méthanogénèse (Juottonen et al., 2005). D'autres études ont pour objectif d'analyser l'effet de perturbations et l'impact de la restauration sur la diversité des communautés de microorganismes et leur résilience (Preston et al., 2012 ; Basiliko et al. 2013). Ces études souvent menées sur le gène d'ARNr 16S ne prennent donc pas en compte la diversité des protistes qui contrôlent et régulent les communautés bactériennes. De plus, en raison des biais d'amplification, liés aux amorces utilisées, cette méthode est assez peu quantitative. Elle reste cependant un bon moyen de montrer la grande diversité des communautés de procaryotes des écosystèmes.

3.4 Séquencer sans amplifier

Une autre approche, plus complexe et plus coûteuse, qualifiée de « shotgun métagénomique » ou « séquençage aléatoire » consiste à fragmenter le pool d'ADN environnemental total et à le séquencer afin d'obtenir une image globale et *a priori* quantitative puisqu'elle est exempte des inconvénients liés à la sélection d'un marqueur de gène et des biais d'amplification (Tveit et al., 2013 ; Lin et al., 2014 ; Bragina et al., 2014 ; Hultman et al., 2015). Tous les organismes (diversité) et toutes les fonctions représentés dans l'échantillon peuvent à priori être identifiés. Cette technologie reste tout de même souvent tributaire de la capacité à assigner les séquences métagénomiques obtenues, les assignations taxonomiques (qui) et fonctionnelles (quoi) des séquences dépendant de leur ressemblance avec des séquences annotées dans les bases de données. La métagénomique permet aussi de relever des nouvelles séquences sans aucune

référence dans les banques de données (Lopez-Garcia and Moreira, 2008 ; Singh et al., 2009 ; Hultman et al., 2015). Cette technologie a permis de démontrer que les communautés de procaryotes de la sphagnosphère possédaient une diversité fonctionnelle qui les différenciait des microbiomes des autres végétaux (Bragina et al., 2014).

Enfin, les approches métagénomiques basées sur le séquençage de l'ADN détectent la présence d'un organisme ou d'un gène mais n'indiquent pas si ces gènes sont exprimés et en conséquence si les microorganismes sont actifs, c'est-à-dire si les gènes sont transcrits en ARN puis traduits en protéines. L'étude « omic » de l'ARN est plus délicate car cet acide nucléique est beaucoup plus labile que l'ADN, et est très sensible à de nombreuses enzymes, comme les RNAses. Cependant, si le challenge technique de l'extraction d'ARN est surmonté, les mêmes technologies (sélection/amplification ou séquençage aléatoire) peuvent s'appliquer au pool d'ARN. Un métranscriptome représente alors une image de la diversité taxonomique et fonctionnelle d'une communauté microbienne active présente dans un échantillon de tourbe (Tveit et al., 2013 ; Tveit et al., 2014 ; Hultman et al., 2015 ; Tveit et al., 2015).

4. L'émergence des concepts de l'écologie virale et de la viriosphère

4.1. La découverte des virus

Comme pour les bactéries, les premiers verrous liés à l'identification des virus nécessitaient le développement de technologies avancées afin de « voir » et de « caractériser » ces entités. Les virus ont été découverts par Dimitri Ivanovski à la fin du XIX° alors qu'il cherchait à identifier un parasite incultivable, destructeur des plants de tabac et qui résistait aux filtrations successives avec des mailles en porcelaine de plus en plus fines (Bos, 1999). Les premiers travaux sur la nature de ces agents infectieux ont rapidement permis de déterminer qu'ils pouvaient infecter toutes sortes de cellules (bactériennes, animales, végétales), et ont pu conclure approximativement à la description des virus comme des particules de petites tailles, dépourvues d'activité métabolique et composés d'une capside protéique et d'un acide nucléique (historique dans l'introduction du travail de thèse de Simon Roux, 2013).

4.1.1. Classification des virus

C'est sur la base de la structure de la capside, de la longueur et du type de génome et du domaine des hôtes que l'on classe toujours les virus (King et al., 2011). La classification de Baltimore sépare les différents groupes de virus en fonction de la nature du génome et de la façon dont il est exprimé et répliqué. On distingue ainsi les virus à ADN (double ou simple brin), les virus à ARN (double brin, et simple brin, positif ou négatif), et les « retro-transcribing virus » à ARN (comme celui

du VIH). La structure 3D de la capside a plus récemment été proposée comme un meilleur moyen de classifier les virus, car sa structure détermine la quantité d'acides nucléiques et donc la taille du génome viral. Sa forme reflète aussi les hôtes potentiels du virion (d'Abrescia et al., 2012 ; Koonin and Dolja, 2014). Cette classification des virus est en perpétuelle évolution avec les progrès de la virologie (King et al., 2011 ; d'Abrescia et al., 2012 ; Koonin and Dolja, 2014).

4.1.2. Les principales stratégies d'infection virales

La réplication d'un virus se fait au dépend de la machinerie cellulaire de l'hôte avec laquelle le virus interagit selon différentes stratégies, liées en partie au métabolisme de ce dernier (Figure 10). Les deux stratégies principalement étudiées en écologie virale sont la stratégie lytique et la stratégie lysogénique (Weinbauer and Suttle, 1996 ; Fuhrman, 1999 ; Paul, 2008 ; Payet and Suttle, 2008 ; Maurice et al., 2013 ; Payet and Suttle 2013, Palesse et al., 2014). Dans les deux cas, le virus se fixe sur la membrane de l'hôte, et injecte son génome dans la cellule.

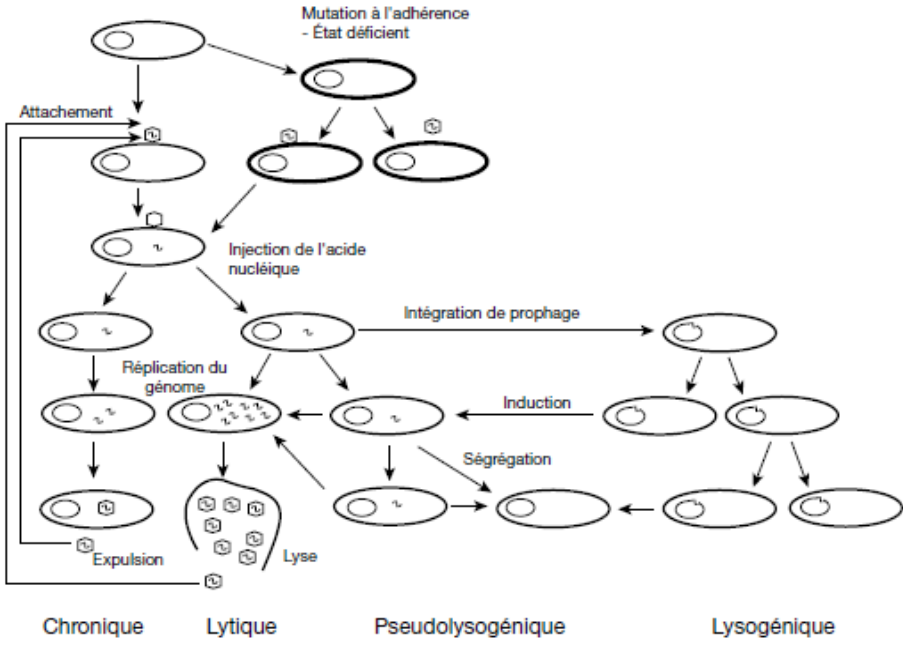


Figure 10 - Représentation schématique des différents cycles de vie des virus (D'après Weinbauer 2004 in Berdjeb et Jacquet 2009a).

Dans le cas de l'infection lytique, la machinerie de la cellule hôte permet la réplication du génome viral et la production des protéines de capsides. La cellule hôte se remplit de virions qui vont mener à la lyse de la cellule, à la libération des virus capables d'infecter une nouvelle cellule et à la mise à disposition des composés lysés dans le milieu (Weinbauer and Peduzzi, 1994 ; Parada et al., 2006 ; Weinbauer et al., 2011 ; Ankrah et al., 2014). Dans le cas de la stratégie lysogénique ou « tempérée », le génome du phage est intégré à celui de l'hôte. Le génome viral est alors appelé

« prophage » et la bactérie est qualifiée de « lysogène ». Le prophage est ensuite maintenu dans l'hôte et répliqué, dupliqué et transmis avec et par la machinerie de l'hôte. Ce type de cycle se perpétue jusqu'à ce qu'un élément induise la sortie de lyse, c'est-à-dire l'extraction du prophage, la réplication du génome viral, la production de virions et la lyse de la cellule hôte. Il existe par ailleurs d'autres stratégies d'infection virales, comme le « budding » chez les archées (production continuelle de virions sans lyse de la cellule hôte), mais celles-ci sont moins documentées (Furhman 1999 ; Quemin and Quax, 2015 ; Weinbauer, 2004 ; Sime-Ngando, 2014).

4.2. Les apports de la microscopie électronique sur la morphologie des virus

Pour mieux comprendre le fonctionnement des interactions entre les virus et leur hôte, il était nécessaire de développer des techniques permettant de « cultiver » les virus en présence de leur hôte, puisqu'ils en sont dépendants. L'utilisation de la microscopie électronique a permis de confirmer la présence des virus dans tous les écosystèmes. (Figure 9). Certaines formes de virus sont caractéristiques d'un type d'hôte. Les phages, par exemple, n'infectent que des procaryotes et sont caractérisés par une tête, une queue et un pied. La forme et la taille du pied sont utilisées pour décrire différents types de phages et comparer la diversité morphologique virale entre les écosystèmes (Weinbauer and Peduzzi, 1994 ; Ashelford et al., 2003 ; Zhong et al., 2014).

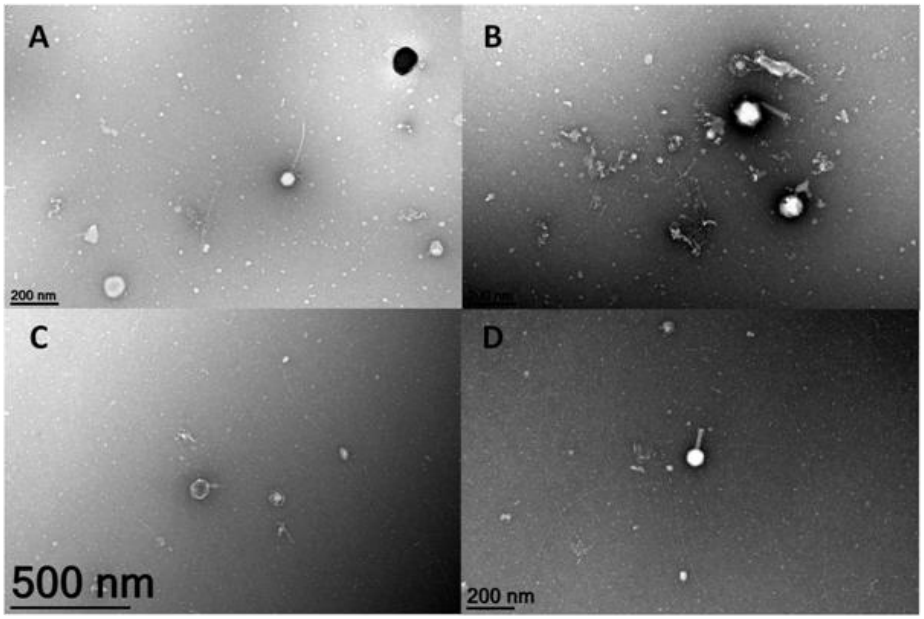


Figure 9 – Images MET (Transmission Electronic Microscopy) de phages observés dans le tapis de *Sphagnum fallax* fen en août 2013, (photographies Agnès Burel, Plateforme Microscopie UR1). A : *Siphovirus*, B et D: *Myovirus*, C: *Podovirus*.

Cependant, du fait de la fragilité des « pieds » des virus, qui sont souvent perdus lors des étapes de filtrations, et du fait que ces pieds ne soient pas toujours visibles en microscopie électronique, cette technologie n'est pas considérée comme fiable pour une représentation exhaustive de la diversité virale. Elle reste cependant une première étape importante pour confirmer

et qualifier la présence de virus dans un écosystème dont la diversité virale n'a jamais été approchée (Lopez-Bueno et al., 2009 ; Sawnson et al., 2009 ; Borrel et al., 2012).

4.3. Les apports de la microscopie à épifluorescence et de la cytométrie

Si la microscopie électronique a permis de découvrir de nouvelles morphologies virales, son utilisation pour la quantification est limitée (Sime-Ngando, 2014). La découverte et la mise à disposition de molécules fluorescentes, telles que le SYBR-green, s'intercalant dans les acides nucléiques (ADN et ARN) ont permis de différencier et de compter plus facilement les virus et les procaryotes dans des échantillons environnementaux (Noble and Fuhrman, 1998 ; Marie et al., 1996 ; Brussaard and Bratbak, 2000 ; Duhamel and Jacquet, 2006 ; Patel et al., 2007). La différenciation par fluorescence se fait sur la taille des organismes : les virus ayant généralement un plus petit génome que les bactéries, celui-ci fluoresce moins que celui des procaryotes (Marie et al., 1999). Une fois « marqués » par la molécule fluorescente, les virus et les procaryotes sont plus facilement comptés, que ce soit en microscopie à épifluorescence ou en cytométrie en flux (Figure 11), ces techniques ont multiplié les capacités de comptage dans de nombreux échantillons et permis la comparaison des abondances virale et bactérienne dans différents écosystèmes (Brussaard, 2004 ; Williamson et al., 2005 ; Duhamel and Jacquet, 2006). Cependant, des doutes sont parfois soulevés sur le fait que les petites particules comptées en cytométrie ne soient pas toutes des virus (Forterre et al., 2013).

Les années 1980 marquent le début des premières estimations quantitatives de virus dans les écosystèmes aquatiques, marins et d'eau douce (Bratbak et al., 1994). Ces études ont révélé que les virus étaient actuellement les entités biologiques les plus abondantes sur Terre (Fuhrman 1999 ; Suttle 2005 ; Weinbauer 2004 ; Rohwer and Thurber 2009 ; Sime-Ngando, 2014) et ont souligné la grande proportion de virus par rapport aux bactéries : de l'ordre de dix particules virales pour une seule hôte (Fuhrman, 1999 ; Suttle 2005 ; Rohwer and Thurber, 2009). Ce rapport est abrégé « VPR » ou « VBR » (« virus to prokaryotes ratio » ou « virus to bacteria ratio »). En raison du lien entre production virale et activité bactérienne (Middelboe, 2000), les VPR très élevés en eaux douce et marine ont amené l'idée d'une forte pression virale sur la communauté bactérienne des écosystèmes aquatiques (Rohwer and Thurber, 2009).

4.4. La « burst size » et le lien activité de l'hôte/production virale

Il a également été découvert que la production virale (le nombre de virions produits par une cellule hôte en fonction du temps) dépendait de l'activité de la communauté microbienne (Middelboe, 2000 ; Middelboe et al ., 2002 ; Parada et al., 2006 ; Payet and Suttle 2008). La quantité de virus produits par une cellule engendrant sa lyse est qualifiée de « burst size » (= « taille de l'explosion »). Elle est encore souvent estimée en microscopie électronique par quantification du nombre de

virions/virus présents autour des cellules visiblement infectés (Weinbauer and Peduzzi 1994 ; Weinbauer, 2003 ; Parada et al., 2006 ; Paterson and Laybourn-Parry, 2012). L'étude de la « burst-size » est souvent complétée par des manipulations visant à diminuer le nombre de particules virales dans l'échantillon (par dilution des virus et concentration des bactéries) afin d'estimer la productivité virale (nombre de virus produits par unité d'hôte et de temps)(Wilhelm, 2002 ; Paterson and Laybourn-Parry, 2012).

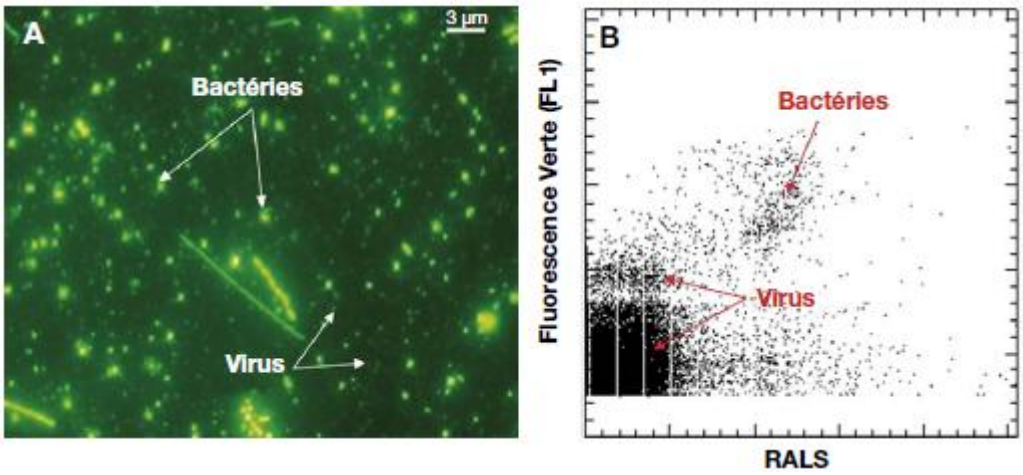


Figure 11 – Visualisation typique de bactéries et de virus. A: par microscopie à épifluorescence (D'après Chen *in* http://www.virusecology.org/MOVE/Method%206.html). B: par cytométrie en flux après marquage au SYBR green (D'après Berdjeb et Jacquet 2009a).

4.6. Les technologies de séquençage et la connaissance de la diversité virale des écosystèmes

Par rapport à la diversité procaryote, l'étude de la diversité virale a longtemps été retardée par le fait que, contrairement aux microorganismes, il n'existe pas de marqueur moléculaire très conservé chez les virus (Edwards and Rohwer, 2005 ; Sakowski et al., 2014). Les premiers travaux de diversité se sont basés sur les acides nucléiques des virus et ont donc commencé par des comparaisons de tailles et de compositions des génomes viraux dans les échantillons environnementaux, et sont toujours utilisés (Zhong et al., 2014). C'est avec l'émergence des technologies de séquençage que les premiers « métaviromes » (pools d'ADN viral issus d'échantillons environnementaux) ont été produits. Les premiers génomes de virus marins ont été séquencés par méthode Sanger en 2002 (Breitbart et al., 2002), puis, avec le pyroséquençage (Edwards and Rohwer , 2005 ; Thurber et al., 2009), la diversité virale a pu être analysée dans des écosystèmes aussi variés que les stromatolithes (Desnues et al., 2008), les étangs du Sahara (Fancello et al., 2012), les lac Antarctiques (Lopez-Bueno et al., 2009), les lacs d'eau douce (Roux et al., 2012)... Ces métaviromes ont révélé que les virus représentaient le plus grand réservoir de diversité génétique actuellement connu sur Terre (Breitbart and Rohwer, 2005 ; Edwards and Rohwer, 2005 ; Rohwer and Thurber 2009). La comparaison de la diversité et de

l'abondance des différents types viraux indiquent que les communautés virales varient fortement entre les écosystèmes, en lien avec les communautés d'hôtes sous-jacentes.

L'importante production de séquences virales issues de différents écosystèmes a permis de décrire et d'identifier la grande diversité des génomes viraux. Si les génomes de certains virus comme les Circoviridae (virus à ADN simple brin infectant les eucaryotes) sont essentiellement constitués d'un gène codant pour la capside et un autre pour la réplication du génome (d'Abrescia et al., 2012), certains virus sont de plus grandes tailles et possèdent des génomes viraux plus longs que les bactéries (Philippe et al., 2013). C'est notamment le cas des virus d'amibes à thèques, comme les Mimivirus, les Pandoravirus et plus récemment le *Pithovirus sibericum* et le *Mollivirus sibericum*, tous deux découverts dans des sols tourbeux (Philippe et al., 2013, Piacente et al., 2012 ; Abergel and Claverie, 2014; Legendre et al., 2015).

Les génomes de ces virus, et ceux d'autres virus comme les cyanophages, peuvent porter des gènes potentiellement impliqués dans les voies métaboliques de l'hôte, comme dans la synthèse de sucre (Piacente et al., 2012) ou dans la photosynthèse (Lindell et al., 2004). Les liens entre les génomes de l'hôte et les génomes viraux qu'ils contiennent (phages et prophages) pourraient aussi mener à l'émergence de nouveaux types viraux, comme dans le cas des chimères, dont le gène de capside est proche de virus à ARN, et la protéine de réplication proche de virus à ADN simple brin (Diemer and Stedman, 2012 ; Roux et al., 2013). On a aussi découvert des virus appelés virophages, qui sont eux-mêmes dépendants d'autres virus (Claverie and Abergel, 2009; Christie and Dokland, 2012). Ces découvertes relancent parfois le débat sur le statut « vivant acellulaire » des virus (Moreira and Lopez-Garcia, 2009 ; Krupovic and Cvirkeite-Krupovic, 2011), ou même sur le statut de simple parasite du métabolisme de l'hôte (Weinbauer ; 2004 ; Paul, 2008 ; Feiner et al., 2015), mais confirment bien les liens très ténus entre le génome viral et celui de l'hôte, mais aussi entre le métabolisme de l'hôte et la diversité virale observable dans les communautés de virus libres (Sime-Ngando, 2014).

Du fait de la petite taille des génomes contenus dans les capsides virales de moins de 200nm, il est délicat d'extraire les particules virales et leurs acides nucléiques, et il faut souvent passer par des étapes d'amplification des acides nucléiques qui ont tendance à 'suramplifier' les séquences de virus à ADN simple brin, qui sont donc actuellement plus décrits (Thurber et al., 2009 ; Steward and Culley, 2010 ; Labonté and Suttle, 2013 ; Roux et al., 2013). Plus récemment certaines études ont ciblé des types de virus environnementaux moins étudiés comme les virus à ARN (Culley et al., 2006), qui infectent notamment les eucaryotes, ou encore les prophages induits (McDaniel et al., 2014; Brum et al., 2015). Il aura malgré tout plus fallu attendre plus d'une dizaine d'année pour la production par

pyroséquençage de métaviromes du sol (Zablocki et al., 2014; Adriaenssens et al., 2015). Les virus infectant les eucaryotes des environnements naturels restent actuellement très peu décrits (Sime-Ngando, 2014)

4.7. La structuration des communautés microbiennes par les virus

Différents concepts ont émergé en lien avec la diversité des communautés microbiennes, et l'abondance et la diversité des virus. Le concept de« killing the winner » (Thingstad et Lignell, 1997) (Figure 12) correspond à la dominance d'un hôte par son abondance et son activité et à sa plus grande susceptibilité de se faire infecter à cause de sa production virale (Schroeder et al., 2003 ; Rodriguez-Brito et al., 2010). Le concept de « seed bank model » (Breitbart and Rohwer, 2005) correspond au fait qu'en parallèle de cette sur-dominance d'un type viral, on retrouve dans les métaviromes une très grande diversité de génomes faiblement abondants mais capables de d'infecter des hôtes faiblement actifs ou peu représentés (Short et al., 2011 ; Pagarete et al., 2013). La conséquence est que le turn-over de la communauté bactérienne est assuré par une très forte productivité bactérienne et virale du « winner », dont la lyse libère de la matière organique utilisée pour le développement et la dominance d'autres bactéries (Middelboe et al., 2003), infectables par des virus « peu abondants », et ainsi de suite. De ce fait, les virus sont considérés comme des éléments structurant la diversité de la communauté microbienne sous jacente (Simek et al., 2001 ; Fuhrman and Schwalbach, 2003 ; Rohwer and Thurber, 2009).

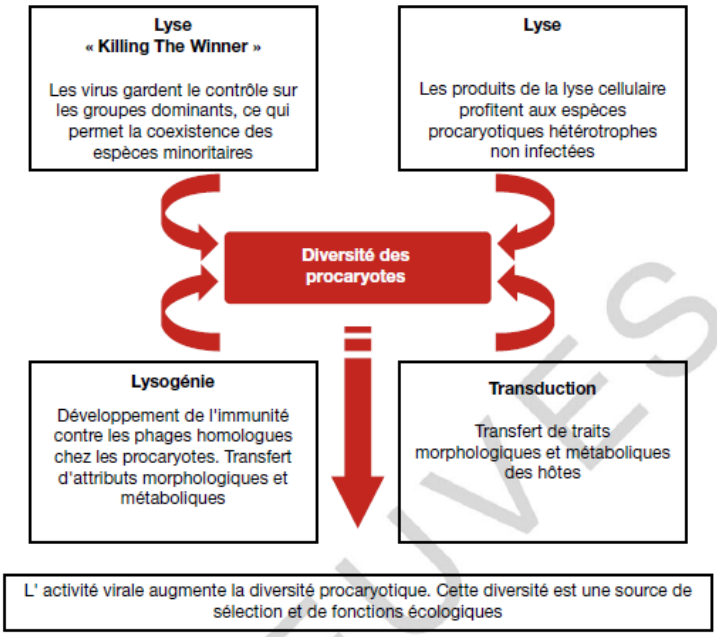


Figure 12 – Schéma conceptuel résumant les actions potentielles des virus sur les communautés de procaryotes (D'après Weinbauer et Rassoulzadegan 2004 *in* Berdjeb et Jacquet 2009b).

4.8. « Viral-shunt », boucles virale et microbienne

Les virus phages et prophages jouent un rôle déterminant (qualitativement et quantitativement) dans le contrôle des communautés bactériennes par leur abondance, leur diversité et leur impact sur la mortalité et la distribution des procaryotes (Bratbak et al., 1990 ; Bratbak et al., 1994 ; Simek et al., 2001; Berdjeb et Jacquet 2009 ; Rodriguez-Valera et al., 2009 ; Lauro et al., 2011). Ils représentent donc au même titre que les Protozoaires un facteur top-down affectant et régulant l'activité bactérienne et le recyclage du carbone (Figure 13) (Bratbak et al., 1990 ; Miki et al., 2008 ; Bouvy et al., 2011 ; Maurice et al., 2013 ; Chow et al., 2014). Cependant, la matière organique des hôtes lysés n'est pas transférée aux niveaux supérieurs du réseau trophique via le broutage des bactéries par les protistes mais est recyclée dans la fraction dissoute et à nouveau disponible pour l'activité bactérienne, d'où le concept de « shunt-viral » (« court-circuit » de la boucle microbienne) ou de boucle virale, ce qui confère aux virus un rôle clé dans les cycles biogéochimiques et le recyclage des éléments (Furhman, 1999 ; Middelboe and Lyck, 2002 ; Middelboe et al., 2003, Weinbauer et al., 2011). Ce processus serait responsable, dans les océans, du détournement de 6 à 26 % du flux de carbone initialement photosynthétisé, allant vers les consommateurs secondaires (Wilhelm and Suttle, 1999).

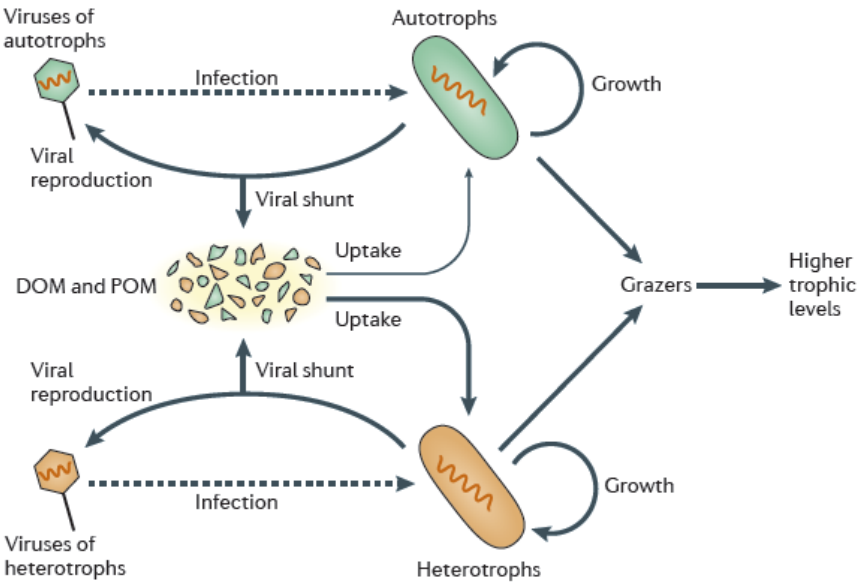


Figure 13 – Schéma du viral-shunt. La lyse virale permet la libération de matière organique dissoute et particulaire, ré-introduite dans la boucle microbienne plutôt que dans les niveaux supérieurs de consommation (D'après Wetz & Wilhelm 2012 *in* Jovel et al. 2014).

La lyse virale constitue une source de molécules organiques labiles rapidement utilisée par les microorganismes, accélérant le turn-over des éléments dans l'écosystème (Middelboe and Lyck,

2002 ; Middelboe et al., 2003 ; Weinbauer et al., 2011). Le lysat se révèle être de nature différente du stock initial de matière organique et de ce fait consommable par d'autres bactéries, engendrant une diversité plus grande des communautés bactériennes (Middelboe et al., 2003 ; Weinbauer et al. 2003, Berdjeb et Jacquet 2009 ; Weinbauer et al., 2011 ; Ankhrah et al., 2015). Enfin, les virus eux-mêmes constituent un stock de matière organique qui entre dans les voies de décomposition, contribuant également de cette façon au fonctionnement de l'écosystème (Dell'Anno et al., 2015). L'ensemble des processus, liés à l'activité des virus en font une sphère d'influence, la viriosphère (Berdjeb et Jacquet 2009), au même titre que les macroorganismes comme les plantes (rhizosphère) ou les vers (drilosphère).

4.9. Les limites actuelles de l'approche 'omic' et de l'écologie virale

Il est important de noter que l'étude de la diversité et de l'écologie virale reste complexe et en perpétuelle évolution, notamment grâce aux rapides avancées sur le plan technique et théorique.

Le concept de « Kill-the-Winner » a été proposé avant même que ne soit révélée la grande diversité des communautés virales environnementales (Fuhrman and Schwalbach, 2003). La modélisation du rôle des virus dans les écosystèmes reste encore limitée par la difficulté à déterminer quel virus infecte quel hôte, difficulté pour laquelle l'une des pistes de recherche est la détection de génomes viraux dans les banques de génomes bactériens (Roux et al., 2015b). Le degré de spécificité d'un virus pour un hôte (Miki et al., 2008) ou une communauté (Vos et al., 2009 ; Koskella and Meaden, 2013 ; Winter et al., 2013) est lui aussi encore peu connu, notamment en fonction de l'activité de ces derniers (Poisot et al., 2011 ; Deng et al., 2012). De plus, on ne connait souvent pas la vitesse à laquelle les virus s'adsorbent sur les hôtes et les infectent, et le devenir de ceux qui ne s'adsorbent pas (Winter et al., 2010 ; Deng et al., 2012 ; Gallet et al., 2012). Enfin, la capacité d'infection des virus retrouvés dans les métaviromes est en effet rarement discutée.

Les concepts du « Kill-the-Winner » et du « shunt viral » s'intéressent aussi aux changements saisonniers d'influence de la viriosphère en lien avec l'activité de communauté d'hôte (Rodriguez-Brito et al., 2010 ; Short at al., 2011 ; Maurice et al., 2013 ; Brum et al., 2015). La question des fluctuations saisonnières de la diversité des communautés virales a aussi été abordée de manière plutôt récente (Lopèz-Bueno et al., 2009 ; Rodriguez-Brito et al., 2010) et porte essentiellement sur certains types de virus dont les hôtes sont connus, via le séquençage d'un gène marqueur d'un type viral (comme celui d'une protéine de capside, ou de l'ADN polymerase) (Short et al., 2011 ; Chow et al., 2012 ; Pagarete et al., 2013). Elle a permis de montrer la saisonnalité des communautés de virus marins, ainsi que des récurrences annuelles (Pagarete et al., 2013).

Dans le cadre du « seed-bank model » et de la grande masse de séquences virales non identifiées, d'autres questions subsistent également. Si les génomes viraux évoluent rapidement, peut-on déterminer les facteurs structurant les communautés virales (Held and Whitaker, 2009, Brum et al., 2015 ; Roux et al., 2015a) ? Les virus peuvent-ils migrer entre des écosystèmes géographiquement distants (Breitbart and Rohwer, 2005 ; Snyder at al., 2007; Hewson et al., 2012, Winter et al., 2013)? Existe-t-il comme pour les procaryotes des sets de gènes communs aux virus d'un même écosystème, révélant leurs interactions avec les communautés d'hôtes de leur environnement (Brum et al., 2015b) ?

Questions de recherche et objectifs de la thèse

Au regard tant des réponses apportées que des questions soulevées par l'utilisation des nouvelles technologies de séquençage, ces dernières représentent actuellement l'un des plus puissants outils pour appréhender la diversité des communautés virales et microbiennes, et augmenter notre compréhension de leur fonctionnement et de leurs interactions. Combiner ces outils à l'utilisation d'autres techniques (microscopie électronique, comptages en cytométrie, lien avec la physico-chimique) représente actuellement le moyen le plus puissant pour l'investigation et la compréhension du fonctionnement des compartiments viraux et microbiens des écosystèmes.

En conséquence, étant donné (i) le rôle des virus dans la structuration des communautés microbiennes , (ii) l'absence de données sur les communautés virales des tourbières et l'importance de ces écosystèmes à l'échelle mondiale dans la régulation du climat (fonctionnement 'puits' de carbone), l'objectif général de ce travail de thèse était de comprendre comment les communautés de virus des tourbières se structurent et quelles pouvaient être les principales variables environnementales de forçage conduisant à leur organisation. Nous nous sommes focalisés sur les virus de petite taille (< 2µm) et notre recherche s'est appuyée sur un réseau de sites expérimentaux en régions tempérée et boréale, permettant des études en conditions naturelles et modifiées (simulation du réchauffement climatique, restauration écologique). Nous avons pour cela : i) utilisé différentes techniques complémentaires : séquençage nouvelle génération, bioinformatique, cytométrie en flux ; ii) caractérisé physico-chimiquement le milieu environnemental et iii) travaillé à différents niveaux d'organisation spatio-temporelle : fluctuations saisonnières, dynamique spatiale (stade dynamique, profondeur).

Le chapitre I a pour objectifs :

i) de dégager des patrons de distribution spatiale et temporelle des communautés de virus et de procaryotes, en comparant les fluctuations saisonnières d'abondance et de diversité, dans deux stades dynamiques clés de la succession écologique en tourbière, en surface dans l'acrotelme (globalement aérobie) et plus en profondeur, dans le mésotelme marqué par l'alternance de périodes aérobies et anaérobies, constituant une interface avec le catotelme, constamment saturé en eau et anaérobie ;

ii) d'identifier des variables physico-chimiques de contrôle des communautés de virus et de procaryotes.

Afin de répondre à ces objectifs, nous avons, dans une tourbière à *Sphagnum* des Monts du Forez (Puy-de-Dôme), effectué des comptages en cytométrie en flux (virus et procaryotes), établi des profils spatio-temporels de variables physico-chimiques clés (T°C, oxygène, pH, nutriments) et identifié la diversité des communautés virales et microbiennes par l'analyse de 12 metaviromes (métagénomique et analyses bioinformatiques).

Le chapitre II a pour objectif de parvenir à mieux connaître le fonctionnement des communautés microbiennes avec lesquelles interagissent les communautés virales observées dans le chapitre I. Cela nécessitait donc de:

i) comprendre la structuration de la communauté microbienne active, d'un point de vue taxonomique et fonctionel, en fonction du stade dynamique, et de la profondeur (acrotelme vs mésotelme)

ii) de déterminer si cette structuration de l'expression de gènes est basée sur des communautés adaptées et différentes en fonction des facteurs structurants.

Les communautés microbiennes ont été observées via 12 métatranscriptomes et 12 métagénomes représentant l'acrotelme et le mésotelme du fen et du bog (analyses bioinformatiques). Ces jeux de données ont été produits depuis la même tourbière que celle étudiée dans le chapitre I.

Les objectifs du **chapitre III** sont de:

i) vérifier si les communautés virales de la matrice tourbe sont distinctes de celles observées dans l'eau circulant dans le tapis de sphaignes et les pores des hyalocytes, et si l'analyse des communautés de la matrice de tourbe permettent d'obtenir une fine résolution spatiale de la diversité virale

i) vérifier l'existence de patrons spatiaux de diversité virale, i.e. fen *vs* bog et surface (acrotelme) *vs* profondeur (mésotelme), et de les confronter aux résultats des deux précédents chapitres afin de renforcer notre compréhension des interactions entre le compartiment virale et microbien

iii) déterminer s'il existe une spécificité de l'écosystème tourbière à Sphaignes en terme de diversité virale par comparaison de métaviromes identifiés sur des échantillons prélevés à la même période dans différentes sites des régions boréales (Canada, Finlande), et par comparaison avec les métaviromes issus d'autres écosystèmes

Pour cela, nous avons adapté une deuxième méthode d'extraction à partir de la matrice de tourbe afin d'optimiser la récupération des particules virales. Nous avons échantillonné l'acrotelme et le mésotelme de 5 tourbières réparties en France (tourbières de Frasne, et de la Guette), au Québec (tourbière de Bois-des-Bel), en Finlande (tourbière de Lakkasuo) et sur l'île subantarctique d'Amsterdam. Nous avons procédé à des comptages en cytométrie, à de la mesure de paramètres physico-chimiques et à la production de métaviromes de virus issus de l'eau des Sphaignes et de la matrice de tourbe, et issus des deux couches (acrotelme et mésotelme) de la tourbière du Québec, et du fen et du bog de Finlande.

Le chapitre IV est consacré à la réponse des communautés virales aux changements globaux et plus précisément au changement climatique et au changement d'usage des terres. Les questions étaient les suivantes :

i) comment un réchauffement modéré de la température de l'air (< 3°C) et les variations de nappe se traduisent-elles à court et moyen terme ?

ii) la restauration d'une tourbière anciennement exploitée par ensemencement de propagules de sphaignes a-t-elle un impact positif ?

iii) quel est l'impact de la ré-humectation d'une tourbière drainée sur les communautés microbiennes via l'observation de la diversité virale

Pour cela, nous avons d'une part :

i) comparé les abondances de virus et de procaryotes dans les sites expérimentaux de Frasne dans le Doubs (INSU-SO Tourbières) et de Mukhrino en Sibérie occidentale (Chaire UNESCO) permettant de mesurer les effets à moyen (Frasne, 6 ans) et court terme (Mukhrino, 1 an). Le dispositif de la station sibérienne permet également de tester l'effet d'un abaissement et d'un réhaussement de la nappe d'eau, ce dernier aspect étant également traitée à la tourbière de La Guette (Cher, INSU-SO Tourbières) dont les drains ont été bouchés pour permettre la remontée de la nappe dans la tourbière.

Et, d'autre part :

ii) comparé le complexe de tourbières restaurée (anciennement exploitée industriellement) et naturelle du site de Bois-des-Bel (Québec).

Enfin une synthèse générale reprend les acquis principaux de cette thèse qui constitue, à notre connaissance, le premier travail sur l'écologie des virus de tourbières. De nouvelles perspectives de recherche concluent ce mémoire.

Références

Abergel C., Claverie J.M. 2014. Pithovirus Sibericum: Awakening of a Giant Virus of More than 30,000 Years. *Médecine Sciences : M/S* 30 (3): 329–31. doi:10.1051/medsci/20143003022.

Abrescia N.G.A., Bamford D.H, Grimes J.M, and Stuart D.I. 2012. Structure Unifies the Viral Universe. *Annual Review of Biochemistry* 81 (January). ANNUAL REVIEWS, 4139 EL CAMINO WAY, PO BOX 10139, PALO ALTO, CA 94303-0897 USA: 795–822. doi:10.1146/annurev-biochem-060910-095130.

Adriaenssens E.M., Van Zyl L., De Maayer P., Rubagotti E., Rybicki E., Tuffin M., Cowan D.A. 2015. Metagenomic Analysis of the Viral Community in Namib Desert Hypoliths. *Environmental Microbiology* 17 (2): 480–95. doi:10.1111/1462-2920.12528.

Andersen R., Francez A.-J., Rochefort L. 2006. The physico-chemical and microbiological status of a restored bog in Québec: identifications of relevant criteria to monitor success. *Soil Biology and Biochemistry* 38: 1375-1387.

Andersen R., Grasset L., Thormann M.N., Rochefort L., Francez A.-J. 2010. Changes in microbial community structure and function following *Sphagnum* peatland restoration. *Soil Biology and Biochemistry* 42: 291-301.

Andersen R., Wells C., Macrae M., Price J. 2013. Nutrient Mineralisation and Microbial Functional Diversity in a Restored Bog Approach Natural Conditions 10 Years Post Restoration. *Soil Biology and Biochemistry* 64 (September). PERGAMON-ELSEVIER SCIENCE LTD, THE BOULEVARD, LANGFORD LANE, KIDLINGTON, OXFORD OX5 1GB, ENGLAND: 37–47. doi:10.1016/j.soilbio.2013.04.004.

Ankrah N.Y.D., May A.L., Middleton J.L., Jones D.R., Hadden M.K., Gooding J.R., LeCleir G.R., Wilhelm S.W., Campagna S.R., Buchan A. 2014. Phage Infection of an Environmentally Relevant Marine Bacterium Alters Host Metabolism and Lysate Composition. *The ISME Journal* 8 (5). Nature Publishing Group: 1089–1100. doi:10.1038/ismej.2013.216.

Artz R.R.E., Anderson I.C., Chapman S.J., Hagn A., Schloter M., Potts J.M., Campbell C.D. 2007. Fungal diversity and community composition change in response to vegetational succession during natural regeneration of cut-over peatlands. *Microbial Ecology* 54: 508-522.

Artz R.R.E., Chapman S.J., Siegenthaler A., Mitchell E.A.D., Buttler A., Bortoluzzi E., Gilbert D., Yli-Petays M., Vasander H., Francez A.-J. 2008. Functional microbial diversity in cutover peatlands responds to vegetation succession and is partly directed by labile carbon. *J. Appl. Ecol.* 45: 1799-1809.

Artz R.E.E. 2009. Microbial community structure and carbon substrate use in northern peatlands. In: Baird A.J., Belyea L.R., Comas X., Reeve A.S., Slater L.D. (eds), Carbon cycling in northern peatlands. *Geophysical Monograph Series* 184: 111-129.

Ashelford K.E., Day M.J., Fry J.C. 2003. Elevated Abundance of Bacteriophage Infecting Bacteria in Soil. *Applied and Environmental Microbiology* 69 (1): 285–89. doi:10.1128/AEM.69.1.285-289.2003.

Azam F., Fenchel T., Field J., Gray J., Meyer-Reil J., Thingstad F. 1983. The ecological role of water-column microbes in the sea. *Marine Ecology Processes Series* 10: 257-263.

Bardgett R.D., Freeman C., Ostle N.J. 2008. Microbila contributions to climate change through carbon cycle feedbacks. *The ISME Journal* 2: 805-814.

Basiliko N., Yavitt J.B., Dees P.M., Merkel S.M. 2003. Methane biogeochemistry and methanogen communities in two northern peatland ecosystems, New York State. *Geomicrobiol. J.* 20: 563–577.

Basiliko N., Henry K., Gupta V., Moore T.R., Driscoll B.T., Dunfield P.F. 2013. Controls on bacterial and archaeal community structure and greenhouse gas production in natural, mined, and restored Canadian peatlands. *Frontiers in Microbiology* 4, article 215, doi: 10.3389/fmicb.2013.00215

Bauer I. 2004. Modelling effects of litter quality and environment on peat accumulation over different time-scales. *J. Ecol.* 92: 661-674.

Berdjeb L., Jacquet S. 2009a. La viriosphère : quelle place dans le fonctionnement et l'évolution des écosystèmes aquatiques (partie 1) ? *Virologie* 13 (3): 133-143.

Berdjeb L., Jacquet S. 2009b. La viriosphère : quelle place dans le fonctionnement et l'évolution des écosystèmes aquatiques (partie 2) ? *Virologie* 13 (4): 1-11.

Björk R.G., Ernfors M. Sikström U., Nilsson M.B., Andersson M.X., Rütting, Klemedtsson L. 2010. Contrasting effects of wood ash application on microbial community structure, biomass and processes in drained forested peatlands. *FEMS Microbiol. Ecol.* 73: 550-562.

Blok D., Heijmans M.P.D., Schaepman-Strub G., van Ruijven J., Parmentier F.J.W., Maximov T.C., Berendse F. 2011. The cooling capacity of mosses: controls on water and energy fluxes in a Siberian tundra site. *Ecosystems* 14: 1055-1065.

Bongowski M. 2003. Protozoa and plant growth: the microbial loop in soil revisited. *New Phytol.* 162: 617-631.

Borrel G., Colombet J., Robin A., Lehours A.-C., Prangishvili D., and Sime-Ngando T. 2012. Unexpected and Novel Putative Viruses in the Sediments of a Deep-Dark Permanently Anoxic Freshwater Habitat. *The ISME Journal* 6 (11). Nature Publishing Group: 2119–27. doi:10.1038/ismej.2012.49.

Bos L. 1999. Beijerinck's Work on Tobacco Mosaic Virus: Historical Context and Legacy. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 354 (1383): 675–85. doi:10.1098/rstb.1999.0420.

Bouvy M., Bettarel Y., Bouvier C., Domaizon I., Jacquet S., Le Floc'h E., Montanié H., et al. 2011. Trophic Interactions between Viruses, Bacteria and Nanoflagellates under Various Nutrient Conditions and Simulated Climate Change. *Environmental Microbiology* 13 (7): 1842–57. doi:10.1111/j.1462-2920.2011.02498.x.

Bragazza L., Bardgett R.D., Mitchell E.A.D., Buttler A. 2015. Linking soil microbial communities to vascular plant abundance along a climate gradient. New Phytol. 205: 1175-1182.

Bragina A., Berg C., Müller H., Moser D., and Berg G. 2013. Insights into Functional Bacterial Diversity and Its Effects on Alpine Bog Ecosystem Functioning. *Scientific Reports* 3 (January). Nature Publishing Group: 1955. doi:10.1038/srep01955.

Bragina, A., Oberauner-Wappis L., Zachow C., Halwachs B., Thallinger G. G., Müller H., and Berg G. 2014. The Sphagnum Microbiome Supports Bog Ecosystem Functioning under Extreme Conditions. *Molecular Ecology*. doi:10.1111/mec.12885.

Bratbak G., Heldal M., Norland S., Thingstad T.F. 1990. Viruses as Partners in Spring Bloom Microbial Trophodynamics. *Applied and Environmental Microbiology* 56 (5): 1400–1405. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=184418&tool=pmcentrez&rendertype=abstract.

Bratbak G., ThingstadF., Heldal M. 1994. Viruses and the Microbial Loop. *Microbial Ecology* 28 (2): 209–21. doi:10.1007/BF00166811.

Breeuwer A.M., Heijmans B.J., Robroek B.J.M., Limpens J., Berendse F. 2008. The effect of increased temperature and nitrogen deposition on decomposition in bogs. *Oikos* 117: 1258-1268.

Breitbart M., Rohwer F. 2005. Here a virus, there a virus, everywhere the same virus ? *Trends in Microbiology* 13: 278-284.

Breitbart M., Salamon P., Andresen B., Mahaffy J.M., Segall A.M., Mead D., Azam F., Rohwer F. 2002. Genomic Analysis of Uncultured Marine Viral Communities. *Proceedings of the National Academy of Sciences of the United States of America* 99 (22): 14250–55. doi:10.1073/pnas.202488399.

Bridgham S.D., Richardson C.J. 1992. Mechanisms controlling soil respiration (CO_2 and CH_4) in southern peatlands. *Soil Biol. Biochem.* 24: 1089–1099.

Bridgham S.D., Pastor J., Janssens J.A., Chapin C., Malterer T., 1996. Multiple limiting gradients in peatlands: a call for a new paradigm. *Wetlands* 16: 45-65.

Brum J.R., Hurwitz B.L., Schofield O., Ducklow H.W., Sullivan M.B. 2015. Seasonal Time Bombs: Dominant Temperate Viruses Affect Southern Ocean Microbial Dynamics. *The ISME Journal*, August. Nature Publishing Group. doi:10.1038/ismej.2015.125.

Brum J.R., Ignacio-Espinoza J.C., Roux S., Doulcier G., Acinas S.G., Alberti A., Chaffron S., et al. 2015. Ocean Plankton. Patterns and Ecological Drivers of Ocean Viral Communities. *Science (New York, N.Y.)* 348 (6237): 1261498. doi:10.1126/science.1261498.

Brussaard C.P.D., Marie D., Bratbak G. 2000. Flow Cytometric Detection of Viruses. *Journal of Virological Methods* 85 (1-2): 175–82. doi:10.1016/S0166-0934(99)00167-6.

Brussaard C.P.D. 2004. Optimization of Procedures for Counting Viruses by Flow Cytometry. *Applied and Environmental Microbiology* 70 (3): 1506–13. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=368280&tool=pmcentrez&rendertype=abstract.

Buttler A., Robroek B., Laggoun-Défarge F., Jassey V.E.J., Pochelon C., Bernatd G., Delarue F., Gogo S., Mariotte P., Mitchell E.A.D., Bragazza L. 2015. Experimental warming interacts with soil moisture to discriminate plant responses in an ombrotrophic peatland. *Journal of Vegetation Science* 26: 964-974.

Chow, C.E.T., and Jed A. Fuhrman. 2012. "Seasonality and Monthly Dynamics of Marine Myovirus Communities." *Environmental Microbiology* 14 (8): 2171–83. doi:10.1111/j.1462-2920.2012.02744.x.

Chow C.E.T., Kim D.Y., Sachdeva R., Caron D.A., Fuhrman J.A. 2014. Top-down Controls on Bacterial Community Structure: Microbial Network Analysis of Bacteria, T4-like Viruses and Protists. *The ISME Journal* 8 (4): 816–29. doi:10.1038/ismej.2013.199.

Christie G.E., Dokland T. 2012. Pirates of the Caudovirales. *Virology*, no. 0. Accessed December 4. doi:10.1016/j.virol.2012.10.028.

Clarholm M., Rosswall T. 1980. Biomass and turnover of Bacteria in a forest soil and apeat. *Soil Biol. Biochem.* 12, 49-57.

Clarholm M. 1994. The microbial loop in soil. In: Ritz K., Dighton J., Giller K.E. (eds), *Beyond the biomass. Compositional and functional analysis of soil microbial communities,* British Society of Soil Science, John Wiley & Sons, Chichester, pp. 221-230.

Claverie. J.M., Abergel C. 2009. Mimivirus and Its Virophage. *Annual Review of Genetics* 43 (January): 49–66. doi:10.1146/annurev-genet-102108-134255.

Clymo R.S., Hayward P.M. 1982. The ecology of *Sphagnum*. In: Smith A. (ed), Bryophyte ecology, Chapman and Hall, London & New-York, pp. 229-289.

Clymo R.S., Bryant C.L. 2008. Diffusion and mass flow of dissolved carbon dioxide, methane, and dissolved organic carbon in a 7-m deep raised peat bog. *Geochim. Cosmochim. Acta* 72:2048-2066.

Collins V.G., D'Sylva B.T., Latter P.M. 1978. Microbial populations in peat. In: Heal O.W. & Perkins D.F. (eds), *Production Ecology of British moors and montane grasslands,* Springer-Verlag, Berlin, pp. 631-650.

Coûteaux M.-M. Faurie G., Palka L., Stienberg C. 1988. La relation prédateur-proie (protozoaires-bactéries) dans les sols : rôle dans la régulation des populations et conséquences sur les cycles du carbone et de l'azote. *Revue d'Ecologie et de Biologie du Sol* 25 : 1-31.

Croft M., Rochefort L., Beauchamp C.J. 2001. Vacuum-extraction of peatlands disturbs bacterial population and microbial biomass carbon. *Applied Soil Ecology* 18: 1-12.

Culley A.I., Lang A.S., Suttle C.A. 2006. Metagenomic Analysis of Coastal RNA Virus Communities. *Science* 312 (5781): 1795–98. doi:10.1126/science.1127404.

De Barjac, H. 1955. *Essai d'interprétation bactériologique de sols tourbeaux acides*. Thèse de doctorat en Sciences Naturelles. Université de Paris (France). 160p.

Dedysh, S. N. 1998. Isolation of Acidophilic Methane-Oxidizing Bacteria from Northern Peat Wetlands. *Science* 282 (5387): 281–84. doi:10.1126/science.282.5387.281.

Dedysh S.N., Pankratov T.A., Belova S.E., Kulichevskaya I.S., Liesack W. 2006. Phylogenetic analysis and *in situ* identification of bacteria community composition in an acidic *Sphagnum* peat bog. *Appl. Environ. Microbiol.* 72 : 2110–2117.

Delarue F., Gogo S., Buttler A., Bragazza L., Jassey V.E.J., Bernard G., Laggoun-Défarge F. 2014. Indirect effects of experimental warming on dissolved organic carbon content in subsurface peat. *J. Soils Sediments* 14: 1800-1805.

Delarue F., Buttler A., Bragazza L., Grasset L., Jassey V., Gogo S., Laggoun-Défarge F. 2015. Experimental warming differentially affects microbial structure and activity in two contrasted moisture sites in a Sphagnum-dominated peatland. *Science of the Total Environment* 511: 576-583.

Dell'Anno A., Corinaldesi C., Danovaro R. 2015. Virus decomposition provides an important contribution to benthic deep-sea ecosystem functioning. *Proc. Natl. Acad. Sci. U.S.A.* 112: E2014-E2019.

Deng, Li, Ann Gregory, Suzan Yilmaz, Bonnie T Poulos, Philip Hugenholtz, and Matthew B Sullivan. 2012. Contrasting Life Strategies of Viruses That Infect Photo- and Heterotrophic Bacteria, as Revealed by Viral Tagging. *mBio* 3 (6). doi:10.1128/mBio.00373-12.

Desnues C., Rodriguez-Brito B., Rayhawk S., Kelley S, Tran T., Haynes M., Liu H., et al. 2008. Biodiversity and Biogeography of Phages in Modern Stromatolites and Thrombolites. *Nature* 452 (7185): 340–43. doi:10.1038/nature06735.

Diemer G.S., and Stedman K.M. 2012. A Novel Virus Genome Discovered in an Extreme Environment Suggests Recombination between Unrelated Groups of RNA and DNA Viruses. *Biology Direct* 7 (January): 13. doi:10.1186/1745-6150-7-13.

Dise N.B., Phoenix G.K. 2011. Peatlands in a changing world. *New Phytol.* 191: 309-311.

Dorrepaal E., Toet S., van Loegtestijn R.S.P., Swart E., van de Weg M.J., Callaghan T.V., Aerts R. 2009. Carbon respiration from subsurface peat accelerated by climate warming in the subarctic. *Nature* 460: 616-619.

Duhamel S, and Jacquet S. 2006. Flow Cytometric Analysis of Bacteria- and Virus-like Particles in Lake Sediments. *Journal of Microbiological Methods* 64 (3): 316–32. doi:10.1016/j.mimet.2005.05.008.

Edwards R.A., and Rohwer F. 2005. Viral Metagenomics. *Nature Reviews. Microbiology* 3 (6): 504–10. doi:10.1038/nrmicro1163.

Ekelund F., Saj S., Vestergard M., Bertaux J., Mikola J. 2009. The "soil microbial loop" is not always needed to explain protozoan stimulation of plants. *Soil Biology and Biochemistry* 41: 2336-2342.

Fancello L., Trape S., Robert C., Boyer M., Popgeorgiev N., Raoult D., Desnues D. 2012. Viruses in the Desert: A Metagenomic Survey of Viral Communities in Four Perennial Ponds of the Mauritanian Sahara. *The ISME Journal*, April. doi:10.1038/ismej.2012.101.

Feiner R., Argov T., Rabinovich L., Sigal N., Borovok I., Herskovits A.A. 2015. A New Perspective on Lysogeny: Prophages as Active Regulatory Switches of Bacteria. *Nature Reviews Microbiology* 13 (10). Nature Publishing Group: 641–50. doi:10.1038/nrmicro3527.

Fenner N., Freeman C., Lock M.A., Harmens H, Reynolds B., and Sparks T. 2007. Interactions between Elevated CO 2 and Warming Could Amplify DOC Exports from Peatland Catchments. *Environmental Science & Technology* 41 (9). American Chemical Society: 3146–52. doi:10.1021/es061765v.

Forterre P., Soler N., Krupovic M., Marguet E., Ackermann H.W. 2013. Fake Virus Particles Generated by Fluorescence Microscopy. *Trends in Microbiology* 21 (1): 1–5. doi:10.1016/j.tim.2012.10.005.

Francez A.-J., Bignon J.-J., Mollet A.-M. 1992. The peatlands in France: localization, characterization, use and conservation. *Suo* 43: 11-24.

Francez A.-J. 1995. Dynamique du carbone et de l'azote chez le *Carex rostrata,* l'*Eriophorum vaginatum* et le *Calluna vulgaris* dans une tourbière des Monts du Forez (France). *Can. J. Bot.* 73: 121-129.

Francez A.-J. 2000. La dynamique du carbone dans les tourbières à Sphagnum, de la sphaigne à l'effet de serre. *L'Année biologique* 39: 205-270.

Frolking S., Roulet N.T., Moore T.R., Richard P.J.H., Lavoie M. and Muller S.D. 2001. Modelling northern peatland decomposition and peat accumulation. *Ecosystems* 4: 479–498.

Fuhrman J.A. 1999. Marine Viruses and Their Biogeochemical and Ecological Effects. *Nature* 399 (6736): 541–48. doi:10.1038/21119.

Fuhrman J.A., and Schwalbach M. 2003. Viral Influence on Aquatic Bacterial Communities. *Biol. Bull.* 204 (2): 192–95. http://www.biolbull.org/content/204/2/192.full.

Gallet R., Lenormand T., Wang I-N. 2012. Phenotypic Stochasticity Protects Lytic Bacteriophage Populations from Extinction during the Bacterial Stationary Phase. *Evolution; International Journal of Organic Evolution* 66 (11): 3485–94. doi:10.1111/j.1558-5646.2012.01690.x.

Gauthier, R. 2001. Les sphaignes boréles. *Le naturaliste canadien*. 25 (3) : 180-185.

Gicquel A. 2012. *Impact des changements globaux sur le fonctionnement des tourbières: couplage C-N-S et interactions biotiques.* Thèse de l'Université de Rennes 1 : 221 p.

Gilbert D., Amblard C., Bourdier G., Francez 1998. The microbial loop at the surface of a peatland: structure, function and impact of nutrient input. *Microb. Ecol.* 35: 83-93.

Gilbert D., Mitchell E.A.D. 2006. Microbial diversity in Sphagnum peatlands. In: Martini I.P., Martinez Cortizas A., Chesworth W. (eds), *Peatlands: evolution and records of environmental and climaqte changes.* Developments in Earth surface processes 9, Elsevier, Amsterdam: pp. 287-318.

Given P.H., Dickinson C.H. 1975. Biochemistry and microbiology of peats. In: Paul P.A. & Mc Laren A.D. (eds), *Soil Biochemistry,* vol. 3, chap. 4, Marcel Dekker, New-York, pp. 123-212.

Glatzel S., Kalbitz K., Dalva M., Moore T. 2003. Dissolved organic matter properties and their relationship to carbon dioxide efflux from restored peat bogs. *Geoderma* 113: 397-411.

Gorham, Eville. 1991. Northern Peatlands: Role in the Carbon Cycle and Probable Responses to Climatic Warming. 1 (2): 182–95. http://www.esajournals.org/doi/abs/10.2307/1941811.

Griffiths B.S. 1989. Enhanced nitrification in the presence of bacteriophagous protozoa. *Soil Biol. Biochem.* 21: 1045-1051.

van der Heijden, E., Boon, J.J., Rasmussen, S. and Rudolph, H. 1997. *Sphagnum* acid and its decarboxylation productisopropenyphenol as biomarkers for fossilized *Sphagnum* in peats. *Anc Biomol*. 1: 93–107.

Hättenschwiler S., Tiunov A., Scheu S. 2005. Biodiversity and litter decomposition in terrestrial ecosystems. *Annu. Rev. Ecol. Evol. Syst.* 36: 191-218.

Held N.L., and Whitaker R.J. 2009. Viral Biogeography Revealed by Signatures in Sulfolobus Islandicus Genomes. *Environmental Microbiology* 11 (2): 457–66. doi:10.1111/j.1462-2920.2008.01784.x.

Hewson, I., J. G. Barbosa, J. M. Brown, R. P. Donelan, J. B. Eaglesham, E. M. Eggleston, and B. A. LaBarre. 2012. Temporal Dynamics and Decay of Putatively Allochthonous and Autochthonous Viral Genotypes in Contrasting Freshwater Lakes. *Applied and Environmental Microbiology* 78 (18): 6583–91. doi:10.1128/AEM.01705-12.

Høj L., Olsen R.A., and Torsvik V.L. 2005. Archaeal Communities in High Arctic Wetlands at Spitsbergen, Norway (78 Degrees N) as Characterized by 16S rRNA Gene Fingerprinting. *FEMS Microbiology Ecology* 53 (1): 89–101. doi:10.1016/j.femsec.2005.01.004.

Høj L., Olsen R.A., and Torsvik V.L. 2007. Effects of Temperature on the Diversity and Community Structure of Known Methanogenic Groups and Other Archaea in High Arctic Peat. *The ISME Journal* 2 (1). International Society for Microbial Ecology: 37–48. doi:10.1038/ismej.2007.84.

Hultman J., Waldrop M.P., Mackelprang R., David M.M., McFarland J., Blazewicz S.J., Harden J., et al. 2015. Multi-Omics of Permafrost, Active Layer and Thermokarst Bog Soil Microbiomes. *Nature* 521 (7551). Nature Publishing Group: 208–12. doi:10.1038/nature14238.

Ingram H.A.P. 1978. Soil layers in mires: function and terminology. J. *Soil Sci.* 29:224-227.

Jassey V. 2011. *Impact d'un réchauffement climatique sur le fonctionnement de la sphagnosphère : relations polyphenols-communautés microbiennes.* Thèse de l'Université de Franche-Comté, Besançon (France): 238 p.

Jassey V.E.J, Chiapusio G., Mitchell E.A.D., Binet P., Toussaint M.-L, Gilbert D. 2011. Fine-Scale Horizontal and Vertical Micro-Distribution Patterns of Testate Amoebae along a Narrow Fen/Bog Gradient. *Microbial Ecology* 61 (2): 374–85. doi:10.1007/s00248-010-9756-9.

Jassey V.E.J., Shimano S., Dupuy C., Toussaint M.-L., Gilbert D. 2012. Characterizing the feeding habits of the Testate Amoebae *Hyalosphenia papilio* and *Nebela tincta* along a narrow "fen-bog" gradient using digestive vacuole content and ^{13}C and ^{15}N isotopic analyses. *Protist* 163: 451-464.

Jones C.G., Lawton J.H., Shachak M. 1994. Organisms as ecosystem engineers. *Oikos* 69, 373–386.

Joosten H., Clarke D. 2002. *Wise use of mires and peatlands: background and principles including a framework for decision-making.* International Mire Conservation Group & International Peat Society, Jyväskylä, 304 p.

Joosten H., Tapio-Biström M.-L., Tol S. (eds). 2012. *Peatlands – Guidance for climate change mitigation through conservation, rehabilitation and sustainable use.* Mitigation of climate change in agriculture series 5, FAO & Wetlands International (2nd ed): 101 p.

Jovel L., Effler T.C., Buchan A., Wilhelm S.W., Weitz J. 2014. The elemental composition of virus particles: implications for marine biogeochemical cycles. *Nature Reviews-Microbiology* 12: 519-528.

Juottonen H., Galand P.E., Tuittila E.-S., Laine J., Fritze H., and Yrjälä K. 2005. Methanogen Communities and Bacteria along an Ecohydrological Gradient in a Northern Raised Bog Complex. *Environmental Microbiology* 7 (10): 1547–57. doi:10.1111/j.1462-2920.2005.00838.x.

Kim S.-Y., Freeman C., Fenner N., Kang H. 2012. Functional and structural responses of bacterial and methanogen communities to 3-year warming incubation in different depths of peat mire. *Appl. Soil Ecology* 57: 23-30.

Kimmel, K., and Mander U. 2010. Ecosystem Services of Peatlands: Implications for Restoration. *Progress in Physical Geography* 34 (4): 491–514. doi:10.1177/0309133310365595.

King A.M.Q., Adams M.J., and Lefkowitz E.J. 2011. Virus Taxonomy: Classification and Nomenclature of Viruses : Ninth Report of the International Committee on Taxonomy of Viruses. *Elsevier*. https://books.google.com/books?id=KXRCYay3pH4C&pgis=1.

Kip N., Fritz C., Langelaan E.S., Pan Y., Bodrossy L., Pancotto V., Jetten M.S.M., Smolders A.J.P., Op den Camp H.J.M. 2012. Methanotrophic activity and diversity in different *Sphagnum magellanicum* dominated habitats in the southernmost peat bogs of Patagonia. *Biogeosciences* 9: 47-55.

Koonin E.V., and Dolja V.V. 2014. Virus World as an Evolutionary Network of Viruses and Capsidless Selfish Elements. *Microbiology and Molecular Biology Reviews* 78 (2): 278–303. doi:10.1128/MMBR.00049-13.

Koskella B., and Meaden S. 2013. Understanding Bacteriophage Specificity in Natural Microbial Communities. *Viruses* 5 (3): 806–23. doi:10.3390/v5030806.

Kotsyurbenko O.R., Chin K.-J., Glagolev M.V., Stubner S., Simankova M.V., Nozhevnikova A.N. Conrad R. 2004., Acetoclastic and hydrogenotrophic methane production and methanogenic populations in acidic West-Siberian peat bogs. *Environ Microbiol* 6: 1159–1173.

Krupovic M., and Cvirkaite-Krupovic V. 2011. Virophages or Satellite Viruses? *Nature Reviews Microbiology* 9 (11): 762–63. doi:10.1038/nrmicro2676.

Labonté J.M., and Suttle C.A. 2013. Previously Unknown and Highly Divergent ssDNA Viruses Populate the Oceans. *The ISME Journal* 7 (11). Nature Publishing Group: 2169–77. doi:10.1038/ismej.2013.110.

Lappalainen, E. (editor); 1996; Global Peat Resources; International Peat Society, Finland Couch, G.R.; 1993; Fuel peat - world resources and utilisation; IEA Coal Research, London.

Larmola T., Tuittila E.-S., Tiirola M., Nykänen H., Martikainen P.J., Yrjälä K., Tuomivirta T., Fritze H. 2010. The role of Sphagnum mosses in the methane cycling of a boreal mire. *Ecology* 91: 2356-2365.

Lauro F.M., DeMaere M.Z., Yau S., Brown M.V., Ng C., Wilkins D., Raftery M.J., et al. 2011. An Integrative Study of a Meromictic Lake Ecosystem in Antarctica. *The ISME Journal* 5 (5). International Society for Microbial Ecology: 879–95. doi:10.1038/ismej.2010.185.

Legendre M., Lartigue A., Bertaux L., Jeudy S., Bartoli J., Lescot M., Alempic J.-M., et al. 2015. In-Depth Study of Mollivirus Sibericum , a New 30,000-Y-Old Giant Virus Infecting Acanthamoeba. *Proceedings of the National Academy of Sciences* 112 (38): 201510795. doi:10.1073/pnas.1510795112.

Limpens J., Granath G., Gunnarsson U., Aerts R., Bayley S., Bragazza L., Bubier J., Buttler A., van den Berg L.J.L., Francez A.-J., Gerdol R., Grosvernier P., Heijmans M. M. P. D., Hoosbeek M. R., S. Hotes S., Ilomets M., Leith I., Mitchell E.A.D., Moore T., Nilsson M.B., Nordbakken J.-F., Rochefort L., Rydin H., Sheppard L.J., Thormann M., Wiedermann M.M., Williams B.L., Xu B. 2011. Climatic modifiers of the response to nitrogen deposition in peat-forming Sphagnum mosses: a meta-analysis. *New phytol* 191: 496–507.

Lin X., Green S., Tfaily M.M., Prakash O., Konstantinidis K.T., Corbett J.E., Chanton J.P., Cooper W.T., and Kostka J.E. 2012. Microbial Community Structure and Activity Linked to Contrasting Biogeochemical Gradients in Bog and Fen Environments of the Glacial Lake Agassiz Peatland. *Applied and Environmental Microbiology* 78 (19): 7023–31. doi:10.1128/AEM.01750-12.

Lin X., Tfaily M.M, Steinweg J.M., Chanton P., Esson K., Yang Z.K., Chanton J.P., Cooper W., Schadt C.W., Kostka J.E. 2014. Microbial Community Stratification Linked to Utilization of Carbohydrates and Phosphorus Limitation in a Boreal Peatland at Marcell Experimental Forest, Minnesota, USA. *Applied and Environmental Microbiology* 80 (11). American Society for Microbiology: 3518–30. doi:10.1128/AEM.00205-14.

Lindell D., Sullivan M.B., Johnson Z.I., Tolonen A.C., Rohwer F., and Chisholm S.W. 2004. Transfer of Photosynthesis Genes to and from Prochlorococcus Viruses. *Proceedings of the National Academy of Sciences of the United States of America* 101 (30). National Academy of Sciences: 11013–18. doi:10.1073/pnas.0401526101.

Lindo Z., Gonzalez A. 2010. The Bryosphere: an integral and influential component of the Earth's Biosphere. *Ecosystems* 13: 612-627.

López-Bueno A., Tamames J., Velázquez D., Moya A., Quesada A., Alcamí A. 2009. High Diversity of the Viral Community from an Antarctic Lake. *Science* 326 (5954): 858–61. doi:10.1126/science.1179287.

López-García P., and Moreira D. 2008. Tracking Microbial Biodiversity through Molecular and Genomic Ecology. *Research in Microbiology* 159 (1): 67–73. doi:10.1016/j.resmic.2007.11.019.

Loreau M. 2001. Microbial diversity, producer-decomposer interactions and ecosystem processes: a theoretical model. *Proc. R. Soc.London Ser. B* 268: 303–309.

Maire N. 1983. Etude du repeuplement d'une tourbe stérilisée par quatre méthodes biologiques globales (ATP, dégagement de CO_2, phosphatase et uréase). *Recherche agronm. en Suisse* 22: 221-246.

Manneville O., Vergne V., Villepoux O., Groupe d'Etudes des Tourbières. 2006. le monde des tourbières et des marais. France, Suisse, Belgique, Luxembourg. Delachaux et Niestlé, Paris, 320 p. (2nd édition)

Marie D., Vaulot D., and Partensky F. 1996. Application of the Novel Nucleic Acid Dyes YOYO-1, YO-PRO-1, and PicoGreen for Flow Cytometric Analysis of Marine Prokaryotes. *Appl. Envir. Microbiol.* 62 (5): 1649–55. http://aem.asm.org/content/62/5/1649.short.

Marie D., Brussaard C.P.D., Thyrhaug R., Bratbak G., Vaulot D. 1999. Enumeration of Marine Viruses in Culture and Natural Samples by Flow Cytometry. *Applied and Environmental Microbiology* 65 (1): 45–52. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=90981&tool=pmcentrez&rendertype=abstract.

Maurice C.F., Bouvier C., de Wit R., Bouvier T. 2013. Linking the Lytic and Lysogenic Bacteriophage Cycles to Environmental Conditions, Host Physiology and Their Variability in Coastal Lagoons. *Environmental Microbiology* 15 (9): 2463–75. doi:10.1111/1462-2920.12120.

McDaniel L.D., Rosario K., Breitbart M., and Paul J.H. 2014. Comparative Metagenomics: Natural Populations of Induced Prophages Demonstrate Highly Unique, Lower Diversity Viral Sequences. *Environmental Microbiology* 16 (2): 570–85. doi:10.1111/1462-2920.12184.

Merilä P., Galand P.E., Fritze H., Tuittila A.-S., Kukko-oja K., Laine J., Yrjälä K. 2006. Methanogen communities along a primary succession transect of mires ecosystems. *FEMS Microbiol. Ecol.* 55: 221-229.

Middelboe M. 2000. Bacterial Growth Rate and Marine Virus-Host Dynamics. *Microbial Ecology* 40 (2): 114–24. http://www.ncbi.nlm.nih.gov/pubmed/11029080.

Middelboe M., and Lyck P.G. 2002. Regeneration of Dissolved Organic Matter by Viral Lysis in Marine Microbial Communities. *Aquatic Microbial Ecology* 27 (2). Inter-Research: 187–94. http://cat.inist.fr/?aModele=afficheN&cpsidt=13562408.

Middelboe M., Nielsen T.G., and Bjørnsen P.K. 2002. Viral and Bacterial Production in the North Water: In Situ Measurements, Batch-Culture Experiments and Characterization and Distribution of a Virus–host System. *Deep Sea Research Part II: Topical Studies in Oceanography* 49 (22-23): 5063–79. doi:10.1016/S0967-0645(02)00178-9.

Middelboe M., Riemann L., Steward G.L., Hansen V., Nybroe O. 2003. Virus-induced transfer of organic carbon between marine bacteria in a model community. *Aquat Microbial Ecol* 33: 1-10.

Miki T., Nakazawa T., Yokokawat T., Nagata T. 2008. Functional Consequences of Viral Impacts on Bacterial Communities: A Food-Web Model Analysis. *Freshwater Biology* 53 (6): 1142–53. doi:10.1111/j.1365-2427.2007.01934.x.

Mitsch W.J., Gosselink J.G. 2007. *Wetlands.* John Wiley & Sons, Inc., Hoboken, New Jersey, 582 p.

Moore T.R., Trofymow J.A., Prescott C., CIDET Working Group. 2005. Patterns of decomposition and carbon, nitrogen, and phosphorus dynamics of litter in upland forest and peatlands sites in central Canada. *Can. J. For. Res.* 35: 133-142.

Morales S.E., Moluser P.J., Ward N., Hudman S.P., Gotelli N.J., Ross D.S., Lewis T.A. 2006. Comparison of bacterial communities in New England *Sphagnum* bogs using terminal restriction fragment length polymorphism (T-RFLP). *Microb. Ecol.* 52: 34-44.

Moreira D., and López-García P. 2009. Ten Reasons to Exclude Viruses from the Tree of Life. *Nature Reviews. Microbiology* 7 (4). Nature Publishing Group: 306–11. doi:10.1038/nrmicro2108.

Myers B., Webster K.L., McLaughin J.W., Basiliko N. 2012. Microbial activity across a boreal peatland nutrient gradient: the role of fungi and bacteria. *Wetlands Ecol. Manage.* 20: 77-88.

Noble R.T., and Fuhrman J.A. 1998. Use of SYBR Green I for Rapid Epifluorescence Counts of Marine Viruses and Bacteria. *Aquatic Microbial Ecology* 14 (2). INTER-RESEARCH, NORDBUNTE 23, D-21385 OLDENDORF LUHE, GERMANY: 113–18. doi:10.3354/ame014113.

Opelt K., Berg C., Schönmann S., Eberl L., Berg G. 2007. High Specificity but Contrasting Biodiversity of Sphagnum-Associated Bacterial and Plant Communities in Bog Ecosystems Independent of the Geographical Region. *The ISME Journal* 1 (6): 502–16. doi:10.1038/ismej.2007.58.

Pagarete, A, Chow C.E. T, Johannessen T., Fuhrman J. A., Thingstad T. F., and Sandaa R. A. 2013. Strong Seasonality and Interannual Recurrence in Marine Myovirus Communities. *Applied and Environmental Microbiology* 79 (20): 6253–59. doi:10.1128/AEM.01075-13.

Palesse S., Colombet J., Ram A.S.P, and Sime-Ngando T. 2014. Linking Host Prokaryotic Physiology to Viral Lifestyle Dynamics in a Temperate Freshwater Lake (Lake Pavin, France). *Microbial Ecology*. Laboratoire Microorganismes: Génome et Environnement, UMR CNRS 6023, Clermont Université, Université Blaise Pascal, BP 80026, 63171, Aubière Cedex, France. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=24910014&retmode=ref&cmd=prlinks.

Pankratov T.A., Ivanova A.O., Dedysh S.N., Liesack W. 2011. Bacterial populations and environmental factors controlling cellulose degradation in an acidic *Sphagnum* peat. *Environ. Microbiol.* 13: 1800-1814.

Parada V., Herndl G.J., and Weinbauer M.G. 2006. Viral Burst Size of Heterotrophic Prokaryotes in Aquatic Systems. *Journal of the Marine Biological Association of the UK* 86 (03). Cambridge University Press: 613. doi:10.1017/S002531540601352X.

Patel A., Noble R.T., Steele J.A., Schwalbach M.S., Hewson I., and Fuhrman J.A. 2007. Virus and Prokaryote Enumeration from Planktonic Aquatic Environments by Epifluorescence Microscopy with SYBR Green I. *Nature Protocols* 2 (2). NATURE PUBLISHING GROUP, MACMILLAN BUILDING, 4 CRINAN ST, LONDON N1 9XW, ENGLAND: 269–76. doi:10.1038/nprot.2007.6.

Paterson H., and Laybourn-Parry J. 2012. Antarctic Sea Ice Viral Dynamics over an Annual Cycle. *Polar Biology* 35 (4): 491–97. doi:10.1007/s00300-011-1093-z.

Paul J.H. 2008. Prophages in Marine Bacteria: Dangerous Molecular Time Bombs or the Key to Survival in the Seas? *The ISME Journal* 2 (6): 579–89. doi:10.1038/ismej.2008.35.

Payet J.P., and Suttle C.A. 2008. Physical and Biological Correlates of Virus Dynamics in the Southern Beaufort Sea and Amundsen Gulf. *Journal of Marine Systems*, Sea ice and life in a river-influenced arctic shelf ecosystem, 74 (3–4): 933–45. doi:10.1016/j.jmarsys.2007.11.002.

Payet J.P., and Suttle C.A. 2013. To Kill or Not to Kill: The Balance between Lytic and Lysogenic Viral Infection Is Driven by Trophic Status. *Limnology and Oceanography* 58 (2): 465–74. doi:10.4319/lo.2013.58.2.0465.

Peltoniemi K., Fritze H., Laiho R. 2009. Response of fungal and actinobacterial communities to water-level drawdown in boreal peatland sites. *Soil Biol. Biochem* 41: 1902-1914.

Philippe N., Legendre M., Doutre G., Couté Y., Poirot O., Lescot M., Arslan D., et al. 2013. Pandoraviruses: Amoeba Viruses with Genomes up to 2.5 Mb Reaching that of Parasitic Eukaryotes. *Science (New York, N.Y.)* 341 (6143): 281–86. doi:10.1126/science.1239181.

Piacente F., Marin M., Molinaro A., De Castro C., Seltzer V., Salis A., .Damonte G., et al. 2012. Giant DNA Virus Mimivirus Encodes Pathway for Biosynthesis of Unusual Sugar 4-Amino-4,6-Dideoxy-D-Glucose (Viosamine). *The Journal of Biological Chemistry* 287 (5): 3009–18. doi:10.1074/jbc.M111.314559.

Poisot, T., Lepennetier G., Martinez E., Ramsayer J., and Hochberg M. E. 2011. Resource Availability Affects the Structure of a Natural Bacteria-Bacteriophage Community. *Biology Letters* 7 (2): 201–4. doi:10.1098/rsbl.2010.0774.

Preston M.D., Smemo K.A., McLaughlin J.W., Basiliko N. 2012. Peatland microbial communities and decomposition processes in the James Bay Lowlands, Canada. *Front. Microbiol.* 3:70 doi: 10.3389/fmicb.2012.00070

Putkinen A., Juottonen H., Juutinen S., Tuittila E.-S., Fritze H., Yrjälä K. 2009. Archaeal rRNA diversity and methane production in deep boreal peat. *FEMS Microbiol. Ecol.* 70: 87–98.

Quemin E.R.J., and Quax T.E.F. 2015. Archaeal Viruses at the Cell Envelope: Entry and Egress. *Frontiers in Microbiology* 6 (January): 552. doi:10.3389/fmicb.2015.00552.

Raghoebarsing A.A., Smolders J.P., Schlid M.C., Rijpstra I., Wolters-Arts M., Deerksen J., Jetten M.S., Schouten S., Sinninghe Damsté J.S., Lamers L.P., Roelefs J.G.M., Op den Camp H.J.M., Strous M. 2005. Methanotrophic symbionts provide carbon for photosynthesis in peat bogs. *Nature* 436: 1153–1156.

Rodriguez-Brito B., Li L., Wegley L., Furlan M., Angly F., Breitbart M., Buchanan J., et al. 2010. Viral and Microbial Community Dynamics in Four Aquatic Environments. *The ISME Journal* 4 (6): 739–51. doi:10.1038/ismej.2010.1.

Rodriguez-Valera F., Martin-Cuadrado A.-B., Rodriguez-Brito B., Pasić L., Thingstad T.F., Rohwer F., and Mira A. 2009. Explaining Microbial Population Genomics through Phage Predation. *Nature Reviews. Microbiology* 7 (11). Nature Publishing Group: 828–36. doi:10.1038/nrmicro2235.

Rohwer F., and Thurber R.V. 2009. Viruses Manipulate the Marine Environment. *Nature* 459 (7244): 207–12. doi:10.1038/nature08060.

Roulet N.T., Lafleur P.M., Richard P.J.H., Moore T.R., Humphreys E.R., Bubier J. 2007. Contemporary carbon balance and late Holocene carbon accumulation in a northern peatland. *Glob. Chang. Biol.* 13: 397–411

Roux S., Enault F. 2012. Assessing the Diversity and Specificity of Two Freshwater Viral Communities through Metagenomics. *PloS One* 7 (3): e33641. doi:10.1371/journal.pone.0033641.

Roux S. 2013. *Diversité, évolution et écologie virale : des communautés aux génotypes. Analyse bioinformatique de métagénomes viraux*. Thèse de doctorat en sciences agronomiques. Université Blaise Pascal-Clermont-Ferrand II (France). <NNT : 2013CLF22380>. <tel-00908344>.

Roux S., Enault F., Bronner G., Vaulot D., Forterre P., and Krupovic M. 2013. Chimeric Viruses Blur the Borders between the Major Groups of Eukaryotic Single-Stranded DNA Viruses. *Nature Communications* 4 (January). Nature Publishing Group: 2700. doi:10.1038/ncomms3700.

Roux S., Enault F., Ravet V., Colombet J., Bettarel Y., Auguet J.C., Bouvier T., et al. 2015a. Analysis of Metagenomic Data Reveals Common Features of Halophilic Viral Communities across Continents. *Environmental Microbiology*, October, n/a – n/a. doi:10.1111/1462-2920.13084.

Roux S., Hallam S.J., Woyke T., and Sullivan M.B. 2015b. Viral Dark Matter and Virus-Host Interactions Resolved from Publicly Available Microbial Genomes. *eLife* 4 (July). eLife Sciences Publications Limited: e08490. doi:10.7554/eLife.08490.

Rydin H., Jerglum J. 2006. *The Biology of Peatlands.* Biology of Habitats Series. Oxford University Press, Oxford & New-York.

Sakowski E.G., Munsell E.V., Hyatt M., Kress W., Williamson S.J., Nasko D.J., Polson S.W., and Wommack K.E. 2014. Ribonucleotide Reductases Reveal Novel Viral Diversity and Predict Biological and Ecological Features of Unknown Marine Viruses. *Proceedings of the National Academy of Sciences* 111 (44): 15786–91. doi:10.1073/pnas.1401322111.

Schroeder D.C., Oke J., Hall M., Malin G., and Wilson W.H. 2003. Virus Succession Observed during an Emiliania Huxleyi Bloom. *Applied and Environmental Microbiology* 69 (5): 2484–90. doi:10.1128/AEM.69.5.2484-2490.2003.

Serkebaeva Y.M., Kim Y., Liesack W., and Dedysh S.N. 2013. Pyrosequencing-Based Assessment of the Bacteria Diversity in Surface and Subsurface Peat Layers of a Northern Wetland, with Focus on Poorly Studied Phyla and Candidate Divisions. *PLoS ONE* 8 (5): e63994. doi:10.1371/journal.pone.0063994.

Short C.M., Rusanova O., and Short S.M. 2011. Quantification of Virus Genes Provides Evidence for Seed-Bank Populations of Phycodnaviruses in Lake Ontario, Canada. *Isme Journal* 5 (5): 810–21. doi:10.1038/ismej.2010.183.

Sime-Ngando S. 2014. Environmental Bacteriophages: Viruses of Microbes in Aquatic Ecosystems. *Frontiers in Microbiology* 5 (January). Frontiers: 355. doi:10.3389/fmicb.2014.00355.

Simek, K., Pernthaler J., Weinbauer M. G., Dolan J. R., Nedoma J., Masi M., and Horna K.. 2001. Changes in Bacterial Community Composition and Dynamics and Viral Mortality Rates Associated with Enhanced Flagellate Grazing in a Mesoeutrophic Reservoir. *Applied and Environmental Microbiology* 6: 2723-2733. doi:10.1128/AEM.67.6.2723.

Singh, A. H, Doerks T., Letunic I., Raes J., and Bork. P. 2009. Discovering Functional Novelty in Metagenomes: Examples from Light-Mediated Processes. *Journal of Bacteriology* 191 (1): 32–41. doi:10.1128/JB.01084-08.

Sizova M.V., Panikov N.S., Tourova T.P., Flanagan P.W. 2003. Isolation and characterization of oligotrophic acido-tolerant methanogenic consortia from a *Sphagnum* peat-bog. *FEMS Microbiol. Ecol.* 45: 301-315.

Smirnov N.N. 1961. Food cycles in sphagnous bogs. *Hydrobiologia* 17: 175-181.

Snyder, J. C., Blake W., Matthew L., Roberto F. F., Spuhler J., Ortmann A. C., Douglas T., and Young. M. 2007. Virus Movement Maintains Local Virus Population Diversity. *Proceedings of the National Academy of Sciences of the United States of America* 104 (48): 19102–7. doi:10.1073/pnas.0709445104.

Soudilovskaia N.A., Cornelissen J.H.C., During H.J., van Logtestijn R.S.P., Lang S.I., Aerts R. 2010. Similar cation exchange capacities among bryophyte species refute a presumed mechanism of peatland acidification. *Ecology* 91: 2716-2726.

Sparling G.P., Feltham C.W., Reynalds J., West A.W., Singelton P. 1990. Estimation of soil microbial C by a fumigation-extraction method: use on soils of high organic matter content, and a reassessment of the k_{ec} -factor. *Soil Biol. Biochem.* 22: 301–307.

Stewart J.M., Wheatley R.E., 1990. Estimates of CO_2 production from eroding peat surfaces. *Soil Biol. Biochem.* 22: 65-68.

Steward, G. F., and A. I. Culley. 2010. Extraction and purification of nucleic acids from viruses, p. 154-165. In S. W. Wilhelm, M. G. Weinbauer, and C. A. Suttle [eds.], *Manual of Aquatic Viral Ecology*. ASLO. [DOI 10.4319/mave.2010.978-0-9845591-0-7.154]

Strack M. 2008. *Peatlands and climate change.* International Peat Society, Jyväskyla, 223 p.

Suttle C.A. 2005. Viruses in the Sea. *Nature* 437 (7057): 356–61. doi:10.1038/nature04160.

Swanson M.M., Fraser G., Daniell T.J., Torrance L., Gregory P.J., and Taliansky M. 2009. Viruses in Soils: Morphological Diversity and Abundance in the Rhizosphere. *Annals of Applied Biology* 155 (1): 51–60. doi:10.1111/j.1744-7348.2009.00319.x.

Thingstad T.F., Lignell R. 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. *Aquatic Microbial Ecology* 13: 19-27.

Thormann M.N., Bayley S.E. 1997. Decomposition along a moderate-rich fen-marsh peatland gradient in boreal Alberta, Canada. *Wetlands* 17: 123–137.

Thormann M.N., Currah R.S., Bayley S.E. 2003. Successsion of microfungal assemblages in decomposing peatland plants. *Plant and Soil* 250: 323-333.

Thormann M.N., Bayley S.E. , Currah R.S. 2004. Microcosm tests of the effects of temperature and microbial species number on the decomposition of *Carex aquatilis* and *Sphagnum fuscum* litter from southern boreal peatlands. Can. J. Microbiol. 50:793-802.

Thurber R.V., Haynes M., Breitbart M., Wegley L., and Rohwer F. 2009. Laboratory Procedures to Generate Viral Metagenomes. *Nature Protocols* 4 (4). Nature Publishing Group: 470–83. doi:10.1038/nprot.2009.10.

Trinder C.J., Johnson D., Artz R.R.E. 2008. Interactions among fungal community structure, litter decomposition and depth of water table in a cutover peatland. *FEMS Microbiol. Ecol.* 64:433-448.

Tveit A., Schwacke R., Svenning M.M., and Urich T. 2013. Organic Carbon Transformations in High-Arctic Peat Soils: Key Functions and Microorganisms. *The ISME Journal*, September. doi:10.1038/ismej.2012.99.

Tveit A., Urich T, and Svenning M.M. 2014. Metatranscriptomic Analysis of Arctic Peat Soil Microbiota. *Applied and Environmental Microbiology*. doi:10.1128/AEM.01030-14.

Tveit A., Urich T, Frenzel P., and Svenning M. M. 2015. Metabolic and Trophic Interactions Modulate Methane Production by Arctic Peat Microbiota in Response to Warming. *Proceedings of the National Academy of Sciences of the United States of America* 112 (19): E2507–16. doi:10.1073/pnas.1420797112.

Urich T., Lanzén A., Qi J., Huson D.H., Schleper C., and Schuster S.C. 2008. Simultaneous Assessment of Soil Microbial Community Structure and Function through Analysis of the Meta-Transcriptome. *PloS One* 3 (6). Public Library of Science: e2527. http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0002527.

van Breemen N. 1995. How *Sphagnum* bogs down other plants. TREE 10: 270-275.

Vile M.A., Wieder R.K., Zivkovic T., Scott K.D., Vitt D.H., Hartsock J.A., Iosue C.L., Quinn J.C., Petix M., Fillingim H.M., Popma J.M.A., Dynarski K.A., Jackman T.R., Albright C.M., Wykoff D.D. 2014. N2-fixation by methanotrophs sustains carbon and nitrogen accumulation in pristine peatlands. *Biogeochemistry* 121: 317-328.

von Post L. 1924. Das genestische System der organogenen Bildungen Schwedens. *Proc.Int. Soc. Soil Sci. Congr., Helsingfors,* pp. 286-304.

Vos M., Birkett P.J., Birch E., Griffiths R.I., and Buckling A. 2009. Local Adaptation of Bacteriophages to Their Bacterial Hosts in Soil. *Science (New York, N.Y.)* 325 (5942): 833. doi:10.1126/science.1174173.

Waddington M., Rottenberg P.A., Warren F.J. 2001. Peat CO_2 production in a natural and cutover peatland: implications for restoration. *Biogeochemistry* 54: 115-130.

Waksman S.A., Stevens K.R. 1929. Contribution to the chemical composition of peat V.The role of microorganisms in peat formation and decomposition. *Soil Sci.* 28,: 315-340.

Ward S.E., Ostle N.J., Oaldey S., Quirk H., Henrys P.A., Bardgett R.D. 2013. Warming effects on greenhouse gas fluxes in peatlands are modulated by vegetation composition. Ecol. Letters

Weinbauer M.G., and Peduzzi P. 1994. Frequency, Size and Distribution of Bacteriophages in Different Marine Bacterial Morphotypes. *Marine Ecology Progress Series* 108 (1-2). INTER-RESEARCH, NORDBUNTE 23, D-21385 OLDENDORF LUHE, GERMANY: 11–20. doi:10.3354/meps108011.

Weinbauer M.G., and Suttle C.A. 1996. Potential Significance of Lysogeny to Bacteriophage Production and Bacterial Mortality in Coastal Waters of the Gulf of Mexico. *APPLIED AND ENVIRONMENTAL MICROBIOLOGY* 62 (12). AMER SOC MICROBIOLOGY, 1325 MASSACHUSETTS AVENUE, NW, WASHINGTON, DC 20005-4171: 4374–80. http://apps.webofknowledge.com.gate1.inist.fr/full_record.do?product=UA&search_mode=GeneralSearch&qid=10&SID=S2NmBoTqAAfWRtxNexx&page=1&doc=1.

Weinbauer M.G., Brettar I., Höfle M.G. 2003. Lysogeny and Virus-Induced Mortality of Bacterioplankton in Surface, Deep, and Anoxic Marine Waters. *Limnology and Oceanography* 48 (4): 1457–65. http://doi.wiley.com/10.4319/lo.2003.48.4.1457.

Weinbauer M.G. 2004. Ecology of Prokaryotic Viruses. *FEMS Microbiology Reviews* 28 (2): 127–81. doi:10.1016/j.femsre.2003.08.001.

Weinbauer M.G., Bonilla-Findji O., Chan A.M., Dolan J.R., Short S.M., Simek K., Wilhelm S.W., and Suttle C.A. 2011. Synechococcus Growth in the Ocean May Depend on the Lysis of Heterotrophic Bacteria. *Journal of Plankton Research* 33 (10): 1465–76. doi:10.1093/plankt/fbr041.

Wieder R.K., Lang G.E. 1982. A critique of the analytical methods used in examiningdecomposition data obtained from litter bags. *Ecology* 63: 1636-1642.

Wilhelm S.W., Brigden S.M., and Suttle C.A. 2002. A Dilution Technique for the Direct Measurement of Viral Production: A Comparison in Stratified and Tidally Mixed Coastal Waters. *Microbial Ecology* 43 (1): 168–73. doi:10.1007/s00248-001-1021-9.

Williams B.L., Sparling G.H. 1984. Extractable nitrogen and phosphorus in relation to microbial biomass in acid organic soils. *Plant and Soil* 76: 139–148.

Williamson K.E., Radosevich M., and Wommack K.E. 2005. Abundance and Diversity of Viruses in Six Delaware Soils. *Applied and Environmental Microbiology* 71 (6): 3119–25. doi:10.1128/AEM.71.6.3119-3125.2005.

Winter C., Bouvier T., Weinbauer M.G., and Thingstad T.F. 2010. Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: The 'Killing the Winner' Hypothesis Revisited. *Microbiology and Molecular Biology Reviews : MMBR* 74 (1): 42–57. doi:10.1128/MMBR.00034-09.

Winter C., Matthews B., and Suttle C.A. 2013. Effects of Environmental Variation and Spatial Distance on Bacteria, Archaea and Viruses in Sub-Polar and Arctic Waters. *The ISME Journal* 7 (8). Nature Publishing Group: 1507–18. doi:10.1038/ismej.2013.56.

Woese C.R. 1987. Bacterial Evolution. *Microbiological Reviews* 51 (2): 221–71. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=373105&tool=pmcentrez&rendertype=abstract.

Yavitt J.B., Lang G.E., Wieder R.K. 1987. Control of carbon mineralization to CH_4 and CO_2 in anaerobic *Sphagnum*-derived peat from Big Run Bog, West Virginia. *Biogeochemistry* 4: 141-157.

Zablocki O., van Zyl L., Adriaenssens E.M., Rubagotti E., Tuffin M., Cary S.C., and Cowan D. 2014. High-Level Diversity of Tailed Phages, Eukaryote-Associated Viruses, and Virophage-like Elements in the Metaviromes of Antarctic Soils. *Applied and Environmental Microbiology* 80 (22): 6888–97. doi:10.1128/AEM.01525-14.

Zhong X., Ram A.S.P., Colombet J., and Jacquet S. 2014. Variations in Abundance, Genome Size, Morphology, and Functional Role of the Virioplankton in Lakes Annecy and Bourget over a 1-Year Period. *Microbial Ecology* 67 (1): 66–82. doi:10.1007/s00248-013-0320-2.

Chapitre 1 :

A. Dynamics of viral abundance and diversity in a *Sphagnum*-dominated peatland: temporal fluctuations prevail over habitat

(en revision pour Frontiers in Microbiology)

Flore Ballaud[1], Alexis Dufresne[1], André-Jean Francez[1], Jonathan Colombet[2], Télesphore Sime-Ngando[2] and Achim Quaiser[1]

[1]Université de Rennes 1, UMR CNRS 6553 ECOBIO, Campus de Beaulieu, 35042 Rennes, France

CNRS, UMR 6553 ECOBIO, Campus de Beaulieu, 35042 Rennes, France

[2] Université Clermont Auvergne, Université Blaise Pascal, BP 10448, F-63000 Clermont-Ferrand, France

CNRS, UMR 6023, Laboratoire Microorganismes: Génome et Environnement (LMGE), 24 avenue des Landais, F-63171 Aubière, France

B. Diversity and comparative genomics of *Microviridae* in *Sphagnum*-dominated peatlands

Frontiers in Microbiology 2015 volume 6 , article 375, doi: 10.3389/fmicb.2015.00375

Achim Quaiser[1]*, Alexis Dufresne[1], Flore Ballaud[1], Simon Roux[2], Yvan Zivanovic[3], Jonathan Colombet[4], Télesphore Sime-Ngando[4] and André-Jean Francez[1]

[1]Université de Rennes 1, CNRS UMR6553 EcoBio, Rennes, France

[2]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA

[3]Université Paris Sud, CNRS UMR8621, 91405 Orsay, France

[4]Université Blaise Pascal, UMR CNRS 6023 LMGE, Clermont-Ferrand, France

Chapitre 1: Dynamics and diversity of the viral communities in a *Sphagnum*-dominated peatland.

Résumé du Chapitre 1

Les virus affectent l'activité microbienne et le recyclage du carbone dans de nombreux écosystèmes, mais malgré l'importance des tourbières à Sphaignes dans le stockage du carbone du sol, et l'importance de l'activité microbienne dans cette fonction, la diversité et le rôle des virus dans ces écosystèmes n'avaient à notre connaissance jamais été étudiés. L'abondance des particules virales et des procaryotes a été suivie de manière bimestrielle à deux profondeurs de l'acrotelme et du mésotelme du fen et du bog d'une tourbière à *Sphagnum*. Au cours de ce suivi, la variation de l'abondance des particules virale était comprise entre $1.7x10^6$ et $5.6x10^8$ particules mL^{-1}. Si l'abondance des particules virales présente d'importances variations temporelles, elle ne diffère apparemment pas entre le fen et le bog. Les fluctuations temporelles apparaissent positivement corrélées aux fluctuations de l'abondance des procaryotes et de la concentration en carbone organique dissout, et négativement corrélées à celles du niveau de la nappe d'eau et de l'oxygène dissout. L'analyse par séquençage aléatoire des pools d'ADN viraux (métaviromes) montre un bouleversement de la diversité virale entre deux grandes périodes : avec d'une part l'hiver et le printemps et d'autre part l'été et l'automne, suggérant, en lien avec l'abondance des particules virales, une succession temporelle des communautés virales associée aux changements saisonniers des facteurs environnementaux (précipitations, température…). Des séquences virales ont été retrouvées dans douze pools d'ADN qui représentent la diversité microbienne de la tourbe au printemps (métagénomes). Ces séquences n'étant pas issues des métaviromes de printemps, mais de ceux d'été et d'automne, le décalage temporel entre les métagénomes et les métaviromes pourrait s'expliquer par transition de stratégie d'infection virale, passant en été de la lysogénie à l'infection lytique en réponse à un changement d'activité de la communauté d'hôtes. Nos résultats montrent que les variations temporelles des conditions environnementales liées aux facteurs météorologiques structurent plus fortement l'abondance et la diversité virale que les différences environnementales associées aux deux grands types d'habitats rencontrés dans les tourbières à Sphaignes. Par ailleurs, si une faible proportion des séquences de ces 12 métaviromes a pu être identifiée taxonomiquement ou fonctionnellement par comparaison avec les génomes viraux référencés, ces ressemblances comprenaient des caractéristiques intéressantes, notamment le fait que la majeure partie des identifications correspondent à des protéines de capside et de réplication. Les séquences virales ont été assemblées en contigs, séparément pour chacun des 12 métaviromes. Ce travail bioinformatique

d'assemblage a permis de produire de nombreux contigs de plus de 3 000 nucléotides de longueur dont l'analyse et la comparaison ont mené à la description et à l'annotation de nouveaux génomes viraux associés à la famille des *Microviridae*.

Les *Microviridae* sont des virus à ADN simples brin (ssDNA viruses) qui infectent les bactéries. Cette famille de bactériophages reste assez mal caractérisée, malgré le fait qu'elle comprenne le phage PhiX174, qui est l'un des principaux modèles en virologie pour l'étude des génomes viraux et des structures de capsides virales. De récentes études ont montré que ces *Microviridae* sont une famille diversifiée et qu'ils constituent une part non négligeable des communautés virales tant des écosystèmes marins et d'eau douce que des microbiomes humains. Cependant, malgré leur importance, nous n'avions aucune connaissance de leur diversité, de leur abondance et de leur rôle écologique dans les écosystèmes de sols. L'analyse phylogénétique de la grande protéine de capside, dont la sequence est bien conservée, révèle que ces nouveaux génomes sont en partie affiliés aux sous-familles des *Gokushovirinae* et des *Pichovirinae*, et permet la distinction et la définition de deux nouvelles sous-familles, les *Aravirinae* et les *Stokavirinae*. La modélisation de la structure de la grande protéine de capside des *Aravirinae* montre que celle-ci présente des similarités avec celles des *Alpavirinae* et des *Pichovirinae*. Elle révèle également l'existence de deux autres régions très variables au sein de cette structure protéique, ces régions étant potentiellement impliquées dans la reconnaissance phage/hôte. De plus, deux nouveaux prophages représentant potentiellement une nouvelle sous-famille de *Microviridae* ont été identifiés dans les génomes de *Parabacteroidetes merdae* et *Parabacteroidetes distasonis*.La distinction phylogénétique des sous-familles a ensuite été confirmée par l'analyse de l'organisation des gènes dans les nouveaux génomes de *Microviridae* et renforcée par des analyses de similarité. L'analyse de l'abondance relative de la grande protéine de capside (VP1) révèle que les *Gokushovirinae* et les *Aravirinae* représentent les plus abondantes sous-familles de *Microviridae* dans 11 des 12 métaviromes de tourbière, les séquences recrutées contre la protéine VP1 des *Gokushovirinae* et des *Aravirinae* constituant respectivement jusqu'à 4.19% et 0.65% du nombre total de séquences de certains métaviromes. Ainsi, cette étude complète la banque actuellement disponible de génomes de *Microviridae* ce qui permet d'envisager l'estimation quantitative des sous-familles de *Microviridae*.

Chapitre 1 A. Dynamics of viral abundance and diversity in a Sphagnum-dominated peatland: temporal fluctuations prevail over habitat

1. Introduction

Viruses are present in virtually all ecosystems, and are considered to be the most abundant biological entities in the biosphere (Suttle, 2005). Through lytic and lysogenic life cycles, viruses affect the metabolism and abundance of their cellular hosts from all three domains of life (synthesis in Kirchman, 2012), impacting the diversity and the structure of microbial communities (Middelboe, 2000; Suttle, 2007; Sime-Ngando, 2014). Because viruses function as a top-down control on microbial production (Pradeep Ram et al., 2011; Chow et al., 2014), they affect biogeochemical cycles through the release of substantial amounts of organic carbon and nutrients in the environment (Middelboe and Lyck, 2002). While resources released by this viral shunt can be reused to sustain microbial biomass, it typically results in a reduction of the overall flow of organic matter and energy towards higher trophic levels (Fuhrman, 1999; Middelboe and Lyck, 2002; Ankrah et al., 2014). Besides the viral-shunt, a significant proportion of the released viruses represent a consequential amount of labile organic matter that can be readily decomposed (Dell'Anno et al., 2015). Viral lysis, along with bacterial grazing are important processes in microbial succession (Fierer et al., 2010), suggesting that coupled viral-host interactions can influence ecosystem-level carbon cycling, depending on the activity of cells and on the balance between lytic and lysogenic strategies. Consequently, considering viral ecology is critical to understanding ecosystem functioning.

Peatlands are a globally relevant component of the carbon cycle, storing a quarter of global soil carbon and more carbon than all vegetation (Yu, 2012; Turetsky et al., 2015). *Sphagnum*-dominated peatlands often develop from a nutrient-rich, non-acidic fen (minerotrophic fen) during early stages, to a nutrient-poor acidic bog (rain-fed ombrotrophic bog)(Rydin and Jeglum, 2006). During this development process, peatlands accumulate large stocks of partly decomposed organic matter as peat. This accumulation is a consequence of the long-term imbalance between carbon uptake from photosynthesis and carbon losses *via* respiration, methanogenesis and DOC leaching, which results in the preservation of up to 15 % of the net primary production as peat (Clymo, 1984; Moore, 1987; Francez and Vasander, 1995; Roulet et al., 2007). The combination of abiotic conditions, such as low temperature, low pH, high soil water content and poor nutrient availability, as well as biotic factors (*Sphagnum* litter quality, allelopathy) constrains microbial activities, limiting

decomposition and mineralization (Clymo, 1984; Rydin and Jeglum, 2006). Accumulation of organic matter over long time periods leads to stratification with a deep permanently anoxic catotelm in contact with underlying bedrock, an intermediate mesotelm layer that is periodically anoxic and oxic and the predominantly oxic acrotelm at the surface (Clymo and Bryant, 2008). The heterogeneity of peat, the variety of dynamic stages and sources of water and nutrients provide a large panel of habitats and niches for a broad diversity of microorganisms transforming different carbon substrates through aerobic and anaerobic pathways of decomposition (Juottonen et al., 2005; Artz, 2009; Andersen et al., 2010; Tveit et al., 2012).

Understanding of microbial diversity in peatlands, especially in regards to spatio-temporal patterns, is currently limited. The development of cultivation-independent approaches such as metagenomics provide a powerful tool to investigate microbial taxa and associated protein-coding genes across ecosystems (Lynch et al., 2004; Vandenkoornhuyse et al., 2010). More recently, sequencing of viral metagenomes (hereafter metaviromes) has dramatically expanded understanding of viral diversity, particularly in marine and freshwater environments (Comeau et al., 2008; Rohwer and Thurber, 2009; Lopez-Bueno et al., 2009; Roux et al., 2012; Smith et al., 2013; Labonté and Suttle, 2013; Wommack et al., 2015) compared to soil ecosystems (Kimura et al., 2008; Zablocki et al., 2014; Adriaenssens et al., 2015). Most analyses have focused on cross-biomes comparisons, while spatio-temporal dynamics remain poorly documented (Pagarete et al., 2013; Chow et al., 2014). Several viral marker genes, conserved within particular viral families, are considered to be reliable for taxonomic affiliation (Roux et al., 2011; Sakowski et al., 2014). Nevertheless, none of these marker genes are ubiquitous among viruses and they are far from the high level of conservation found in rRNA genes that are used for the classification of organisms from the three kingdoms. While several studies have investigated the diversity of microbial communities in peatlands (Dedysh et al., 2006; Artz et al., 2007; Peltoniemi et al., 2009; Dedysh, 2011; Mackelprang et al., 2011; Bragina et al., 2012; Serkebaeva et al., 2013; Mondav et al., 2014; Tveit et al., 2014; Hultman et al., 2015) viral diversity in *Sphagnum*-dominated peatlands remains largely unknown and basic knowledge of virus ecology in these ecosystems is still lacking (Quaiser et al., 2015). In view of the importance of viruses in structuring and regulating prokaryotic communities and the implication of the latter in the carbon sink function of peatland; it is essential to understand the role of viruses in the dynamics of the microbial communities in this ecosystem.

In order to characterize virus ecology in a temperate *Sphagnum*-dominated peatland, we combined and integrated different approaches to study the spatio-temporal patterns of viral abundance and diversity. The goals of this work were: 1) to compare the seasonal abundance of viruses and prokaryotes at two different depths corresponding to a stratified analysis of acrotelm,

the most active layer, of fen and bog; 2) to identify abiotic controls on abundance and distribution of these viruses and prokaryotes; and 3) to characterize the viral diversity over an annual cycle.

2. Materials and methods

2.1. Site description and experimental design

Pore-water samples from peat were collected at Les Pradeaux (3°55E; 45°32N), a *Sphagnum*-dominated peatland situated in the French Massif Central at an altitude of 1 350 m. The fen is dominated by *Sphagnum fallax*, *Carex rostrata, Eriophorum angustifolium* and *Menyanthes trifoliata*, while the bog is mainly colonized by *Sphagnum magellanicum*, *Sphagnum capillifolium, Andromeda polifolia*, *Vaccinium oxycoccos* and *Eriophorum vaginatum*. Complementary approaches of viral ecology and corresponding sampling strategies were applied (Annexe1: Supplementary Table S1):

a.) The seasonal fluctuations of viral-particle and prokaryote abundance were monitored in the pore water from two different depths sampled in fen and bog in triplicate at eight different dates: in 2010 (May), 2011 (June, August (2 dates), and October) and 2012 (June, July and September). This represents a total of 96 samples (2 depths x 2 peats - Fen + Bog - x 8 dates x 3 triplicates). Pore-water, pressed water from peat samples, is widely used to analyze microbial communities in *Sphagnum*-peatlands, and it is expected to recover the large majority of the microorganisms (Gilbert et al., 1998a).

b.) Viral-particles and prokaryotes abundances associated with physico-chemical features were studied from triplicate pore water samples extracted in fen and bog (one depth) in 2012 (March, May, June, July, August, September and November), representing 42 samples (fen + bog x 7 dates x 3 triplicates). The associated viral-particle and prokaryote abundances were analyzed when available in triplicates, but omitted when counts could not be obtained for the different depths due to environmental conditions as too high water logging or not enough humidity to separate the two layers.

c.) Viral diversity using shotgun metagenomics was analyzed in pore water from fen and bog (one depth; 0-15cm) in June 2011, August 2011, October 2011 (one sample per dynamic stage) and March 2012 (triplicate samples for each dynamic stage), which represent a collection of 12 metaviromes.

d.) As viruses are intimately linked to prokaryote communities, we sampled peat (matrix + water) for the metagenome analysis of the microbial diversity. Samples were collected in triplicate for two depths of the two dynamic stages, fen and bog, during the sampling campaign of June 2011 (12 metagenomes).

2.2 Physico-chemical parameters

Peat temperature was measured directly in the peat profiles at 5, 10 and 15 cm under the *Sphagnum* capitula layer. Conductivity, pH and oxygen were measured in the field with filtered (125 µm) peat water. Dissolved organic carbon (DOC) and anion concentrations (nitrate, sulfate and acetate) were measured at the "Biogeochemical Analysis" platform (ECOBIO and GEOSCIENCES - OSU Rennes), following water filtration at 0.22 µm (Whatman), using a Bioritech DOC Analyser and a Dionex Analyzer (Table 1).

Table 1 - Physico-chemical parameters measured in the peat samples from 2012.

		Temperature (C°)	Water-table (cm)	pH	Conductivity (µS cm^{-1})	Sulfate ($mg.L^{-1}$)	Oxygen ($mg.L^{-1}$)	DOC (ppm)
FEN	March	14.5 (1.3)	1 (0)	4.7 (0.2)	17.6 (4.2)	0.70 (0.14)	9.54 (0.55)	15.6 (3.1)
	May	14.3 (0.6)	0.8 (1.8)	4.6 (0.1)	16.1 (1.1)	0.67 (0.07)	6.33 (0.02)	19.7 (2.2)
	June	14 (0.5)	-0.2 (0.6)	6.1 (0.1)	16.8 (4.8)	0.69 (0.06)	6.57 (0.90)	25.3 (3.3)
	July	13.8 (1.6)	-5.6 (1.2)	6.6 (0.2)	12.8 (0.7)	1.08 (0.07)	9.61 (0.84)	18.4 (1.2)
	August	15 (1.0)	-10 (1)	4.6 (0.1)	97.7 (4.8)	1.85 (0.38)	0.44 (0.02)	NA
	September	10.1 (1.4)	-7.3 (0.6)	4.6 (0.1)	34.6 (8.2)	1.64 (0.10)	1.28 (1.11)	38.4 (2.5)
	November	5 (4.4)	1 (0)	4.6 (0.1)	56.9 (32.8)	1.67 (0.09)	9.73 (0.35)	NA
BOG	March	11 (3.0)	-12.3 (2.5)	4.7 (0.1)	44.2 (23.7)	1.24 (0.35)	7.28 (0.20)	19.5 (0.7)
	May	13.7 (0.8)	-10.3 (1.5)	4.4 (0.2)	42.5 (8.1)	2.20 (0.41)	5.69 (0.56)	27.9 (5.3)
	June	12.8 (0.8)	-16.2 (1.6)	4.4 (0.3)	45.3 (13.2)	0.94 (0.19)	5.83 (0.49)	49.7 (9.1)
	July	13.8 (0.3)	-20.3 (2.1)	4.5 (0.1)	29.5 (3.4)	1.46 (0.09)	6.12 (0.34)	17.8 (1.3)
	August	16.8 (1.4)	-23.7 (2.1)	4.2 (0.1)	61.8 (2.6)	2.28 (0.17)	1.77 (1.68)	NA
	September	10.7 (0.6)	-21.3 (2.1)	4.3 (0.1)	62.8 (25.3)	3.01 (0.60)	0.30 (0.07)	47.9 (12.3)
	November	2.7 (0.6)	-22 (2.0)	4.1 (0.1)	47.5 (18.3)	2.13 (0.44)	10.96 (0.58)	NA

Mean (± standard deviation) (n=3), Abbreviations: DOC: Dissolved Organic Carbon; NA: not available.

2.3 Prokaryote and viral-like particle abundances

At each sampling point, a 20 cm peat core was extracted, and the sections from 5 to 10 cm (called "surface") and from 10 to 15 cm (called "depth") below the *Sphagnum* capitula were harvested. For each layer of the peat column, at least 20 mL of peat water was extracted, prefiltered at 125 µm, and fixed in Glutaraldehyde (2%, Grade II, Sigma-Aldrich). Prokaryote abundance (PA) and viral particle abundance (VPA) were obtained using flow cytometry (BD Accuri C6) (Sime-Ngando et al., 2008).

2.4 Statistical analysis

Due to non-homogeneity of variance, one-factor Kruskal-Wallis tests were used on PA, VPA and VPR in order to detect differences between sites (fen *versus* bog), depths (surface *versus* depth) and sampling dates. A principal component analysis (PCA) was performed on the physico-chemical dataset taking into account samples with available DOC (ade4 package)(Dray and Dufour, 2007). The first component was associated with the fluctuation of the physico-chemical variables through the habitats and seasons. We used the sample coordinates on the first component as a variable representing the spatio-temporal gradient (Legendre and Legendre, 1998; Ramette, 2007). Then potential relations between the gradient and log transformed PA, VPA and VPR were tested using

linear regression. All statistical analyses were performed using the open-source statistical software R (version 2.14) (R Development Core Team, 2013).

2.5 Metavirome and metagenome production

Fen and bog viral communities were sampled at 4 dates (Annexe 1: Table S1, Table 2): 07 June 2011 (vFen_June11, vBog_June11; along with peat for a study of the microbial communities), 12 August 2011 (vFen_Aug11, vBog_Aug11), 12 October 2011 (vFen_Oct11, vBog_Oct11) and 23 March 2012 (vFen_Mar12, vBog_Mar12). Due to low water content of the peat, only one sample was collected for each dynamic stage for the three first dates. In March 2012, the fen and bog were sampled in biological triplicates (vFen_Mar12_A, B, C; vBog_Mar12_A, B, C). In total we prepared six fen and six bog metaviromes. *Sphagnum* water was extracted and prefiltered at 125 µm. Viruses were concentrated using PEGylation (Colombet et al., 2007). Viral concentrates were filtered through a 0.22 µm filter (Minisart, Sartorius) and diluted 5x in H_2O (Sigma) to a volume of 5 mL. Remaining non-viral DNA was digested with 10 U DNAse RQ1 RNAse free (Promega) at 37 °C for 1h. Viral DNA was extracted as described previously (Quaiser et al., 2015). DNA quality was checked with the High Sensitivity DNA kit on a 2100 Bioanalyzer (Agilent). Whole genome amplification (WGA) was applied in triplicate for each sample using GenomiPhi V2 (GE Healthcare) following manufacturers' instructions and the triplicates were pooled.

Subsequent pyrosequencing was performed on a GS FLX system (454 Life Sciences, Roche, Branford, CT) at the "Functional and Environmental Genomics" platform (OSU, Rennes, France). Roche/454 filtering tools and stringent filters were used to ensure the highest sequence quality and to remove artificial replicates of sequences and sequences smaller than 250 bp as shown previously (Quaiser et al., 2014).

For the metagenomes, twelve peat samples, distinguishing bog/fen, surface/depth, each in triplicates, collected on 07 June 2011, underwent DNA extraction, pyrosequencing as well as size and quality trimming (Annexe 1: Table S2) as described previously (Quaiser et al., 2014). Briefly, 2g of peat matrix were lysed in 15ml of extraction buffer containing 4% cetyltrimethylam- monium bromide (CTAB), 0.5% polyvinylpyrrolidone (PVP, Sigma-Aldrich), 0.7 M NaCl, 100 mM potassium phosphate (pH 6.8), 20 mM EDTA (pH 8.0), 1% beta-mercaptoethanol, 1 M guanidin thiocyanate and incubated at 65°C for 30 min. Homogenization was obtained by vigorous vortexing every 5 min during 1min in the presence of glass beads. One volume of chloroform-isoamylalcohol (24:1) was added, vortexed for 1 min and incubated at room temperature for 5 min. The samples were centrifuged at 4000 g for 15 min at 4°C and the aqueous phases were transferred to new tubes. Binding conditions for silica-based DNA extraction were adjusted, applied on Nucleo Spin RNA II kit

columns and subsequent purification was performed following the instructions of the manufacturer (Macherey-Nagel). DNA was nebulized to fragments of about 700bp. The DNA was purified with Agencourt AMPur XP magnetic beads (Beckman-Coulter). DNA fragmentation quality was checked with the High sensitivity DNA kit on a 2100 Bioanalyzer (Agilent). Subsequent library construction and pyrosequencing was performed in technical duplicates on a GS FLX system (454/Roche) at the ''Functional and Environmental Genomics'' platform (OSU, Rennes, France). Roche/454 filtering tools and stringent filters developed locally were used to ensure the highest possible sequence quality and to suppress artificial replicates of sequences as well as sequences smaller than 250 bp.

The metaviromes are available under the Metavir IDs (http://metavir-meb.univ-bpclermont.fr/): 1368, 1369, 1370, 1373, 1374, 1375, 1376, 1377, 1378, 1379, 1380 and 1382. In addition, pyrosequencing reads reported in this publication have been deposited in the ENA Sequence Read Archive under the study accession number PRJEB11420 (metaviromes) and PRJEB11421 (metagenomes).

2.6 Metavirome analysis

After pyrosequencing, sequence quality and size trimming, we obtained 481 402 and 615 487 sequences with an average length of 417 bp from fen and bog, respectively (Table 2). The quality of the virome extraction process was assessed by determining the amount of rRNA and tRNA sequences using Meta_RNA that identifies SSU and LSU rRNAs from the three kingdoms (Huang et al., 2009) and tRNAscan-SE (Lowe and Eddy, 1997). In total 69 rRNA (0.0063%) and 1681 tRNA (0.15%) sequences were identified indicating an insignificant level of potential contamination of microbial DNA. To avoid misinterpretation of the results, these sequences were excluded in subsequent analysis.

The viral diversity was analyzed using Metavir (Roux et al., 2014) and the sequences were subjected to tBLASTx (Altschul et al., 1997) against the NCBI RefSeqVirus database (e-value 10^{-7}). Taxonomic assignment of the sequences was determined with MEGAN (Huson et al., 2007). Several accompanying tools were used on the Galaxy/Genouest bioinformatics platform (Le Bras et al., 2013).

To estimate the level of similarity between the viral communities, the proportion of similar sequences of each pair of metaviromes was computed with Compareads 1.2.2 (Maillet et al., 2012). Using this software, two sequences are considered to be similar if they present a defined number (t) of identical *k*-mers (k). To calibrate this analysis we tested 3 different numbers (t=2; 4; 10) of identical 33mers (k=33 nt). The most reliable results were obtained using 4 identical 33mers, parameters that were used for further analyses computed with Compareads. Compareads output is a percentage of similarity between a pair of metaviromes. These percentages were used to build a

distance matrix, on which hierarchical clustering was performed using the R package pvclust package (Suzuki and Shimodaira, 2006) (distance="correlation", method="average").

Comparisons made with Compareads give a global estimate of the similarity between the metaviromes. In order to take into account the diversity of sequences within each metavirome, we analyzed the qualitative distribution (presence-absence) of clusters of highly similar sequences in the 12 metaviromes. This second analysis also allowed removal of potential bias due to the variation of the number of sequences obtained for each metavirome. Sequences from the 12 metaviromes were clustered using CD-HIT-EST (Huang et al., 2010) (c=0.95; n=8). Clustering results were used to compute Sørensen dissimilarity between pairs of metaviromes using MOTHUR (Schloss et al., 2009). Hierarchical clustering (pv-clust package) was used to represent compositional relatedness between metaviromes from the matrix of Sørensen dissimilarities. Clusters of sequences were split into different categories according to the amount of sequences they contained and the same analysis was performed for each size category.

2.7 Analysis of the sequences shared by a metagenome and a metavirome

In order to find viral sequences in the metagenomes, and to analyze the link between the viral and microbial communities, sequences shared by at least one metagenome and one metavirome were retrieved using Compareads (k=33, t=4), and clustered using CD-HIT-EST (see "metaviromes analysis"). To ensure that this selection of sequences did not alter the compositional patterns observed for the total metaviromes, Sørensen dissimilarity was calculated with these clusters for the "shared" sequences originating from the metaviromes. Correspondence Analysis (CA) was performed on the whole dataset of shared sequences and the sample dissimilarities carried on the two first axes was represented with a hierarchical clustering (ade4 package) (Dray and Dufour, 2007). Taxonomic assignment of the sequences was obtained using Metavir tBLASTx output (evalue 10^{-7}) (Roux et al., 2014), and analyzed with MEGAN (Huson et al., 2007).

3. Results

3.1. Viral and prokaryotic abundance

Viral particle abundance (VPA) and prokaryotic abundance (PA) were investigated for the two peatland development stages over two years (Figure 1) aiming for the detection of spatial trends in the abundance of biological entities. VPA ranged from 1.7 ± 0.9 x 10^6 (fen surface, July 2012) to 5.6 ± 2.1 x 10^8 particles mL^{-1} (bog depth, September 2012), and PA ranged from 2.8 ± 1.2 x 10^6 (fen surface, July 2012) to 6.3 ± 1.3 x 10^8 cells mL^{-1} (fen depth, May 2010). VPA and PA were significantly correlated (Spearman, r= 0.76; $P<10^{-15}$; N=95). We did not observe significant differences in PA and

VPA fen and bog; however, we detected significant variations with time (PA: KW-test, P=1.8 x 10^{-5}, N=95; VPA: KW-test, P=1.1 x 10^{-8}, N=95). While PA was significantly higher at depth (average abundance of 9.5 ± 7.1 x 10^{7} cells mL^{-1} at the surface and 1.9 ± 1.6 x 10^{8} cells mL^{-1} at depth) (KW-test, P<0.05, N=95), VPA did not differ significantly with depth (KW-test, P=0.056, N=95). Nevertheless, VPA followed a similar trend, with higher average abundances at depth compared with the surface (1.8 ± 1.6 x 10^{8} particles mL^{-1} at depth and 1.2 ± 1.1 x 10^{8} particles mL^{-1} at the surface). The virus to prokaryote ratio (VPR) differed by sampling date (Figure 2) (KW-test, P<0.01, N=95). No significant differences were observed between development stages or sampling depths. The highest VPR, measured in June 2011 was due to low PA rather than to high VPA.

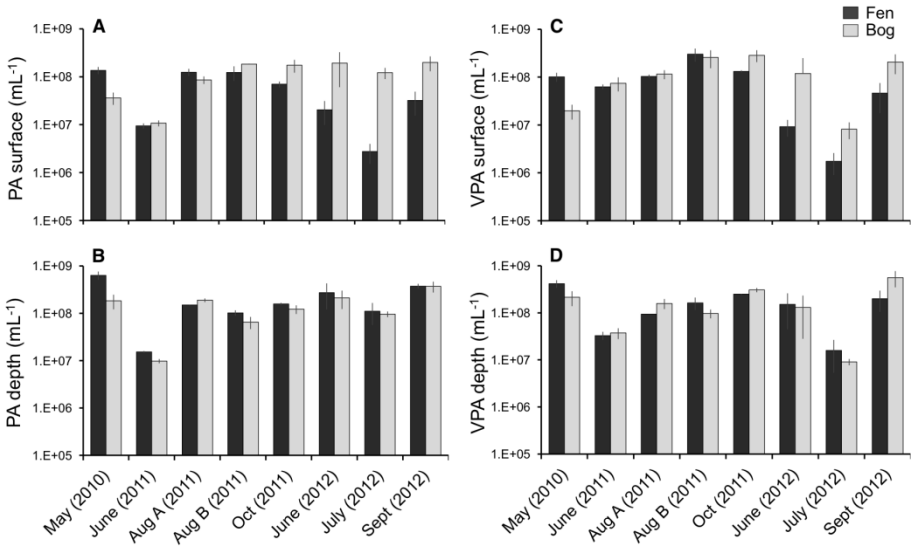


Figure 1 - Prokaryote abundance (PA) and viral particle abundance (VPA) from the fen and bog of the *Sphagnum*-dominated peatland. (A) PA at the surface, (B) PA at depth, (C) VPA at the surface, (D) VPA at depth. Bars represent standard deviations (N=3).

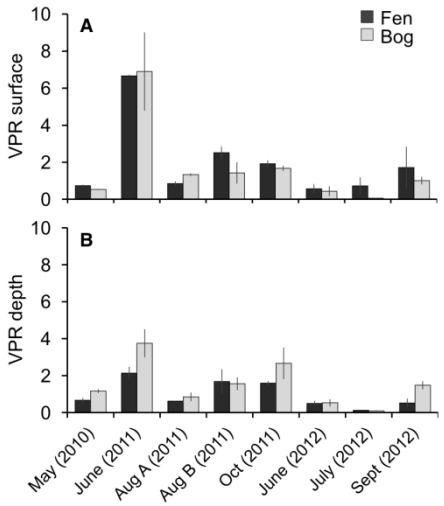


Figure 2 - Comparison of the virus to prokaryote ratio (VPR) at the surface (A) and at depth (B) of the *Sphagnum*–dominated peatland. Bars represent standard deviations (N=3).

3.2. Link between abiotic variables, VPA and PA

The potential relationships between abiotic and biotic environment and virus communities were analyzed combining fluctuation of physico-chemical variables and of viral and prokaryote abundance. The annual mean water-table level was -2.9 ± 4.5 cm in the fen and -18.0 ± 5.0 cm in the bog. Temperature varied more at the surface of the bog (annual mean: 11.6 ± 4.4 °C) than at the surface of the fen (12.4 ± 3.8 °C), but was lower at depth for both stages (fen: 9.3 ± 3.5 °C; bog: 9.2 ± 3.5 °C) (Table 1).

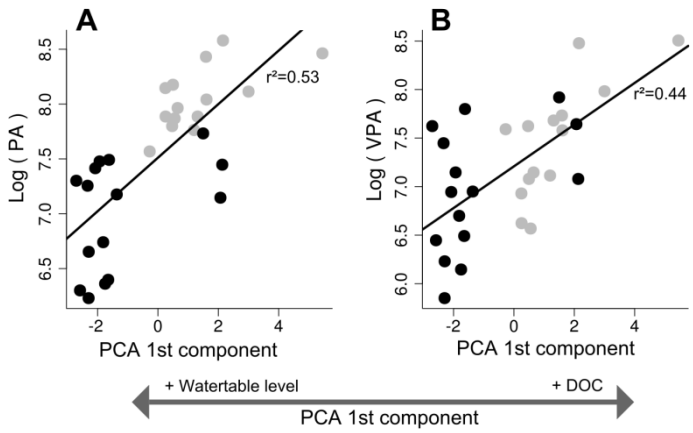


Figure 3 - Linear regressions between log-transformed PA (A), VPA (B) and the abiotic gradient. The abiotic gradient corresponds to the data points' position on the 1st component of the PCA. Black dots: fen samples. Grey dots: bog samples.

We characterized the variation of abiotic physico-chemical parameters with a principal component analysis (PCA) (Annexe 1: Figure S1). The first component of the PCA accounted for 49.8% of the variance, and was positively correlated with conductivity, sulfate (SO_4^{2-}) and dissolved organic carbon (DOC) and negatively correlated with water-table level and oxygen (O_2). The second component accounted for 16.4% of the variance and was mainly correlated with nitrate concentration, which was higher in the fen in June, July and September. The distribution of data points across the first two components emphasized the differences between fen and bog but also highlighted similar temporal trends within the two development stages, distinguishing the March and May samples from September. The seasonal fluctuations of the water-table were closely linked with variation in water-chemistry, suggesting that water sources and flowpaths affect nutrient concentrations, potentially due to dilution of a limited solute stock (for example, Spearman's rank correlations; water-table to conductivity: R=-0.67, P= 6 x 10^{-5}, N=29; water-table to sulfate: R=-0.66, P= 9 x 10^{-5}, N=29) (Annexe 1: Table 1, Annexe 1: Figure S1). Because the first component represents an integrated variable of spatio-temporal variations of the physico-chemical and hydrological parameters, we used sample coordinates on this axis to model the seasonal abiotic fluctuations for both fen and bog. Log-

transformed PA and VPA were both positively, linearly correlated with this abiotic gradient (N=29, PA: r^2=0.53, P<10^{-5}; VPA: r^2=0.41, P<10^{-3}) (Figure 3), whereas VPR was unrelated to the abiotic gradient (r^2=-0.03, P=0.68).

3.4. Viral community composition and diversity

Pyrosequencing of the 12 metaviromes yielded 1 096 889 sequences with an average length of 417 bp (Table 2). Analysis revealed that sequences of ribosomal RNA genes accounted for 0.0063% (69 rRNA sequences), which were excluded from subsequent analysis. In addition, the predicted protein coding sequences that matched the functional category databases (KEGG) and the cluster of orthologous genes database (COG) accounted for only 0.38% and 0.45%, respectively.

Table 2 - Main characteristics of the 12 fen and bog metaviromes.

Date	June 2011		Aug. 2011		Oct. 2011		Mar. 2012					
Origin	*Fen*	*Bog*	*Fen*	*Bog*	*Fen*	*Bog*	*Fen*			*Bog*		
Sample	vFen_June11	vBog_June11	vFen_Aug11	vBog_Aug11	vFen_Oct11	vBog_Oct11	vFen_Mar12_A	vFen_Mar12_B	vFen_Mar12_C	vBog_Mar12_A	vBog_Mar12_B	vBog_Mar12_C
Metavir Id	1368	1369	1373	1374	1375	1376	1377	1378	1379	1380	1382	1370
N° of sequences > 250pb	90 983	36 165	107 181	171 483	67 303	105 296	75 274	62 492	78 169	114 591	122 251	65 701
Average size (bp)	415	418	416	415	409	413	415	416	416	416	414	415
Total nucleotides (Mbp)	37.8	16.4	44.65	71.24	27.55	43.56	31.27	26.05	32.58	47.76	50.71	27.33
N° rRNA	0	0	1	3	0	22	8	3	11	9	7	5
N° of tRNA	50	40	207	249	5	245	119	138	77	109	101	272
Hits to KO_KEGG	19	16	111	151	37	1134	314	132	957	1115	310	76
Hits to COG	69	59	113	276	46	1053	535	278	873	1011	550	167
% "No Hit"	91.30%	88.30%	92.99%	92.83%	94.10%	95.47%	93.64%	94.09%	94.08%	91.90%	89.52%	90.01%
N° shared sequences	1027	109	1492	817	8335	3792	329	195	1786	66	336	42
% shared	1.13%	0.30%	1.39%	0.48%	12.38%	3.60%	0.44%	0.31%	2.28%	0.06%	0.27%	0.06%

This corresponds to 10-50 times fewer matches than are typically found in conventional short read metagenomes (Table 2) (Quaiser et al., 2011), indicating a very low level of contamination by genomic DNA from microorganisms (Roux et al., 2013b). This allows the precise characterization of viral diversity and variation in fen and bog through the year.

3.4. Taxonomic composition

Sequences were compared against viral genomes from the NCBI RefSeqVirus database. Only a small proportion of sequences, ranging from 4.2% (vBog_Oct11, v=virus/Fen or Bog/sampling date) to 10.9% (vBog_June11) matched the available viral genomes indicating the presence of currently undetected viruses (Table 2, Figure 4). Matches associated with ssDNA viruses were most common, accounting for a mean of 4.5% of the total number of sequences, with primary assignment to the bacteriophage family *Microviridae* (1.7 ± 1.3%) and to the eukaryal ssDNA family *Circoviridae* (0.9 ± 0.7%). Matches with dsDNA viruses appeared mostly affiliated with the order of *Caudovirales* (1.3 ± 0.8%), which can be hosted by both Bacteria and Archaea. While the protocol was not designed to preserve RNA viruses, we detected sequences matching to ssRNA viruses affiliated with *Tombusviridae* and *Sclerophtora macrospora* virus A representing likely the recently identified so

called "chimeric viruses" (Diemer and Stedman, 2012; Roux et al., 2013a). They were present in all samples and accounted for 0.1% (vFen_Aug11) to up to 2.8% (vBog_Aug11) of the total metavirome sequences (Figure 4).

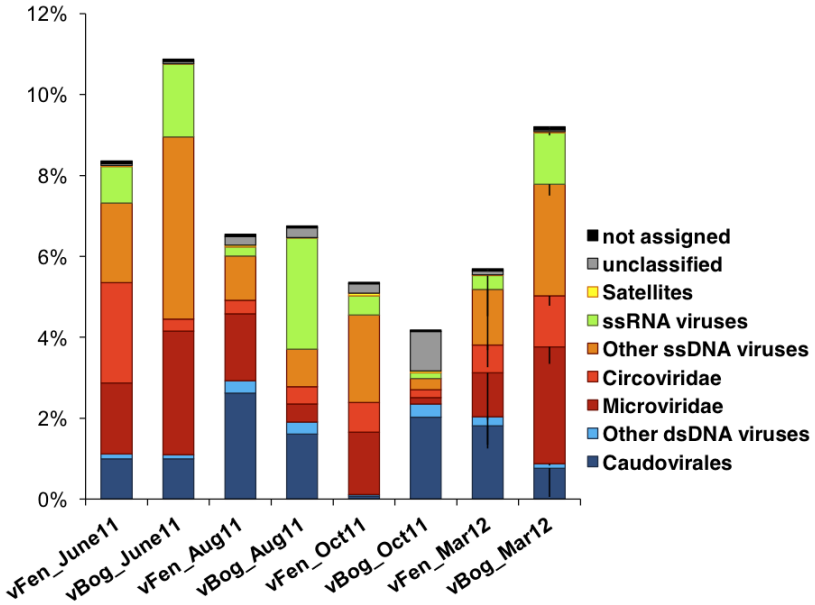


Figure 4 - Relative abundances of the different viral types in the 6 fen and 6 bog metaviromes. Taxonomic affiliations were obtained by tBLASTx against RefSeqVirus (e-value 10^{-7}). Relative abundances were averaged for March fen and bog metavirome triplicates. Error bars are shown only for negative values and represent standard deviation (N=3).

3.5. Genomic diversity based comparisons of the metaviromes

Due to the lack of virus reference sequences in the databases the majority of sequences (88.3 to 95.4%) remained unaffiliated to known viral taxa. To characterize the remaining, unidentified metaviromes sequences, we analyzed the proportion of similar sequences (4 identical kmers of 33bp) between each pair of metaviromes with Compareads (Maillet et al., 2012). The dendrogram built from the similarity matrix showed two well-separated clusters (Annexe 1: Figure S2A). One group was composed of metaviromes collected during summer and autumn 2011 (vBog_Aug2011, vFen_Oct11, vBog_Oct2011), while the second group included the communities sampled in winter 2012 (March 2012) and spring 2011 (June 2011) regardless of the peatland development stage.

Sequence comparisons with Compareads provide a global estimate of the proportion of similar sequences without taking into account the internal structure of the sequence sets. Therefore we clustered sequences with a 95% identity threshold to assess the diversity of protein-coding gene sequences and to determine which groups of sequences drive the similarity between metaviromes in the Compareads analysis. Sørensen dissimilarity was calculated for every pair of metaviromes. This distance is based on the distribution (presence-absence) of sequences from the different clusters in the 12 metaviromes. Sørensen dissimilarities between pairwise metaviromes were uncorrelated with the amount of sequences in pairwise comparisons (Spearman's rank correlation, R=0.08, P=0.495,

N=66). The pattern of similarity between the metaviromes supported the first analysis done with Compareads (Annexe 1: Figure S2B), with nearly identical grouping of summer and autumn communities and winter and spring communities.

Unique sequences represented 28% of the total metaviromes and 49% of the sequences belonged to clusters consisting of 10 to more than 1 000 sequences. We split clusters of highly similar sequences into different size categories, depending on the number of sequences they included (Figure 5) and computed hierarchical clustering based on Sørensen dissimilarities for each size category. Dendrograms indicated that the summer and autumn group of metaviromes (vBog_Aug11, vBog_Oct11 and vFen_Oct11) was distinct from the other metaviromes for all size categories, though this distinction was not significant for clusters smaller than 5 sequences (Figure 5). This suggests a fundamental change in viral community between the two main groups of metaviromes with low intergroup Sorensen similarity (Annexe 1: Figure S3). Resemblance between winter and spring metaviromes (Fen and Bog from June 2011 and March 2012) was only significant for clusters larger than 250 sequences. Thus, the resemblance between June and March metaviromes appears to be due to a small number of large clusters.

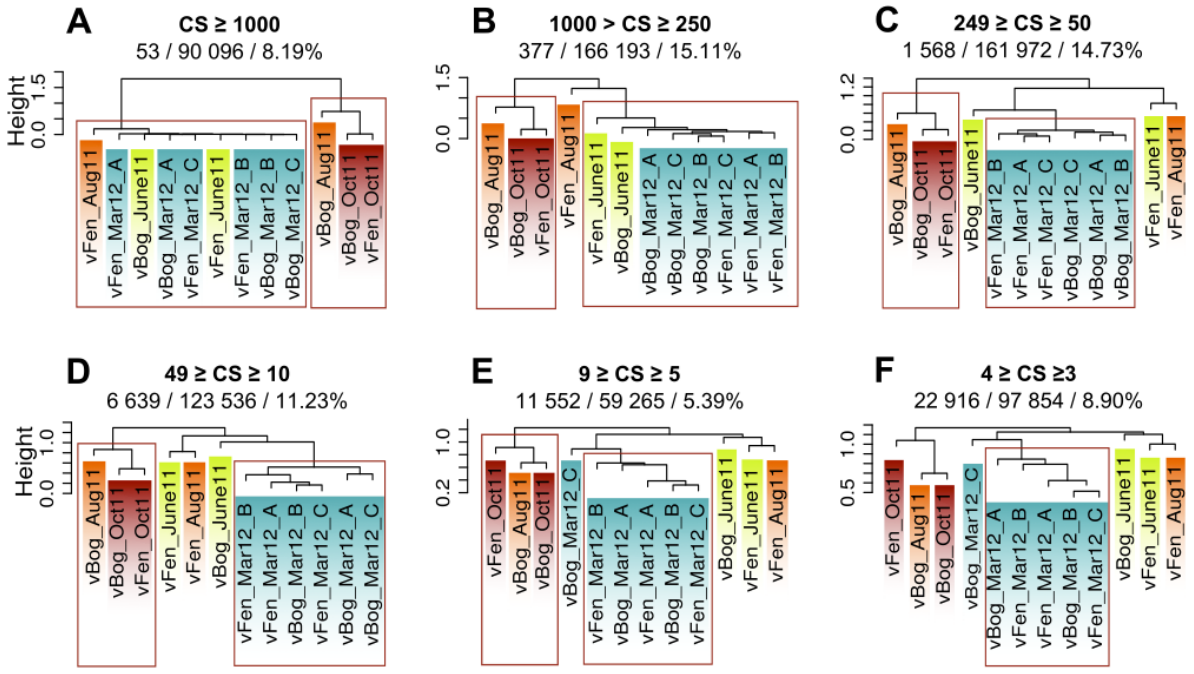


Figure 5 - Metaviromes similarities based on Sørensen Index and on the number of sequences gathered in the clusters. (A), (B), (C), (D), (E), (F) represent different size categories, ranging from the category of clusters that gathered more than 1000 sequences (A) to the category of clusters represented by 3 or 4 sequences (F). Abbreviations: CS: Cluster Size. Numbers below the class size: N° of clusters represented in the size category / Nb of sequences represented in the size category / % of the total metaviromes sequences. Red rectangles represent the most robust clusters (pvclust. au=0.05).

To analyze the genetic similarities with other metaviromes, we compared the 12 peatland metaviromes with 49 available metaviromes from 8 different ecosystem types by hierarchical clustering and tBLASTx (Annexe 1: Figure S4, see Materials and Methods). Peatland viruses formed a distinct group, clearly separated even from geographically close viral communities originating from freshwater lakes, indicating that these metaviromes represent a unique community characteristic of and structured by its ecosystem (Roux et al., 2012).

3.6. Link between viral and microbial communities

To investigate the interactions between viruses and the microbial communities in *Sphagnum*-dominated peatlands, we sequenced twelve metagenomes from the fen and bog prokaryotic communities (Annexe 1: Table S2) from the same day and site as the metaviromes vFen_June11 and vBog_June11 (Annexe 1: Table S1). Metagenome DNA was extracted from the peat matrix allowing finer spatial sampling. In addition, the peat matrix contained the peat pore-water from which the viral particles were sampled. Based on taxonomic affiliations fen and bog prokaryotic communities (hereafter called pFen and pBog) appeared to be predominantly composed of the same main phyla (Annexe 1: Figure S5). However, regardless of depth, fen and bog appeared to harbor prokaryotic communities with distinct structures as shown by non-metric multidimensional scaling ordination of euclidean distances between metagenomes (Annexe 1: Figure S6) and with the analysis of similarity (ANOSIM, Euclidean distance, $R=0.68$, $P<0.01$). In order to identify viral signatures in these metagenomes, we identified sequences shared by metagenomes and metaviromes using Compareads (4 identical kmers of 33bp). We obtained 18,676 "shared" sequences, of which 30 were from bog metagenomes (pBog), 320 were from fen metagenomes (pFen) (Annexe 1: Table S2), and 18,326 from the 12 metaviromes (Table 2). In most metaviromes, the number of "shared" sequences represented less than 1% of the total sequences but reached up to 12% in vFen_Oct11 (Table 2). We clustered sequences with a 95% identity threshold, built a contingency matrix based on the number of sequences from each sample in the different clusters and performed a correspondence analysis (CA) on this contingency matrix.

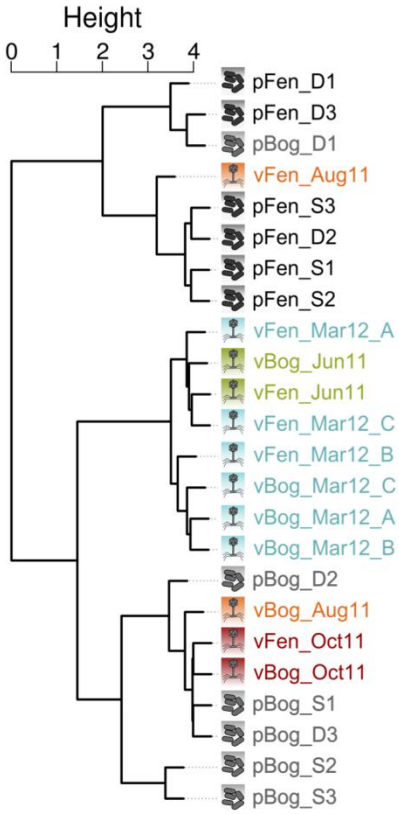


Figure 6 – Hierarchical clustering representing metagenome and metavirome similarity based on sequences shared by at least one metagenome and one metavirome (Compareads1.2.2, k=33. t=4). Shared sequences have been clustered using CD-HIT-EST, and a Correspondence Analysis (CA) was performed on the clusters. Hierarchical clustering is based on the two first axes of the CA. Abbreviations: pFen/pBog: metagenome from fen/bog sampled in June 2011; S/D: surface/depth; 1,2,3: metagenome replicates; vFen/vBog: metavirome from fen/bog; A,B,C: metavirome replicates.

The first two axes of the CA contained 21% of the total information. Hierarchical clustering based on the first two components revealed three groups of samples (Figure 6). Two groups included metagenomes sampled in June 2011 from fen and bog respectively (pFen and pBog_ samples) as well as metaviromes from August and October 2011. The third group contained the metaviromes from June 2011 and March 2012. Thus metaviromes from June 2011 did not cluster with the metagenomes sampled on the same day. Clustering and Sorensen dissimilarity analysis based solely on the subset of metavirome sequences shared with metagenomes revealed the same summer and autumn and winter and spring groupings as obtained for complete metaviromes (Annexe 1: Figure S7). Among these 18 676 "shared" sequences, a total of 774 sequences (4%) were assigned to references in RefSeqVirus (tBLASTx, e-value 10^{-7}) indicating that the vast majority (96%) originate from currently unidentified viruses. Most hits were associated with the *Microviridae* subfamily *Gokushovirinae* (ssDNA viruses, 409 hits) and to a lesser extend with the ssDNA viruses *Circoviridae* (173 hits), *Caudovirales* (dsDNA viruses, 134 hits), and *Sclerophtora macrospora* virus A (ssRNA viruses, 19 hits).

4. Discussion

4.1. Quality of the metaviromes

To explore the diversity and potential ecological role of viruses in *Sphagnum*-dominated peatland, we analyzed and compared 6 fen and 6 bog metaviromes covering the seasonal periods. The quality of the viromes is essential for comprehensive analysis, since contaminations with microbial genomic DNA would falsify the results. To assure that microbial DNA contamination was satisfactorily low, we applied the pegylation procedure to enrich viral particles (Colombet et al., 2007), DNAse treatment to degrade "free" DNA not protected by capsids, and triplicate whole genome amplification to balance potential amplification bias. The high quality of the virome sequences was shown by the very low abundance of rRNA sequences as well as the low number of matches to functional databases (*i.e.* KEGG_KO and COG). This is in accordance with the high diversity present in the viral genomic pool and with the high rates of evolutionary changes in viral genomes that are much less conserved than microbial genes (Duffy et al., 2008).

It has been hypothesized that viruses infecting eukaryotes might be more important in terrestrial ecosystems and wetlands, where protozoan and fungal biomass is higher, while bacteriophages dominate viral consortia in marine and freshwater ecosystems (Farnell-Jackson and Ward, 2003; Jackson and Jackson, 2008; Kimura et al., 2008). Concerning the peatland metaviromes, among the sequences that matched viral genome databases, similar proportions of sequences were assigned to eukaryotic viruses, such as *Circoviridae* and *Sclerophtora macrospora* virus A-like viruses, and to prokaryotic viruses, such as *Caudovirales* or *Microviridae*. However, due to the vast majority of sequences being unassigned and to the high variability of viral genes, it remains impossible to determine whether viral communities in peatlands are dominated by prokaryote or eukaryote infecting viruses.

4.2. Successional patterns of viral and microbial communities

Ecological integration of the viral compartment into ecosystem functioning is mainly obtained through approaches combining virome sequences and viral abundance analysis (Wommack et al., 2015). In order to characterize viral ecology of *Sphagnum*-dominated peatlands, we monitored seasonal abundance and diversity of viruses and prokaryotes at two different depths of fen and bog and attempted to identify whether these were correlated to abiotic factors. Fens and bogs are development stages of peatlands that differ fundamentally in vegetation (*Sphagnum* and vascular plants) and associated physico-chemistry (Rydin and Jeglum, 2006). Thus, as already observed for microbial diversity (Opelt et al., 2007; Bragina et al., 2012), we hypothesized that peatland

development stage and associated *Sphagnum* habitat would be the major driver in the distribution of virus and prokaryote abundance and diversity. Our results confirm that prokaryotic communities differ in structure between the 2 dynamic stages and physcio-chemical conditions but, surprisingly, we did not observe any significant difference in VPA, PA and viral diversity between fen and bog. The viral communities showed no systematic spatial trend and high variability even within replicates.

While we did not detect any significant spatial differences, we observed a significant seasonal fluctuation of virus diversity and abundance. For both fen and bog, VPA and PA (log-transformed data) were strongly correlated with the seasonal fluctuations of water-table, DOC conductivity and sulfate: i.e. VPA and PA were higher when water-table was low and DOC and sulfate were high (Figure 3). DOC has been recognized as a key factor in the C-balance of *Sphagnum*-dominated peatlands (Billett et al., 2004) and its patterns are driven by both biological activity (microbial production and consumption, plant exudation) and abiotic variables such temperature, water-table level or acidity (Clark et al., 2009) with seasonal fluctuations as a consequence (Moore, 1987). While the pH is recognized as an integrated physico-chemical variable, we did not register strong influence on PA and VPA. The increase observed during June and July in the fen is potentially due to photosynthetic activity with strong assimilation of dissolved inorganic carbon by microalgae, which significantly develop in submerged *Sphagnum*-fen at the beginning of summer (Gilbert et al., 1998b). As viral activity is dependent on bacterial production (Middelboe, 2000), the correlations between DOC and VPA and PA suggest a net production of DOC with increasing microbial activity including viral lysis. DOC concentrations also depend on temperature and water flows, which play a key role in the production and redistribution of carbon in the peat (Waddington and Roulet, 1997; Clark et al., 2009), the 2 factors interacting during drawdown and flooding periods but we did not evidence a clear combination of DOC and temperature. The fluctuations of the water-table depend largely on hydrologic inputs that is rainfall in the bog and rainfall and runoff in the fen, which occur at the yearly scale (seasons) and at a short-time scale (episodic events). Sulfate also interacts with DOC concentrations, especially during drought periods when significant production of sulfate may increase peat acidification and ionic strength (Clark et al., 2009). In our study, sulfate concentrations were significantly higher during water-table drawdowns and were positively linked to DOC. In addition, we could not show whether the production of sulfate was linked to nitrate through microbial S° oxidation as already demonstrated in peatlands (Burgin and Hamilton, 2008). Nitrate seems to play a key role only during summer in the fen (see component 2 of the ACP), a dynamic stage in which nitrogen mineralization significantly occurs at this season (Francez and Loiseau, 1999). Our results suggest that viruses and prokaryotes are more abundant at depth regardless of peatland development stage. This is in accordance with a previous study concerning prokaryote abundance

(Dedysh et al., 2006) and potentially due to more buffered temperature and water-table fluctuations at depth providing more stable conditions. Altogether, these results indicate that seasonal changes in temperature and precipitations (allogenic variables) influence PA and VPA *via* water-table fluctuations and consequently nutrient concentrations with larger effects at the surface than at depth.

Comparisons of metaviromes with metagenomes showed that temporal variations are more influential than differences in peatland habitat in structuring viral communities. There was a particularly substantial shift in sequence composition from spring to autumn with distinct patterns of composition and abundance, suggesting the existence of ecological succession of viral communities at the seasonal scale. This pattern appears to be consistent inter-annually, with metaviromes sampled about a year apart (June 2011 and March 2012) clustering together. These findings suggest, for the first time, that a cyclic succession in peatlands affects free-occurring viruses at the community level. Recent studies on marine ecosystems also described seasonal fluctuations of viral communities at the ocean surface (Chow and Fuhrman, 2012; Pagarete et al., 2013). In these studies, which focused on the diversity of a viral gene marker, seasonality was mainly characterized by fluctuations of dominant viral types while in our study, seasonality was associated with a general change in the composition of viral communities.

4.3. Viruses and carbon cycling in Sphagnum-dominated peatlands

Viruses are believed to be key components of the carbon cycle in many ecosystems, both altering carbon fluxes and contributing to C-redistribution through bacterial lysis (Fuhrman, 1999; Middelboe and Lyck, 2002; Ankrah et al., 2014). Despite recent analyses of peatland microbial food-webs (Lamentowicz et al., 2013), the significance of viruses in the functioning of *Sphagnum*-dominated peatlands remains unknown.

In viral ecology, VPR is generally considered as an indicator of the bacterial hosts metabolic state (Williamson et al., 2005; Kimura et al., 2008) because viral burst size, and thus viral abundance is positively correlated with microbial growth rate (Middelboe, 2000). In the studied peatland, the VPR was low and did not differ between fen and bog, despite the differences between dead organic matter produced in the 2 dynamic stages, in relation to the dominant *Sphagnum* species (Francez, 1995; Thormann et al., 2003). The low VPR compared with other ecosystems is likely due to lower metabolic activity of microorganisms, that is in accordance with the functioning of *Sphagnum*-dominated peatlands where decomposition is slowed down due to constraining conditions (Rydin and Jeglum, 2006; Artz, 2009) and the presence of a significant proportion of dormant cells in the community (Dedysh et al., 2006; Pankratov et al., 2011).

We detected a VPR peak in June 2011, just before a broad modification of viral community composition in fen and bog. Viruses interact with their hosts through at least two main strategies: the lytic and the lysogenic life cycles, the latter is believed to be favored when microbial activity is low (Danovaro et al., 2002; Payet and Suttle, 2013; Sime-Ngando, 2014). The change in the viral community composition in summer could result from a seasonal shift in the active part of the microbial community and related C-cycling processes *via* decomposition that show seasonal patterns (Basiliko et al., 2005; Sun et al., 2012). This illustrates a transition from lysogenic to lytic strategies of the viruses infecting the newly active prokaryotes. This hypothesis is supported by the low PA associated with the VPR peak in June 2011, which could result from virus-mediated bacterial lysis, and by the similarities between spring metagenomes (June) and summer and autumn viromes, suggesting the presence of prophages in the microbial genomes in June, that were later released and detected in the metaviromes in August and October.

5. Conclusion

We applied an integrated approach linking virome sequence analysis, viral particle and prokaryote abundance, physico-chemical parameters and metagenome-virome comparison to get insights into the ecological functioning of the viral community in peatlands. We found that viral community abundance and diversity in *Sphagnum*-dominated peatlands express an ecological succession, and that viruses, as well as their hosts, are strongly influenced by the temporal fluctuations at the peatland surface. The observed low VPR, compared with other ecosystems, is in accordance with slowed down decomposition processes in *Sphagnum*-dominated peatlands. The observed shift in viral diversity suggests a seasonal change of the microbial community and the associated switch of viral life-cycle strategies during summer and autumn highlighting the importance of virus-host interactions as they control the dynamics of microbial communities. These patterns may be related to changes in C-cycling processes but further studies are needed to strictly link microbial and viral diversity with C-transformations in the peat. These should focus on the seasonality of viruses that infect different kind of hosts (prokaryotes/eukaryotes) in order to identify the main factors driving this succession and changes in functioning processes, as already suggested for seasonal fluctuations of plankton (Sommer et al., 2012).

To assess the ecological relevance of the diversity of viral communities, analysis of the spatio-temporal dynamics of ecosystem-specific metaviromes as applied here, rather than cross-biomes comparisons, represent a powerful approach to overcome the lack of viral genomes in the databases, and to take advantage of the whole diversity carried in the sequenced viral communities.

Acknowledgements

We thank members of the 'Environmental and Functional Genomics Platform' (Biogenouest Genomics) for sequencing and the molecular ecology facility (UMR 6553 ECOBIO) for support. We thank Benjamin W. Abbott for manuscript editing. We thank Patricia Madec, Nathalie Le Bris and Sophie Coudouel (OSU Rennes) for technical assistance. This work was supported by funding from the INSU-CNRS EC2CO program to A-JF ("Virtou") (Program MICROBIEN, 2011-2012) and grants from the University of Rennes/CNRS and from the "Observatoire des Sciences de l'Univers de Rennes"(OSUR) to AQ. The authors declare no conflict of interest related to the manuscript.

Chapitre 1 B. Diversity, distribution and comparative genomics of *Microviridae* in *Sphagnum*-peat soils

1. Introduction

Viruses, in particular bacteriophages, have been identified in several ecosystems, and represent the most abundant biological entities on Earth. In marine ecosystems and freshwater lakes they are present at high concentrations of about 10^7 particles/mL on average (Wommack and Colwell, 2000). Viruses can control the microbial abundance and influence the composition of microbial communities by lysing their host organisms (Fuhrman and Schwalbach, 2003). In addition, they carry a highly diverse pool of genetic elements that might be exchanged with their hosts, thereby contributing to adaptation processes through the acquisition of new functions (Pedulla et al., 2003). Despite these potential influences the importance is almost solely illustrated in aquatic ecosystems (Wommack and Colwell, 2000; Suttle, 2007), and little is known about the diversity of viruses and in particular about ssDNA viruses in soil ecosystems. ssDNA viruses are divided into two families of bacteriophages (*Inoviridae* and *Microviridae*) and five eukaryotic viruses (King et al., 2012). Recent studies revealed that *Microviridae* seem to be ubiquitous. Through metagenomic approaches, *Microviridae* were identified in marine environments (Angly et al., 2006; Labonté and Suttle, 2013a; 2013b; Tucker et al., 2011), freshwater habitats (Roux et al., 2012a; 2012b; López-Bueno et al., 2009), human gut or feces (Roux et al., 2012b), stromatolites (Desnues et al., 2008), dragonflies (Rosario et al., 2012), sewage and sediments (Hopkins et al., 2014), and even as temperate phages integrated in the genomes of Bacteroidetes species (Krupovic and Forterre, 2011). These studies contribute significantly to the accumulation of *Microviridae* specific genomic information allowing more precise comparative genome analyses. These revealed that all *Microviridae* genomes possess several homologous genes including a well-conserved gene coding for the major capsid protein VP1, a roughly 500 amino acid long protein. This protein can be used as phylogenetic marker facilitating delineation of *Microviridae* clades or subfamilies (Wommack and Colwell, 2000; Desnues et al., 2008; Roux et al., 2012b; Labonté and Suttle, 2013b; Hopkins et al., 2014). All these studies applied whole-genome amplification methodology that preferentially amplifies small circular DNA templates such as ssDNA bacteriophages, which consequently introduces biases (Fuhrman and Schwalbach, 2003; Kim and Bae, 2011). We took advantage of this bias in that the ssDNA bacteriophages were likely preferentially enriched allowing for more targeted diversity analyses and exploration of new *Microviridae* genomes.

In this study, we analyzed 18 complete *Microviridae* genomes reconstructed through *de novo* assembly of reads from twelve viromes one microbial metagenome and obtained from peat soil and water samples collected in a *Sphagnum*-dominated peatland. Detailed phylogenetic analysis and genome comparisons revealed two new subfamilies. Combined with the available genomic information of *Microviridae* we provide new insights into the diversity, distribution and abundance of the *Microviridae* subfamilies.

2. Materials and methods

2.1. Sampling and accession to 12 peat viromes

Twelve samples were recovered from a *Sphagnum*-dominated peatland at "les Pradeaux mire" in the French Massif Central (3°55E; 45°32N) at an altitude of 1,250 m. Water and peat soils were both sampled in young states of peatland dynamics, called 'Fen' with *Sphagnum fallax* and *Carex rostrata* vegetation, and in older states of dynamics, called 'Bog' with *Sphagnum magellanicum* and *Sphagnum capillifolium, Andromeda polyfolia* and *Eriophorum vaginatum* (Pedulla et al., 2003; Francez and VASANDER, 1995). One sample for each Fen and Bog was collected in June, August and October 2011 as well as three biological replicates of Sphagnum water that was extracted and filtered at 125µm from both Fen and Bog (March 2012). Viruses were concentrated using PEGylation (Wommack and Colwell, 2000; Colombet et al., 2007; Suttle, 2007). Extracellular DNA was digested with DNAse RQ1 (Promega) at 37°C for 1h. Viral DNA was extracted with the Nucleospin Extract II kit (Macherey-Nagel). Whole genome amplification (WGA) reactions were run in triplicate for each sample using Genomi-Phi following the instructions of the manufacturer (GE Healthcare) and subsequently pooled. Library construction and pyrosequencing was performed at the "Functional and Environmental Genomics" platform (OSUR, Rennes, France). After sequence quality and size trimming we obtained 481,402 and 618,487 reads from Fen and Bog respectively with an average length of 415 bp. *Microviridae* genomes reported in this publication have been deposited in the Sequence Read Archive under the study accession number KM589498-KM589516.

2.2. Assembly, annotation and comparative genome analysis

The peat viromes were assembled separately into contigs using the meta-assembler of CAMERA under standard conditions (King et al., 2012; Seshadri et al., 2007). The circularity of the genomes was checked manually. Only *Microviridae* genomes that showed overlapping reads at the start and end of the contigs, confirming their circular genomes were retained. Sequence analysis was performed using an integrated WEB-based annotation platform adapted to metagenomic sequence analysis (Genomic toolbox) (Angly et al., 2006; Zivanovic et al., 2014; Labonté and Suttle, 2013a;

2013b; Tucker et al., 2011) (http://www-archbac.u-psud.fr/projects/aq_virome/aq_virome.html). Candidate coding sequences (CDS) were defined with PRODIGAL (version 2.60) (Roux et al., 2012a; Hyatt et al., 2010; Roux et al., 2012b; López-Bueno et al., 2009). Semi-automated annotation was performed to identify genes by sequence similarity and coding probability using BLASTP (Roux et al., 2012b; Altschul et al., 1997) against the RefSeq NR protein database (GenBank), SWISSPROT (version 57.11), and COG databases (COG + KOG, 7 eukaryotic genomes). Manual annotation was completed within the above-mentioned platform. Gene order comparison was performed using the platform tool GENOMAPPER (Desnues et al., 2008; Zivanovic et al., 2014) combined with multiple alignment and phylogenetic analysis. To allow the interactive visualization of genomic fragment comparisons, we used Artemis Comparison Tool ACTv.6 (Rosario et al., 2012; Carver et al., 2012).

The presence of *Microviridae* in peat metagenomes that were constructed from corresponding peat samples without virus enrichment (Quaiser et al., in preparation) was analyzed by BLASTN and reads matching the virome-assembled *Microviridae* genomes were selected. The *Microviridae* contig from the metagenome was assembled manually using 21 reads. Since the reads showed nearly 100% identity to the *Gokushovirinae* genome Fen7875_21, it was used as a reference for recruitment of metagenome reads.

2.3. Phylogenetic analysis

Sequences of full length major capsid protein and replication protein as well as the complete nucleic acid sequences of the *Gokushovirinae* genomes were aligned using MUSCLE (Hopkins et al., 2014; Edgar, 2004). Alignments were manually edited using ARB (Krupovic and Forterre, 2011; Ludwig et al., 2004). Gaps and ambiguously aligned positions were excluded from phylogenetic analysis. Maximum likelihood trees were reconstructed using TREEFINDER (Jobb et al., 2004) applying a JTT model (amino acid) and GTR3 (nucleic acid) of sequence evolution with a four category discrete approximation of a distribution plus invariant sites. The best-fit models were determined using TREEFINDER. Maximum likelihood bootstrap proportions were inferred using 1000 replicates.

2.4. Major capsid protein structural modeling

I-TASSER (Roy et al., 2010) was used to assess a first model of a VP1 protein from each of the two new clades (Bog5275_51 for *Aravirinae*, Fen51_42 for *Stokavirinae*). These initial models were processed through multiple steps of loops refinement with MODELLER (Eswar et al., 2008) to improve their quality (PROSA-WEB) (Wiederstein and Sippl, 2007). The final quality Z-scores of the obtained models was -5.15 and -4.59 for Bog5275_51 (*Aravirinae*) and Fen51_42 (*Stokavirinae*) respectively, which according to the ProSA-web server is in the range of X-ray based models, though slightly higher than the two reference VP1 models (-6.4 for ϕX174 F and -6.14 for SpV4 VP1).

Visualization of the structural models and sequence conservation was performed with UCSF CHIMERA (Pettersen et al., 2004).

2.5. Distribution of Microviridae

We compiled available full-length major capsid protein sequences from the complete *Microviridae* genomes to a total of 88 sequences and attributed the taxonomic affiliation according to the phylogenetic analysis in Figure 1. BLASTX analysis against the major capsid protein sequences was performed for the 12 *Sphagnum*-peat viromes and 69 viromes from public databases. To assure the reliability of matches a strict cut-off value of $1e^{-10}$ was applied. The number of matches was normalized to the total number of reads in the different viromes.

3. Results and discussion

3.1. Assembly, identification and analysis of complete Microviridae-like genomes

To get insights into the genome structure and the distribution of *Microviridae*, reads from 12 peat viromes were assembled separately into contigs using the meta-assembler of CAMERA (Seshadri et al., 2007). In total 840 contigs longer than 3 kbp were obtained. Of these 107 contigs showed high sequence similarity to the major capsid protein VP1 of *Microviridae* identified with TBLASTX analysis and by GENE RELATIONS and SYNTENY ANALYSIS tools from the Genomics toolbox (see Material and Methods). Only contigs corresponding to complete circular genomes were retained. In total, we obtained 17 new complete bacteriophage genomes affiliated to *Microviridae*. Sequencing coverage ranged from 8.31 to 94.63 times (Annexe 2: Supplementary Table S1). One additional genome was assembled from reads of a microbial metagenome constructed from a peat sample without viral enrichment (Quaiser et al. 2014). Interestingly, several genomes assembled independently from different samples exhibited very high levels of similarity, thus validating the assembly process of viral genomes from metagenomic reads used in this study and confirming that the obtained assemblies were not chimeric and represented real *Microviridae* genomes.

3.2. Diversity of peat Microviridae accessed by phylogenetic analysis of the conserved major capsid protein VP1 and replication protein VP4

To assess the diversity of the new *Microviridae*-affiliated genomes, phylogenetic analysis was performed using the translated sequence of the gene coding for the major capsid protein VP1 (Figure 1). Eight different clades were formed using representative sequences from already identified *Microviridae* subfamilies and VP1 sequences from the 18 peat *Microviridae* genomes.

Five *Microviridae* genomes recovered from the viral fraction of the peat samples as well as the genome assembled from a peat metagenome clustered within the subfamily of *Gokushovirinae*. Two of these genomes derived from Fen and three from Bog samples. Peat *Gokushovirinae* were clearly separated from cultured ssDNA bacteriophage representatives (Chlamydiaphages: Chp 1-4, phiCPG1, CPAR39; Bdellovibriophage: phiMJ2K) and clustered with recently assembled *Gokushovirinae* genomes from planktonic microbial communities (GM1) (Labonté and Suttle, 2013a), freshwater lake (Bourget_248, Bourget_259) (Roux et al., 2012a) and human feces (Roux et al., 2012b).

One new clade was formed by three genomes from three different viromes (Fen4707_41, Bog1249_12, Fen51_42) representing a novel subfamily. We propose to name it *Stokavirinae* (Stoka: small in Sanskrit). A second new clade, which we propose to name *Aravirinae* (Ara: little in Sanskrit), consisted of 7 new genomes assembled from five different peat viromes (Bog9017_22, Bog5275_51, Fen685_11, Fen7895_21, Fen418_41, Fen2266_11, Fen7786_21). Only one genome, Fen7918_21, was affiliated with the recently defined subfamily of *Pichovirinae* that, to date, consists of 7 assembled genomes isolated from freshwater and marine environments (Roux et al., 2012b). Genome Fen7940_21 was affiliated with the dragonfly-associated microvirus representing the second member of this clade of ssDNA microviruses (Rosario et al., 2012) and thereby confirming its distinction into a potentially novel subfamily (Group D). *Alpavirinae* that were identified as integrated prophages in Bacteroidetes genomes associated with human gut and oral microbiota (Krupovic and Forterre, 2011; Roux et al., 2012b) as well as sequences affiliated to the genus Microvirus (proposed subfamily *Microvirinae*) could not be identified in any of the peat viral genomes. Phylogenetic analysis of VP1 including recently amplified sequences from *Gokushovirinae* changed slightly the tree topology (Annexe 2: Supplementary Figure S1)(Labonté and Suttle, 2013a). In this case Bog9017_22 was clearly separated from *Aravirinae* and Fen7940_21 was apart from dragonfly-associated microvirus. This is probably due to the shorter alignment imposed by the PCR-generated sequences that lower the resolution of the phylogenetic analysis.

Taxonomic affiliation was additionally determined using sequences of the replication protein (VP4), which is shared by all the *Microviridae* genomes (Annexe 2: Supplementary Figure S2). While the replication protein is shorter and less conserved than the major capsid protein, phylogenetic analysis revealed similar clades in both with the exception of Bog9017_22, which moved from *Aravirinae* (major capsid protein phylogeny) to *Pichovirinae* in replication protein phylogeny. While this could hint for a chimeric assembly of the genome, the replication protein sequence phylogeny is much less reliable due to the reduced length and lower overall sequence similarity (Annexe 2: Supplementary Figure S2).

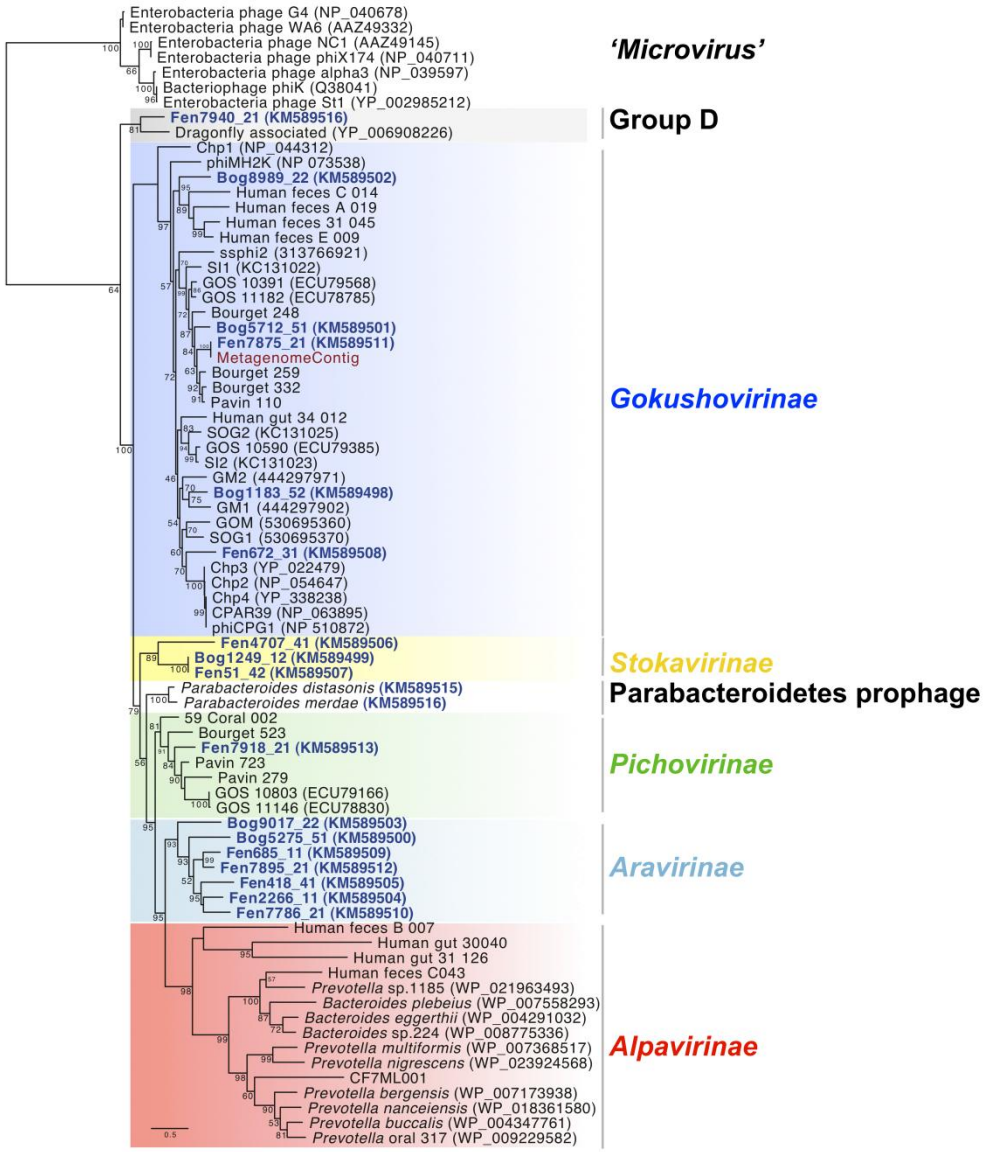


Figure 1 - Maximum likelihood phylogenetic analysis of full-length major capsid protein sequences present in the *Microviridae* genomes from *Sphagnum*-dominated peat viromes. A total of 253 unambiguously aligned positions from 76 sequences were used in the analysis. Bootstrap values are indicated at the nodes. The scale bar indicates the number of substitutions per position for a unit branch length. Chlamydiaphages: Chp 1-4, CPAR39, phiCPG1; Bdellovibriophage: phiMH2K.

3.3. Structure of the Aravirinae and Stokavirinae major capsid protein

Structural modeling of the major capsid protein from the two new *Microviridae* subfamilies indicated that they harbor a conserved eight-stranded β-barrel core ("viral jelly- roll") and a loop extension (Figure 2a) similar to *Alpavirinae* and *Pichovirinae* which is known to form mushroom-like protrusions in SpV4 (Roux et al., 2012b). These protrusions, found in every type of *Microviridae* except for the *Microvirinae*, are thought to bind to host receptors (Roux et al., 2012b).

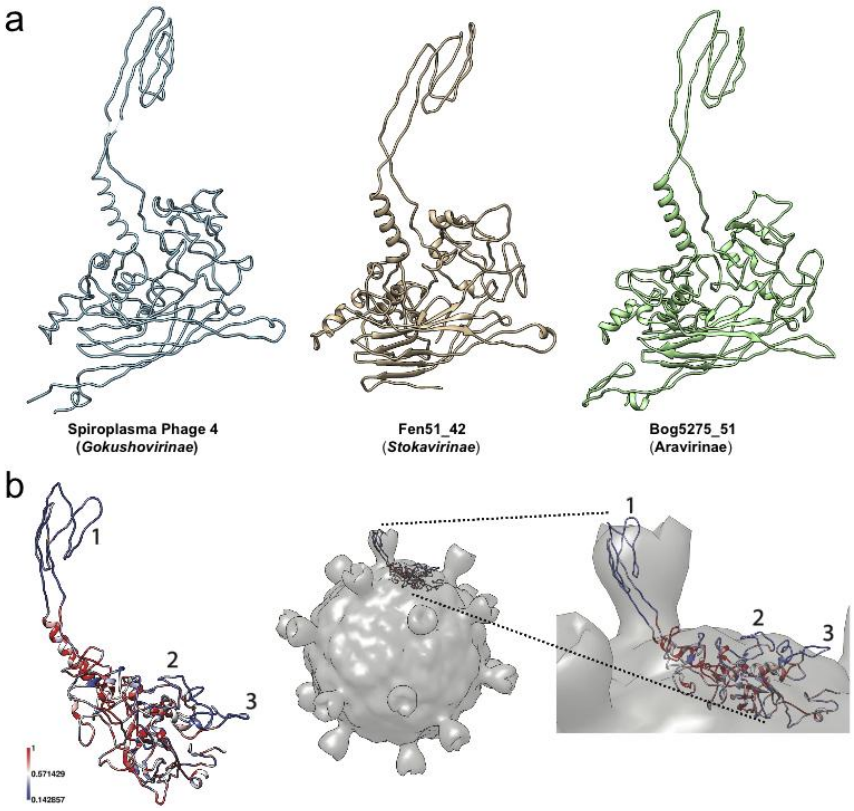


Figure 2 – Modeling of major capsid proteins. (A) Three-dimensional models of *Microviridae* major capsid protein. Different subfamilies are depicted with, from left to right, the reference model from *Spiroplasma* Phage 4 (Pdb Id: 1KVP), representatives from *Stokavirinae* and *Aravirinae*. (B) Hypervariable regions of *Aravirinae* major capsid protein. The sequence conservation across the 7 *Aravirinae* sequences was mapped on the three-dimensional model obtained for the Bog5275_51 genome. The three hypervariable regions were numbered (left panel). This model was mapped on the capsid structure of *Spiroplasma* phage 4 to estimate the position of these hypervariable regions relative to the whole virion (right panel).

Based on the seven different *Aravirinae*, we analyzed the level of residue conservation along the major capsid protein sequences. As expected, the protrusion loop is highly variable (region 1 on Figure 2b). Moreover we identified two additional variable regions backing the separation into a new subfamily (Figure 2b). Interestingly, when mapping the protein model on the virion structure, these two additional variable regions appear to be situated on the virion surface. It is thus tempting to speculate that these regions are forming additional outer structures involved in phage-host recognition.

3.4. Identification of Microviridae prophages in bacterial genomes

Blast searches with the major capsid protein sequences from the assembled genomes showed similarities to VP1 proteins encoded in the genomes of the bacteria *Parabacteroides distasonis* and *Parabacteroides merdae* (Sakamoto and Benno, 2006). A detailed inspection of the genomic environment of the *Parabacteroidetes* VP1 gene showed that genes coding for homologs of VP2 and VP4 were located next to the VP1 gene in a 5 kbp region. This indicates the presence of a prophage

affiliated to *Microviridae* in both *Parabacteroidetes* species. In order to compare their genomic organization, these identified prophage regions were cut from the two genomes and considered as circular genomes. For comparative analysis the genome start was arbitrarily fixed at VP1. Synteny analysis confirmed that the two prophages shared a common ancestor showing the same gene order (VP1-ORF1-ORF2-ORF3-VP2-VP4 for *Parabacteroides distasonis* and VP1-ORF1-VP2-VP4 for *Parabacteroides merdae*) (Figure 3). The gene coding for an uncharacterized protein (ORF1) located downstream of the VP1 coding gene was specific to the two prophages and did not match with genes from other *Microviridae* genomes or from the NCBI non-redundant database. This strengthens the hypothesis that these prophages represent a distinct subfamily of *Microviridae* as suggested by the phylogenetic analysis of the major capsid protein sequences (Figure 1).

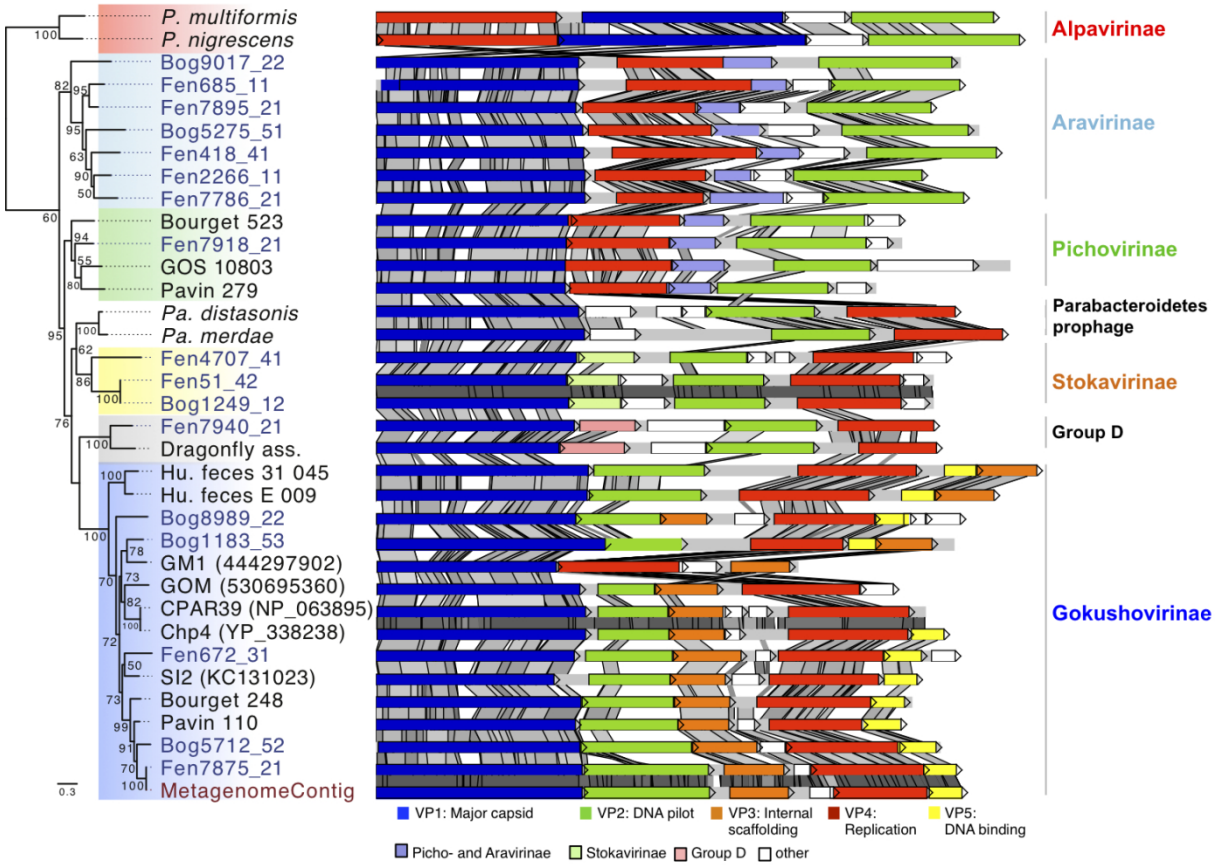


Figure 3 – Major capsid protein phylogeny and genome structure of major subfamilies of *Microviridae* bacteriophages. The affiliation of the peat genomes to *Microviridae* is shown by major capsid protein phylogeny (Maximum likelihood, 252aa positions, 1000 iterations, JTT+G model). Bootstrap values above 50% are indicated at the nodes. The genome structures of the assembled viral peat genomes (blue) were compared to other *Microviridae* genomes. Pairwise comparison (TBLASTX) was visualized with ACT (Carver et al., 2005). Gray shading indicates the level of similarities. Homologous genes specific for some subfamilies are color-coded as indicated in the legend. *P. multiformis, Prevotella multiformis* (WP_007368517)*; P. nigrescens, Prevotella nigrescens* (WP_023924568)*; Pa. distasonis, Parabacteroides distasonis* (ZP_17317332)*; Pa. merdae, Parabacteroides merdae* (WP_022322420). Dragonfly, Dragonfly-associated phage (YP_006908226); GOS_10803 (ECU79166); GM1 (444297902); GOM (530695360); CPAR39 (NP_063895); Chp4 (YP_338238); SI2 (KC131023); Metagenome: genome assembled from the metagenome.

3.5. Comparative genome analysis

To get further insights into the diversity and evolution of the new *Microviridae* genomes the genome structures and gene sequence conservation levels were compared (Figure 3). The characteristic genes encoding the major capsid protein (VP1), replication protein (VP4) and DNA pilot protein (VP2) were identified in all *Microviridae* genomes. All analyzed *Gokushovirinae* possessed one additional gene coding for an internal scaffolding protein (VP3) and in most cases a DNA binding protein (VP5). Both genes were absent in all other subfamilies of *Microviridae*. The sequence similarity of VP3 and VP5 among the peat *Gokushovirinae* ranged from 31-100% and 26-100% identity, respectively. The gene order in *Gokushovirinae* was relatively well conserved, but we could distinguish two clades. In one clade the gene order consists of VP1, VP2, VP3, VP4 and VP5 (i.e. Fen672_31) while in the other clade the organization was VP1, VP2, VP4, VP5 and VP3 (Bog1183_53, Human_feces_E_009 and Human_feces_31_045). However, this distinction was not supported by phylogenetic analysis of the major capsid protein or the replication protein.

The two genomes, representative for the Group D (Dragonfly-associated and Fen7940_21), contain 5 genes. Compared to other groups, they possess two additional genes organized in the same order between VP1 and VP2. The gene situated downstream of VP1 was conserved with up to 50% similarity but only over 41% of their amino acid sequences (Figure 3, Group D, pink). Since no other matches to this gene were found, neither in the NCBI non-redundant protein database nor in other peat *Microviridae* genomes, this gene was a unique feature for this clade confirming the distinction from other *Microviridae* clades. The second protein encoded upstream of VP2 appeared to be unique to each genome as they exhibited no similarity between each other or with proteins found in the NCBI nr database.

The three *Stokavirinae* genomes showed a similar organization. The protein encoded downstream of the major capsid protein gene was homologous within the *Stokavirinae* (Figure 3, light green) but without similarity to *Parabacteroidetes* prophage proteins or to the NCBI NR database. *Stokavirinae* possess a characteristic additional ORF downstream of VP4 that is absent in other subfamilies confirming that they present a distinct clade. Interestingly, the two genomes Bog1249_12 and Fen51_42 showed 99% nucleic acid identity while the third Fen4704_41 was more distant. Both were assembled independently not only from different viromes but also from bog and fen samples recovered at different dates. This indicates that the viral community was at least partially shared among different peat samples.

3.6. Whole genome phylogeny and comparative genome analysis of Gokushovirinae

In order to determine the precise affiliation of the *Gokushovirinae* peat genomes, whole genome phylogeny was performed for this subfamily (Figure 4). The phylogenetic tree obtained reveals three groups that could not be identified based on analysis of the sole VP1 sequences: *Chlamydiaphage Gokushovirinae*, Group 2 and Group 3. As reported recently (Labonté and Suttle, 2013a) all cultured *Gokushovirinae* deriving from *Chlamydia* species, with the exception of Chp1 that appeared more distantly related, clustered together and their genome sequences were strongly conserved (Chp2, Chp3, Chp4, CPAR39 and phiCPG1). Group 2 contained phiMH2K (*Bdellovibrio*) and SpV4 (*Spiroplasma*) as well as recently assembled genomes from marine seawater (SOG1, GOM, ssphi2, SOG2, SI2) (Labonté and Suttle, 2013a) and marine sediments (GM1) (Yoshida et al., 2013). Group 3

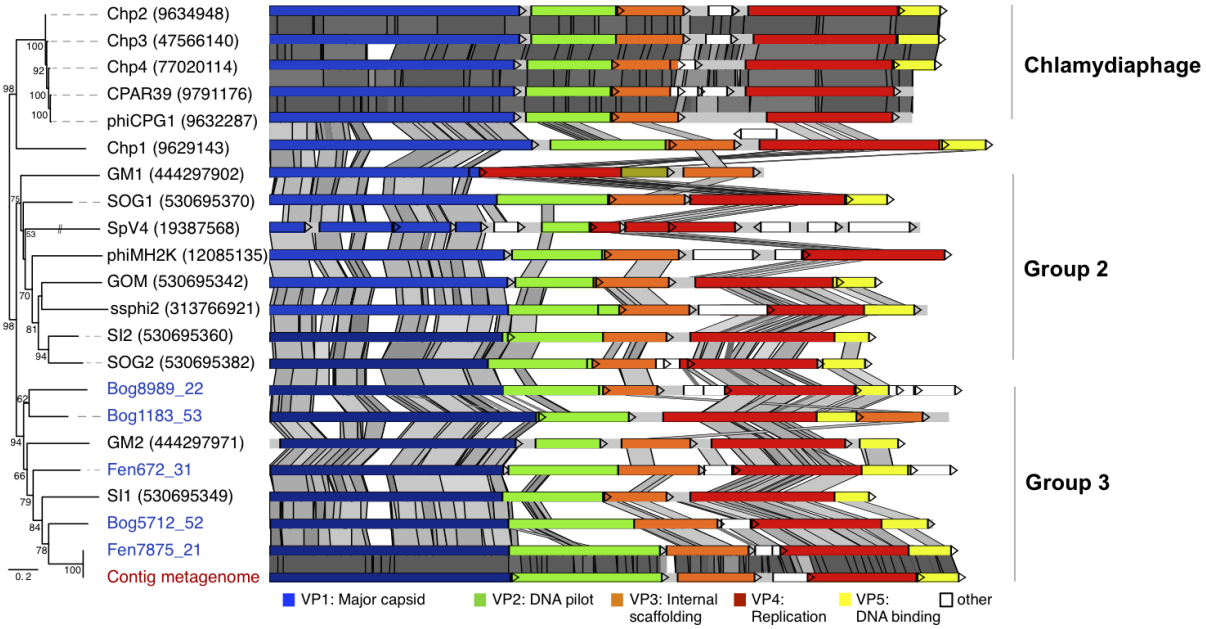


Figure 4 – Whole genome phylogeny and conserved genome structure of peat *Gokushovirinae*. The affiliation of the peat genomes to *Gokushovirinae* was shown by whole genome phylogeny (Maximum likelihood, 2052 positions, 1000 iterations, GTR3 model). Maximum likelihood bootstrap values above 50% are indicated at the nodes. The genome structure of the assembled viral peat genomes (virome red, metagenome blue) was compared to *Gokushovirinae*. Pairwise comparisons (TBLASTX) were visualized by ACT (Carver et al., 2005). Gray shading indicates the level of similarities.

contained the five peat *Gokushovirinae* and two assembled marine genomes SI1 (Labonté and Suttle, 2013a) and GM2 (Yoshida et al., 2013). Four peat *Gokushovirinae* exhibited the typical gene ordering, but in Bog1183_53 the VP3 gene was found downstream of the gene encoding VP5. Identities among the analyzed genomes ranged from 44.1% to 100% identical positions (Annexe 2: Supplementary Table S2). Interestingly, the metagenome derived genome showed 100% similarity to the Fen7875_21 genome.

3.7. Distribution and abundance of Microviridae subfamilies based on matches to the major capsid protein sequence

In an attempt to estimate the distribution and abundance of *Microviridae* subfamilies the relative abundance of the major capsid protein coding genes was determined in the 12 peat viromes and in 69 additional viromes from 12 types of ecosystems (Figure 5). Quantitative estimations of viruses based on the detection of particular signature sequences is currently the best representative proxy, but should be considered with a grain of salt as whole genome amplification leads to overrepresentation of small ssDNA viruses (Kim and Bae, 2011), Nevertheless, to date whole genome amplification is the best way to recover sufficient viral genomic DNA for sequencing. Accordingly, all published viromes and the viromes included in this analysis were generated using this technique. Based on the phylogenetic analyses a database containing 88 major capsid protein sequences representing the different subfamilies of *Microviridae* was constructed. Each virome was searched with BLASTX against the major capsid protein sequence database and best matches were counted using strict count conditions (e-value 10^{-10}). To take into account virome size variation, relative proportions of *Microviridae* subfamilies were obtained through the normalization of the matches by the total number of sequences in each virome. Taxonomic affiliation was determined according to the eight different subfamilies established in phylogenetic analyses (Figure 1). *Alpavirinae* represented the major subfamily of *Microviridae* in all human feces and human saliva viromes reaching up to 18.78% in the virome Human feces A (Kim et al., 2011)(Figure 5A) while they were absent in all peat viromes as well as in most other analyzed viromes. *Gokushovirinae* were the most abundant subfamily in most peat viromes ranging from 0.11% in sample vBog_Oct11 to 4.19% in sample vBog_Mar12_B (Figure 5B). *Gokushovirinae* also represented the majority of the *Microviridae* retrieved in Shimokita and Ogasaware marine sediment viromes (Yoshida et al., 2013) representing 6.00% and 7.47%, respectively, and even reaching 21.66% of the Coral A5 virome (Soffer et al., 2014) (Figure 5A).

With major capsid protein matches accounting for 0.27% of the total virome sequences, *Aravirinae* were the most abundant *Microviridae* subfamily in vBog_Aug11. They represented the second most abundant subfamily in vFen_Jun12, vFen_Aug12, vFen_Mar12_C, vBog_Mar12_A and vBog_Mar12_B. *Stokavirinae* were the second most abundant in vBog_June11, vFen_Mar12_B and vBog_Mar12_B. *Stokavirinae* and *Aravirinae* were not restricted to peatland. They were present in other viromes but seemed to be much less frequent as only 10 and 270 matches were detected respectively from a total of 282,508 major capsid protein matches identified in the 69 non-peat viromes. Matches to *Parabacteroidetes* prophages were low, ranging from 0 in vBog_June11 to 59 in

vFen_Aug11, which corresponds to 0.055% of total viral sequences. Group D was the second most abundant subfamily in the virome vFen_Mars12_A with major capsid protein sequences accounting

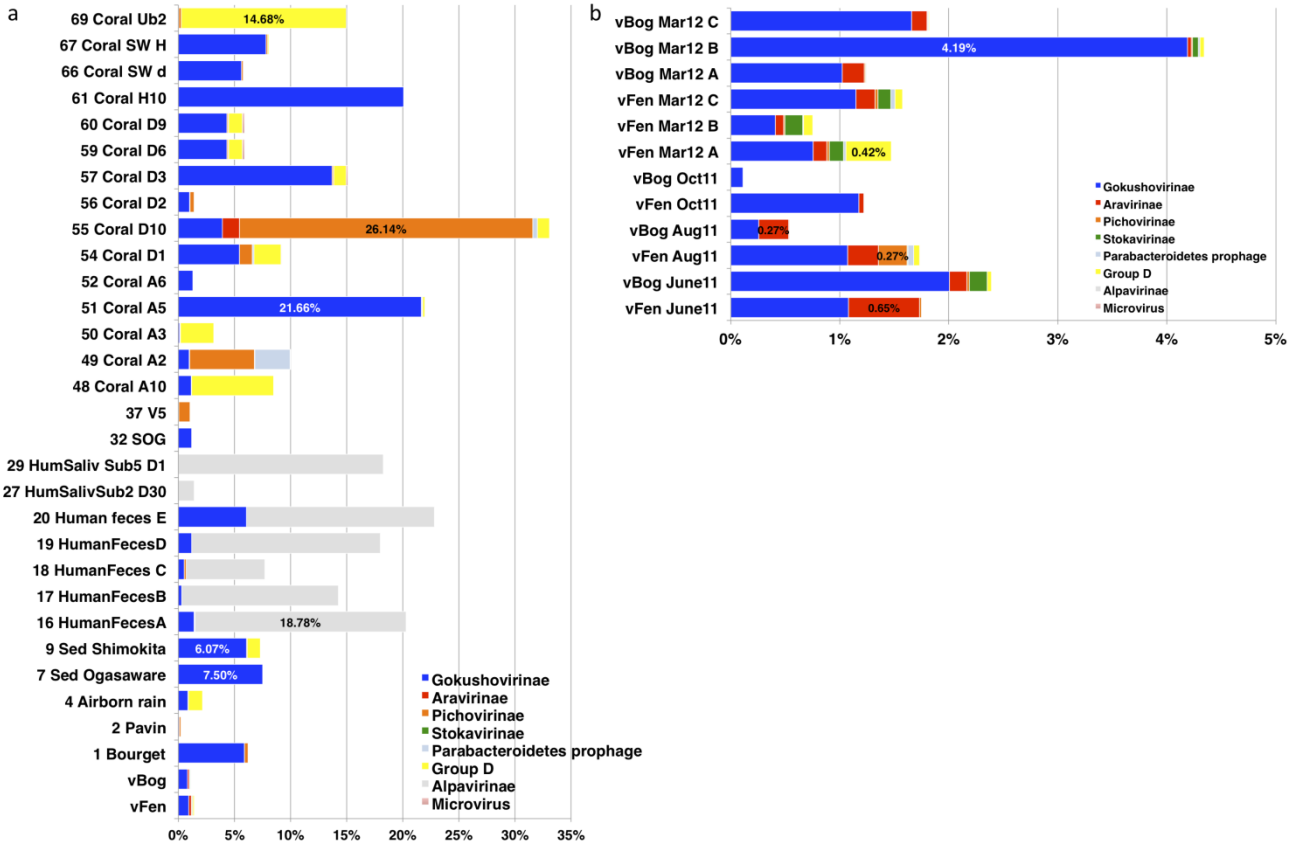


Figure 5 - Relative proportions of *Microviridae* subfamilies based on their content of major capsid protein sequences. (A) VP1 match counts in 69 viromes from public databases; (B) in the 12 peat viromes. Matches to major capsid proteins were normalized to the number of total sequences in the viromes. Best matches were determined by BLASTX with an e-value 10^{-10} cut off.

for 0.42%. As for other viromes, Group D affiliated major capsid protein sequences represented up to 14.68% in the virome Coral Ub2 (Soffer et al., 2014). *Pichovirinae* were present in all peat viromes with the exception of vBog Oct11 reaching up to 0.27% in vFen_Aug11. In the virome Lake Bourget (Roux et al., 2012b), where *Pichovirinae* were first detected, they accounted for 0.31%. The highest proportion was found in Coral D10 (Soffer et al., 2014), where the major capsid protein sequences of *Pichovirinae* accounted for 26.14% (Figure 5A).

4. Conclusion

While bacteriophages are already considered as important biological actors in seawater, freshwater, and human gut ecosystems, this study sheds light onto the specific diversity of *Microviridae* in soil. Namely, we identified two completely new *Microviridae* subfamilies from Sphagnum-peat habitats that are distinct by phylogenetic analysis and by their genome structure. By combining whole genome phylogeny, major capsid protein sequence phylogeny and genome structure analysis; we gained new detailed insights into the evolution and genomic diversity of *Microviridae* subfamilies.

Notably, we highlighted new hypervariable regions in the major capsid protein sequence as well as the co-existence of different genome organizations. Diversity and distribution analysis using the major capsid protein as a marker gene revealed that *Gokushovirinae* have the largest environmental distribution and contributed the most to the pool of *Microviridae*-affiliated reads in a majority of viromes. Moreover, these new genomes allowed for the detection of a second group of *Microviridae* prophages, along with the *Alpavirinae* (Krupovic and Forterre, 2011), in two genomes of Bacteroidetes species that likely represent a novel *Microviridae* subfamily. These new *Microviridae* genomes significantly augment the currently available genome information facilitating subsequent comparative and diversity analyses.

Acknowledgements

We thank members of the 'Environmental and Functional Genomics Platform' (Biogenouest Genomics) for sequencing. We thank Mart Krupovic for assistance and advice on capsid modeling. We thank Daniel Revillini for manuscript editing. This work was supported by funding from the INSU-CNRS EC2CO program to AJF ("VIRTOU"), a grant from the University of Rennes/CNRS and from the ''Observatoire des Sciences de l'Univers de Rennes'' (OSUR) to A.Q.

Références du Chapitre 1

Adriaenssens, E. M., Van Zyl, L., De Maayer, P., Rubagotti, E., Rybicki, E., Tuffin, M., and Cowan, D. A. 2015. Metagenomic analysis of the viral community in Namib Desert hypoliths. *Environ. Microbiol*. 17: 480-95.

Altschul, S. F., Madden, T. L., Schäffer, A. a., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*. 25: 3389-3402. doi:10.1093/nar/25.17.3389.

Andersen, R., Grasset, L., Thormann, M. N., Rochefort, L., and Francez, A. J. 2010. Changes in microbial community structure and function following Sphagnum peatland restoration. Soil Biol. *Biochem*. 42: 291-301. doi:10.1016/j.soilbio.2009.11.006.

Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., Chan, A. M., Haynes, M., Kelley, S., Liu, H., et al. 2006. The marine viromes of four oceanic regions. *PLoS Biol*. 4: e368. doi:10.1371/journal.pbio.0040368.

Ankrah, N. Y. D., May, A. L., Middleton, J. L., Jones, D. R., Hadden, M. K., Gooding, J. R., LeCleir, G. R., Wilhelm, S. W., Campagna, S. R., and Buchan, A. 2014. Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J*. 8: 1089-100. doi:10.1038/ismej.2013.216.

Artz, R., Anderson, I. C., Chapman, S. J., Hagn, A., Schloter, M., Potts, J. M., and Campbell, C. D. 2007. Changes in fungal community composition in response to vegetational succession during the natural regeneration of cutover peatlands. *Microb. Ecol*. 54: 508-22. doi:10.1007/s00248-007-9220-7.

Artz, R. 2009. Microbial Community Structure and Carbon Substrate use in Northern Peatlands, in *Carbon Cycling in Northern Peatlands* (American Geophysical Union): 111-129. Available at: http://dx.doi.org/10.1029/2008GM000806.

Basiliko, N., Moore, T. R., Lafleur, P. M., and Roulet, N. T. 2005. Seasonal and Inter-Annual Decomposition, Microbial Biomass, and Nitrogen Dynamics In a Canadian Bog. *Soil Sci*. 170: 902-912.

Billett, M. F., Palmer, S. M., Hope, D., Deacon, C., Storeton West, R., Hargreaves, K. J., Flechard, C., and Fowler, D. 2004) Linking land-atmosphere-stream carbon fluxes in a lowland peatland system. *Global Biogeochem. Cycles* 18.

Bragina, A., Berg, C., Cardinale, M., Shcherbakov, A., Chebotar, V., and Berg, G. 2012. Sphagnum mosses harbour highly specific bacterial diversity during their whole lifecycle. *ISME J*. 6 : 802-813. doi:10.1038/ismej.2011.151.

Burgin, A. J., and Hamilton, S. K. 2008. NO3- Driven SO4 2- Production in Freshwater Ecosystems: Implications for N and S Cycling. *Ecosystems* 11: 908-922.

Carver, T., Harris, S. R., Berriman, M., Parkhill, J., and McQuillan, J. A. 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics* 28: 464-469. doi:10.1093/bioinformatics/btr703.

Chow, C.-E. T., and Fuhrman, J. A. 2012. Seasonality and monthly dynamics of marine myovirus communities. *Environ. Microbiol*. 14: 2171-2183. doi:10.1111/j.1462-2920.2012.02744.x.

Chow, C.-E. T., Kim, D. Y., Sachdeva, R., Caron, D. a, and Fuhrman, J. A. 2014. Top-down controls on bacterial community structure: microbial network analysis of bacteria, T4-like viruses and protists. *ISME J*. 8: 816-29. doi:10.1038/ismej.2013.199.

Clark, J. M., Ashley, D., Wagner, M., Chapman, P. J., Lane, S. N., Evans, C. D., and Heathwaite, A. L. 2009. Increased temperature sensitivity of net DOC production from ombrotrophic peat due to water table draw-down. *Glob. Chang. Biol*. 15: 794-807. doi:10.1111/j.1365-2486.2008.01683.x.

Clymo, R. S. 1984. The Limits to Peat Bog Growth. *Philos. Trans. R. Soc. B Biol. Sci*. 303: 605-654. doi:10.1098/rstb.1984.0002.

Clymo, R. S., and Bryant, C. L. 2008. Diffusion and mass flow of dissolved carbon dioxide, methane, and dissolved organic carbon in a 7-m deep raised peat bog. *Geochim. Cosmochim*. Acta 72: 2048-2066. Available at: http://www.sciencedirect.com/science/article/pii/S0016703708000835 [Accessed March 19, 2015].

Colombet, J., Robin, A., Lavie, L., Bettarel, Y., Cauchie, H. M., and Sime-Ngando, T. 2007. Virioplankton "pegylation": use of PEG (polyethylene glycol) to concentrate and purify viruses in pelagic ecosystems. *J Microbiol Methods* 71: 212-219. doi:10.1016/j.mimet.2007.08.012.

Comeau, A. M., Hatfull, G. F., Krisch, H. M., Lindell, D., Mann, N. H., and Prangishvili, D. 2008. Exploring the prokaryotic virosphere. *Res. Microbiol*. 159, 306-313. doi:10.1016/j.resmic.2008.05.001.

Danovaro, R., Manini, E., and Dell'Anno, A. 2002. Higher abundance of bacteria than of viruses in deep Mediterranean sediments. *Proc. Natl. Acad. Sci. U. S. A*. 68: 1468-1472.

Dedysh, S. N., Pankratov, T. A., Belova, S. E., Kulichevskaya, I. S., and Liesack, W. 2006. Phylogenetic Analysis and In Situ Identification of Bacteria Community Composition in an Acidic Sphagnum Peat Bog. Appl. Environ. Microbiol. 72: 2110-2117. doi:10.1128/AEM.72.3.2110-2117.2006.

Dedysh, S. N. 2011. Cultivating uncultured bacteria from northern wetlands: knowledge gained and remaining gaps. *Front. Microbiol*. 2: 184.

Dell'Anno, A., Corinaldesi, C., and Danovaro, R. 2015. Virus decomposition provides an important contribution to benthic deep-sea ecosystem functioning. *Proc. Natl. Acad. Sci. U. S. A*. 112 (16): E2014-9. doi:10.1073/pnas.1422234112.

Desnues, C., Rodriguez-Brito, B., Rayhawk, S., Kelley, S., Tran, T., Haynes, M., Liu, H., Furlan, M., Wegley, L., Chau, B., et al. 2008. Biodiversity and biogeography of phages in modern stromatolites and thrombolites. *Nature* 452: 340-343. doi:10.1038/nature06735.

Diemer, G. S., and Stedman, K. M. 2012. A novel virus genome discovered in an extreme environment suggests recombination between unrelated groups of RNA and DNA viruses. *Biol. Direct* 7: 13.

Dray, S., and Dufour, A. B. 2007. The ade4 Package: Implementing the Duality Diagram for Ecologists. *J. Stat. Softw*. 22: 1-20. doi:10.1.1.177.8850.

Duffy, S., Shackelton, L. A., and Holmes, E. C. 2008. Rates of evolutionary change in viruses: patterns and determinants. *Nat. Rev. Genet*. 9: 267-276.

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32: 1792-1797. doi:10.1093/nar/gkh340.

Eswar, N., Eramian, D., Webb, B., Shen, M.-Y., and Sali, A. 2008. Protein structure modeling with MODELLER. *Methods Mol Biol* 426: 145-159. doi:10.1007/978-1-60327-058-8_8.

Farnell-Jackson, E. A., and Ward, A. K. 2003. Seasonal patterns of viruses, bacteria and dissolved organic carbon in a riverine wetland. *Freshw. Biol*. 48: 841-851. doi:10.1046/j.1365-2427.2003.01052.x.

Fierer, N., Nemergut, D., Knight, R., and Craine, J. M. 2010. Changes through time: Integrating microorganisms into the study of succession. *Res. Microbiol*. 161: 635-42. doi:10.1016/j.resmic.2010.06.002.

Francez, A.-J. 1995. Dynamique du carbone et de l'azote chez le Carex rostrata, l'Eriophorum vaginatumet le Calluna vulgarisdans une tourbière à sphaignes des monts du Forez (France). *Can. J. Bot*. 73: 121-129.

Francez, A.-J., and VASANDER, H. 1995. Peat accumulation and peat decomposition after human disturbance in French and finnish mires. *Acta oecologica* 16: 599-608.

Francez, A.-J., and Loiseau, P. 1999. Devenir de l'azote mineral dans une tourbiere a Sphagnum fallax Klinggr. et Carex rostrata Stokes du Massif central (France). *Can. J. Bot*. 77: 1136-1143.

Fuhrman, J. A. 1999. Marine viruses and their biogeochemical and ecological effects. *Nature* 399: 541-548. doi:10.1038/21119.

Fuhrman, J. A., and Schwalbach, M. 2003. Viral influence on aquatic bacterial communities. *Biol. Bull*. 204: 192-195.

Gilbert, D., Amblard, C., Bourdier, G., and Francez, a J. 1998a. Short-term effect of nitrogen enrichment on the microbial communities of a peatland. *Hydrobiologia* 374: 111-119.

Gilbert, D., Amblard, C., Bourdier, G., and Francez, A.-J. 1998b. The Microbial Loop at the Surface of a Peatland:Structure, Function, and Impact of Nutrient Input. *Microb. Ecol.* 35: 83-93. doi:10.1007/s002489900062.

Hopkins, M., Kailasan, S., Cohen, A., Roux, S., Tucker, K. P., Shevenell, A., Agbandje-McKenna, M., and Breitbart, M. 2014. Diversity of environmental single-stranded DNA phages revealed by PCR amplification of the partial major capsid protein. *ISME J.* 8 (10): 2093-103. doi:10.1038/ismej.2014.43.

Huang, Y., Gilna, P., and Li, W. 2009. Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* 25: 1338-1340. doi:10.1093/bioinformatics/btp161.

Huang, Y., Niu, B., Gao, Y., Fu, L., and Li, W. 2010. CD-HIT Suite: A web server for clustering and comparing biological sequences. *Bioinformatics* 26: 680-2. doi:10.1093/bioinformatics/btq003.

Hultman, J., Waldrop, M. P., Mackelprang, R., David, M. M., McFarland, J., Blazewicz, S. J., Harden, J., Turetsky, M. R., McGuire, A. D., Shah, M. B., et al. 2015. Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* 521: 208-212.

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. 2007. MEGAN analysis of metagenomic data. *Genome Res.* 17: 377-386. doi:10.1101/gr.5969107.

Hyatt, D., Chen, G.-L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11: 119. doi:10.1186/1471-2105-11-119.

Jackson, E. F., and Jackson, C. R. 2008. Viruses in wetland ecosystems. *Freshw. Biol*. 53: 1214-1227. doi:10.1111/j.1365-2427.2007.01929.x.

Jobb, G., Haeseler, von, A., and Strimmer, K. 2004. TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol Biol* 4: 18. doi:10.1186/1471-2148-4-18.

Juottonen, H., Galand, P. E., Tuittila, E.-S., Laine, J., Fritze, H., and Yrjälä, K. 2005. Methanogen communities and Bacteria along an ecohydrological gradient in a northern raised bog complex. *Environ. Microbiol*. 7: 1547-1557.

Kim, K.-H., and Bae, J.-W. 2011. Amplification methods bias metagenomic libraries of uncultured single-stranded and double-stranded DNA viruses. *Applied and Environmental Microbiology* 77: 7663-7668. doi:10.1128/AEM.00289-11.

Kim, M.-S., Park, E.-J., Roh, S. W., and Bae, J.-W. 2011. Diversity and abundance of single-stranded DNA viruses in human feces. *Applied and Environmental Microbiology* 77: 8062-8070. doi:10.1128/AEM.06331-11.

Kimura, M., Jia, Z. J., Nakayama, N., and Asakawa, S. 2008. Ecology of viruses in soils: Past, present and future perspectives. *Soil Sci. Plant Nutr*. 54: 1-32. doi:10.1111/j.1747-0765.2007.00197.x.

King, A., Adams, M. J., Lefkowitz, E. J., and Carstens, E. B. 2012. Virus Taxonomy: Ninth Report of the International Commitee on Taxonomy of Viruses. Internation *Union of Microbiological Societies.*

Kirchman, D. L. 2012. Processes in Microbial Ecology. *Oxford University Press.*

Krupovic, M., and Forterre, P. 2011. Microviridae Goes Temperate: Microvirus-Related Proviruses Reside in the Genomes of Bacteroidetes. *PLoS ONE* 6: e19893. doi:10.1371/journal.pone.0019893.

Labonté, J. M., and Suttle, C. A. 2013a. Metagenomic and whole-genome analysis reveals new lineages of gokushoviruses and biogeographic separation in the sea. *Frontiers in Microbiology* 4: 404. doi:10.3389/fmicb.2013.00404.

Labonté, J. M., and Suttle, C. A. 2013b. Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J* 7: 2169-2177. doi:10.1038/ismej.2013.110.

Lamentowicz, M., Bragazza, L., Buttler, A., Jassey, V. E. J., and Mitchell, E. A. D. 2013. Seasonal patterns of testate amoeba diversity, community structure and species-environment relationships in four Sphagnum-dominated peatlands along a 1300m altitudinal gradient in Switzerland. *Soil Biol. Biochem*. 6:, 1-11. doi:10.1016/j.soilbio.2013.08.002.

Le Bras, Y., Roult, A., Monjeaud, C., Bahin, M., and Quénez, O. 2013. Towards a Life Sciences Virtual Research Environment. *e-biogenouest.org*. Available at: https://www.e-biogenouest.org/resources/129/download/jobim_YLeBras_2013.pdf.

Legendre, P., and Legendre, L. 1998. Numerical ecology. Numer. *Ecol. Second English Ed*. 20: 870. doi:10.1021/ic050220j.

López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A., and Alcamí, A. 2009. High diversity of the viral community from an Antarctic lake. *Science* 326: 858-861. doi:10.1126/science.1179287.

Lowe, T. M., and Eddy, S. R. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 25: 955-964. doi:10.1093/nar/25.5.955.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, Buchner, A., Lai, T., Steppi, S., Jobb, G., et al. 2004. ARB: a software environment for sequence data. *Nucleic Acids Res* 32: 1363-1371. doi:10.1093/nar/gkh293.

Lynch, J. M., Benedetti, a., Insam, H., Nuti, M. P., Smalla, K., Torsvik, V., and Nannipieri, P. 2004. Microbial diversity in soil: Ecological theories, the contribution of molecular techniques and the impact of transgenic plants and transgenic microorganisms. *Biol. Fertil. Soils* 40: 363-385. doi:10.1007/s00374-004-0784-9.

Mackelprang, R., Waldrop, M. P., DeAngelis, K. M., David, M. M., Chavarria, K. L., Blazewicz, S. J., Rubin, E. M., and Jansson, J. K. 2011. Metagenomic analysis of a permafrost microbial community reveals a rapid response to thaw. *Nature* 480: 368-371.

Maillet, N., Lemaitre, C., Chikhi, R., Lavenier, D., and Peterlongo, P. 2012. Compareads: comparing huge metagenomic experiments. *BMC Bioinformatics* 13: S10. doi:10.1186/1471-2105-13-S19-S10.

Middelboe, M. 2000. Bacterial Growth Rate and Marine Virus-Host Dynamics. *Microb. Ecol.* 40: 114-124. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11029080 [Accessed January 27, 2015].

Middelboe, M., and Lyck, P. G. 2002. Regeneration of dissolved organic matter by viral lysis in marine microbial communities. *Aquat. Microb. Ecol*. 27: 187-194. doi:10.3354/ame027187.

Mondav, R., Woodcroft, B. J., Kim, E.-H., McCalley, C. K., Hodgkins, S. B., Crill, P. M., Chanton, J., Hurst, G. B., Verberkmoes, N. C., Saleska, S. R., et al. 2014. Discovery of a novel methanogen prevalent in thawing permafrost. *Nat. Commun*. 5: 3212.

Moore, T. R. 1987. Patterns of dissolved organic matter in subarctic peatlands. *Earth Surf. Process. Landforms* 12: 387-397.

Opelt, K., Berg, C., Schönmann, S., Eberl, L., and Berg, G. 2007. High specificity but contrasting biodiversity of Sphagnum-associated bacterial and plant communities in bog ecosystems independent of the geographical region. *ISME J*. 1: 502-516. doi:10.1038/ismej.2007.58.

Pagarete, A., Chow, C., Johannessen, T., Fuhrman, J. A., Thingstad, T. F., and Sandaa, R. A. 2013. Strong seasonality and interannual recurrence in marine myovirus communities. Appl. Environ. *Microbiol*. 79: 6253-6259. doi:10.1128/AEM.01075-13.

Pankratov, T. A., Ivanova, A. O., Dedysh, S. N., and Liesack, W. 2011. Bacterial populations and environmental factors controlling cellulose degradation in an acidic Sphagnum peat. Environ. *Microbiol*. 13: 1800-1814.

Payet, J., and Suttle, C. A. 2013. To kill or not to kill: The balance between lytic and lysogenic viral infection is driven by trophic status. *Limnol. Oceanogr*. 58: 465-474. doi:10.4319/lo.2013.58.2.0465.

Pedulla, M. L., Ford, M. E., Houtz, J. M., Karthikeyan, T., Wadsworth, C., Lewis, J. A., Jacobs-Sera, D., Falbo, J., Gross, J., Pannunzio, N. R., et al. 2003. Origins of highly mosaic mycobacteriophage genomes. *Cell* 113: 171-182.

Peltoniemi, K., Fritze, H., and Laiho, R. 2009. Response of fungal and actinobacterial communities to water-level drawdown in boreal peatland sites. *Soil Biol. Biochem*. 41: 1902-1914. doi:10.1016/j.soilbio.2009.06.018.

Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., and Ferrin, T. E. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25: 1605-1612. doi:10.1002/jcc.20084.

Pradeep Ram, A. S., Rasconi, S., Jobard, M., Palesse, S., Colombet, J., and Sime-Ngando, T. 2011. High Lytic Infection Rates but Low Abundances of Prokaryote Viruses in a Humic Lake (Vassiviere, Massif Central, France). *Appl. Environ. Microbiol*. 77 : 5610-5618.

Quaiser, A., Bodi, X., Dufresne, A., Naquin, D., Francez, A. J., Dheilly, A., Coudouel, S., Pedrot, M., and Vandenkoornhuyse, P. 2014. Unraveling the stratification of an iron-oxidizing microbial mat by metatranscriptomics. *PLoS One* 9. doi:10.1371/journal.pone.0102561.

Quaiser, A., Dufresne, A., Ballaud, F., Roux, S., Zivanovic, Y., Colombet, J., Sime-Ngando, T., and Francez, A.-J. 2015. Diversity and comparative genomics of Microviridae in *Sphagnum*- dominated peatlands. *Front. Microbiol*. 6: 375. doi:10.3389/fmicb.2015.00375.

Quaiser, A., Zivanovic, Y., Moreira, D., and López-García, P. 2011. Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. *ISME J*. 5: 285-304. doi:10.1038/ismej.2010.113.

R Development Core Team 2013. R Development Core Team. *R A Lang. Environ. Stat. Comput*. Available at: http://www.r-project.org/.

Ramette, A. 2007. Multivariate analyses in microbial ecology. *FEMS Microbiol. Ecol*. 62: 142-160.

Rohwer, F., and Thurber, R. V. 2009. Viruses manipulate the marine environment. *Nature* 459 : 207-212. doi:10.1038/nature08060.

Rosario, K., Dayaram, A., Marinov, M., Ware, J., Kraberger, S., Stainton, D., Breitbart, M., and Varsani, A. 2012. Diverse circular ssDNA viruses discovered in dragonflies (Odonata: Epiprocta). *J. Gen. Virol*. 93: 2668-2681. doi:10.1099/vir.0.045948-0.

Roulet, N. T., Lafleur, P. M., Richard, P. J. H., Moore, T. R., Humphreys, E. R., and Bubier, J. 2007. Contemporary carbon balance and late Holocene carbon accumulation in a northern peatland. *Glob. Chang. Biol*. 13: 397-411. doi:10.1111/j.1365-2486.2006.01292.x.

Roux, S., Faubladier, M., Mahul, A., Paulhe, N., Bernard, A., Debroas, D., and Enault, F. 2011. Metavir: a web server dedicated to virome analysis. *Bioinformatics* 27: 3074-3075. doi:10.1093/bioinformatics/btr519.

Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime-Ngando, T., and Debroas, D. 2012a. Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS One* 7. doi:10.1371/journal.pone.0033641.

Roux, S., Krupovic, M., Poulet, A., Debroas, D., and Enault, F. 2012b. Evolution and diversity of the Microviridae viral family through a collection of 81 new complete genomes assembled from virome reads. *PLoS ONE* 7, e40418. doi:10.1371/journal.pone.0040418

Roux, S., Enault, F., Bronner, G., Vaulot, D., Forterre, P., and Krupovic, M. 2013a. Chimeric viruses blur the borders between the major groups of eukaryotic single-stranded DNA viruses. *Nat. Commun*. 4: 2700.

Roux, S., Krupovic, M., Debroas, D., Forterre, P., and Enault, F. 2013b. Assessment of viral community functional potential from viral metagenomes may be hampered by contamination with cellular sequences. *Open Biol.* 3: 130160.

Roux, S., Tournayre, J., Mahul, A., Debroas, D., and Enault, F. 2014. Metavir 2: new tools for viral metagenome comparison and assembled virome analysis. *BMC Bioinformatics* 15: 76. doi:10.1186/1471-2105-15-76.

Roy, A., Kucukural, A., and Zhang, Y. 2010. I-TASSER: a unified platform for automated protein structure and function prediction. *Nat Protoc* 5: 725-738. doi:10.1038/nprot.2010.5.

Rydin, H., and Jeglum, J. 2006. The biology of Peatlands. *Biol. Habitats*: 354. doi:10.1177/004057368303900411.

Sakamoto, M., and Benno, Y. 2006. Reclassification of *Bacteroides distasonis*, *Bacteroides goldsteinii* and *Bacteroides merdae* as *Parabacteroides distasonis* gen. nov., comb. nov., *Parabacteroides goldsteinii* comb. nov. and *Parabacteroides merdae* comb. nov. *Int J Syst Evol Microbiol* 56: 1599-1605. doi:10.1099/ijs.0.64192-0.

Sakowski, E. G., Munsell, E. V., Hyatt, M., Kress, W., Williamson, S. J., Nasko, D. J., Polson, S. W., and Wommack, K. E. 2014. Ribonucleotide reductases reveal novel viral diversity and predict biological and ecological features of unknown marine viruses. *Proc. Natl. Acad. Sci. U. S. A*. 111: 15786-15791.

Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., Lesniewski, R. A., Oakley, B. B., Parks, D. H., Robinson, C. J., et al. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol*. 75: 7537-7541. doi:10.1128/AEM.01541-09.

Serkebaeva, Y. M., Kim, Y., Liesack, W., and Dedysh, S. N. 2013. Pyrosequencing-based assessment of the bacteria diversity in surface and subsurface peat layers of a northern wetland, with focus on poorly studied phyla and candidate divisions. *PLoS One* 8: e63994.

Seshadri, R., Kravitz, S. A., Smarr, L., Gilna, P., and Frazier, M. 2007. CAMERA: a community resource for metagenomics. *PLoS Biol*. 5: e75. doi:10.1371/journal.pbio.0050075.

Sime-Ngando, T. 2014. Environmental bacteriophages: viruses of microbes in aquatic ecosystems. Front. *Microbiol*. 5: 355. doi:10.3389/fmicb.2014.00355.

Smith, R. J., Jeffries, T. C., Roudnew, B., Seymour, J. R., Fitch, A. J., Simons, K. L., Speck, P. G., Newton, K., Brown, M. H., and Mitchell, J. G. 2013. Confined aquifers as viral reservoirs. *Environ. Microbiol. Rep*. 5: 725-730.

Soffer, N., Brandt, M. E., Correa, A. M. S., Smith, T. B., and Thurber, R. V. 2014. Potential role of viruses in white plague coral disease. *ISME J* 8: 271-283. doi:10.1038/ismej.2013.137.

Sommer, U., Adrian, R., De Senerpont Domis, L., Elser, J. J., Gaedke, U., Ibelings, B., Jeppesen, E., Lürling, M., Molinero, J. C., Mooij, W. M., et al. 2012. Beyond the Plankton Ecology Group (PEG) Model: Mechanisms Driving Plankton Succession. *Annu. Rev. Ecol. Evol. Syst*. 43: 429-448.

Sun, C. L., Brauer, S. L., Cadillo-Quiroz, H., Zinder, S. H., and Yavitt, J. B. 2012. Seasonal changes in methanogenesis and methanogenic community in three peatlands, Newyork State. *Front. Microbiol*. 3. doi:10.3389/fmicb.2012.00081.

Suttle, C. A. 2005. Viruses in the sea. *Nature* 437: 356-361. doi:10.1038/nature04160.

Suttle, C. A. 2007. Marine viruses-major players in the global ecosystem. *Nat. Rev. Microbiol*. 5: 801-812. doi:10.1038/nrmicro1750.

Suzuki, R., and Shimodaira, H. 2006. Pvclust: An R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics* 22: 1540-1542. doi:10.1093/bioinformatics/btl117.

Thormann, M. N., Currah, R. S., and Bayley, S. E. 2003. Succession of microfungal assemblages in decomposing peatland plants. *Plant Soil* 250: 323-333.

Tucker, K. P., Parsons, R., Symonds, E. M., and Breitbart, M. 2011. Diversity and distribution of single-stranded DNA phages in the North Atlantic Ocean. *ISME J* 5: 822-830. doi:10.1038/ismej.2010.188.

Turetsky, M. R., Benscoter, B., Page, S., Rein, G., van der Werf, G. R., and Watts, A. 2015. Global vulnerability of peatlands to fire and carbon loss. *Nat. Geosci*. 8: 11-14. Available at: http://www.nature.com.gate1.inist.fr/ngeo/journal/v8/n1/full/ngeo2325.html.

Tveit, A., Schwacke, R., Svenning, M. M., and Urich, T. 2012. Organic carbon transformations in high-Arctic peat soils: key functions and microorganisms. *ISME J*. 7: 299-311. doi:10.1038/ismej.2012.99.

Tveit, A., Urich, T., and Svenning, M. M. 2014. Metatranscriptomic analysis of Arctic peat soil microbiota. Appl. *Environ. Microbiol*. doi:10.1128/AEM.01030-14.

Vandenkoornhuyse, P., Dufresne, A., Quaiser, A., Gouesbet, G., Binet, F., Francez, A. J., Mahé, S., Bormans, M., Lagadeuc, Y., and Couée, I. 2010. Integration of molecular functions at the ecosystemic level: Breakthroughs and future goals of environmental genomics and post-genomics. *Ecol. Lett*. 13: 776-791. doi:10.1111/j.1461-0248.2010.01464.x.

Waddington, J. M., and Roulet, N. T. 1997. Groundwater flow and dissolved carbon movement in a boreal peatland. *J. Hydrol*. 191: 122-138. doi:10.1016/S0022-1694(96)03075-2.

Wiederstein, M., and Sippl, M. J. 2007. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. *Nucleic Acids Res* 35: W407-10. doi:10.1093/nar/gkm290.

Williamson, K. E., Radosevich, M., and Wommack, K. E. 2005. Abundance and diversity of viruses in six Delaware soils. *Appl. Environ. Microbiol*. 71: 3119-3125. doi:10.1128/AEM.71.6.3119-3125.2005.

Wommack, K. E., and Colwell, R. R. 2000. Virioplankton: viruses in aquatic ecosystems. *Microbiol. Mol. Biol*. Rev. 64: 69-114.

Wommack, K. E., Nasko, D. J., Chopyk, J., and Sakowski, E. G. 2015. Counts and sequences, observations that continue to change our understanding of viruses in nature. *J. Microbiol*. 53: 181-92.

Yoshida, M., Takaki, Y., Eitoku, M., Nunoura, T., and Takai, K. 2013. Metagenomic analysis of viral communities in (hado)pelagic sediments. *PLoS ONE* 8: e57271. doi:10.1371/journal.pone.0057271.

Yu, Z. C. 2012. Northern peatland carbon stocks and dynamics: a review. *Biogeosciences* 9: 4071-4085.

Zablocki, O., van Zyl, L., Adriaenssens, E. M., Rubagotti, E., Tuffin, M., Cary, S. C., Cowan, D., Zyl, L. Van, Adriaenssens, E. M., Rubagotti, E., et al. 2014. High-Level Diversity of Tailed Phages, Eukaryote-Associated Viruses, and Virophage-Like Elements in the Metaviromes of Antarctic Soils. *Appl. Environ. Microbiol*. 80 : 6888-6897. doi:10.1128/AEM.01525-14.

Zivanovic, Y., Confalonieri, F., Ponchon, L., Lurz, R., Chami, M., Flayhan, A., Renouard, M., Huet, A., Decottignies, P., Davidson, A. R., et al. 2014. Insights into bacteriophage T5 structure from analysis of its morphogenesis genes and protein components. *J. Virol*. 88: 1162-1174. doi:10.1128/JVI.02262-13.

.

Chapitre 2:

Metagenomic and metatranscriptomic analysis of microbial communities in the acrotelm and the mesotelm of a *Sphagnum*-dominated peatland

Chapitre 2: Metagenomic and metatranscriptomic analysis of microbial communities in the acrotelm and the mesotelm of a *Sphagnum*-dominated peatland

Résumé du Chapitre 2

La plupart des études qui utilisent les nouvelles technologies de séquençages ont porté sur la description de la structure de la diversité des séquences du gène d'ARNr, qui sert de marqueur universel. La structure des communautés de bactéries a généralement été étudiée selon deux grands facteurs : i) l'effet de la profondeur, lié au changement de la matière organique et à l'anoxie, ou ii) l'effet de l'espèce de Sphaigne, puisque chaque espèce de Sphaigne se développe dans des conditions de pH et d'apports en nutriments différent, deux facteurs identifiés comme des éléments majeurs structurant les communautés de procaryotes.

Il semble donc avéré que chaque espèce de Sphaigne abrite une communauté de procaryote spécifique, mais également que la structure des communautés change avec la profondeur. Cependant, la diversité fonctionnelle des communautés de procaryotes des Sphaignes est encore plus mal connue que la diversité taxonomique, et semble présenter des spécificités lui permettant d'évoluer dans l'écosystème constitué par les Sphaignes. Afin de mieux comprendre les interactions procaryotes/Sphaignes/tourbière, nous avons cherché à déterminer par quel facteur (type de Sphaigne ou profondeur) était structurée la diversité active et l'expression de gènes fonctionnels (métatranscriptomique) à deux profondeurs très proches de (l'acrotelme et le mésotelme)du fen (stade jeune, minérotrophique, dominé par *Sphagnum fallax*) et du bog (stade mature, ombrotrophe et plus acide dominé par *Sphagnum magellanicum*). Nous avons ensuite comparé ces résultats au potentiel génétique (diversité taxonomique et fonctionnelle par métagénomique) pour déterminer si ce facteur était associé à une adaptation taxonomique et fonctionnelle des communautés microbiennes.

L'analyse de 12 pools d'ARN de tourbe (2 stades x 2 profondeurs x 3 réplicats) (métatranscriptomes) montre que la diversité taxonomique active (étudiée avec les séquences d'ARNr) présente une forte structuration par la profondeur, confirmée par l'analyse de l'abondance des ARNr d'organismes avec une niche identifiée : avec plus d'organismes aérobes en surface, et des organismes anaérobes dans le mésotelme. La diversité fonctionnelle (analyse taxonomique des séquences des ARNm ou ARN codant des séquences protéiques) présente aussi un patron surface/profondeur, avec plus de

fonctions associées à l'activité autotrophe en surface et plus de fonctions liées à l'activité hétérotrophe en profondeur.

L'analyse des 12 pools d'ADN environnementaux (métagénomes) (même plan d'échantillonnage) présente un patron différent. La diversité taxonomique (séquences de gène d'ARNr) et fonctionnelle (analyse taxonomique des séquences de gènes codant des protéines) est structurée par le stade dynamique.

La comparaison de la diversité fonctionnelle des métagénomes et des métatranscriptomes indique des structures de communautés très proches dans les métagénomes, malgré la distinction fen/bog, et des structures beaucoup plus variables dans les métatranscriptomes. La variation de structure des communautés entre les métagénomes et les métatranscriptomes semble très fortement associée au facteur surface/profondeur. Ce patron suggère que des fonctions proches sont présentes entre les deux stades, avec des fonctions portées par des organismes différents entre les deux stades et une sélection similaire de l'expression de ces fonctions le long du gradient de profondeur, expliquant la spécificité des communautés en fonction de l'espèce de Sphaigne, et la différence de structure entre les deux couches étudiées.

Ce patron de variation de structure des communautés, associé au fait que certains génomes connus soient très abondamment représentés dans tous les échantillons (métagénomes et métatranscriptomes) suggère donc l'existence d'un core-microbiome fonctionnel des Sphaignes, permettant aux communautés de procaryotes de vivre en interactions avec l'espèce végétale dominante et les conditions environnementales qui lui sont associées.

Metagenomic and metatranscriptomic analysis of microbial communities in the acrotelm and the mesotelm of a *Sphagnum*-dominated peatland

Introduction

Peatlands are key components of the global carbon cycle and represent up to one third of the soil carbon stock as a consequence from the imbalance between the fixation of atmospheric carbon through photosynthesis and organic carbon mineralization through heterotrophic respiration. Due to the high water-content, abundant precipitations and cold temperature that favor the development of peatlands, *Sphagnum* mosses constitute engineer species that can colonize and grow in these ecosystems, in which they maintain nutrient-poor and acidic conditions through their high water-retention rates and high cation exchange capacity. Most peatlands develop from the minerotrophic mesotrophic fen, represented by species such as *Sphagnum fallax* and *S. balticum*, which is later colonized by species such as *Sphagnum magellanicum* and *S*. *fuscum* that constitute the ombrotrophic oligotrophic and more acidic bog (Rydin and Jeglum, 2006). Fen and bog *Sphagnum* species have differential ecological needs, depending on their ability to grow under more or less acidic conditions, or different nutrients and water supplies. The high-water content of the peatlands is also associated to a vertical gradient of oxic/anoxic conditions and carbon susbstrate degradation, from the more dessicated and oxic acrotelm to the anoxic catotelm where peat accumulates most and which harbor among others the activity of methanogenic Archaea (Artz, 2009).

Prokaryotic communities are of crucial importance in these ecosystems, as they potentially contribute to both plant and *Sphagnum* growth through nitrogen fixation (Bragina et al., 2013; Vile et al., 2014) and to the degradation of the complex organic compounds produced by the mosses (Pankratov et al., 2011; Lin et al., 2014). Due to the complex (allelopathic) and tenuous relationships between the biology of *Sphagnum* species and their microbial communities, the functioning of the *Sphagnum*-microbial community interactions (Bragina et al., 2014) is referred to as a micro-ecosystem called the “Sphagnosphere” (Jassey, 2011).

Influence of both the dynamics stage and depth has been found on the structure of peatland microbial communities, but it is yet uncertain which of the two factors influences it most. Analysis of rRNA gene diversity revealed a higher influence of depth related conditions on the structure of the prokaryotic and fungal communities (Dedysh et al., 2006; Morales et al., 2006; Lin et al., 2012; Preston et al., 2012; Serkebaeva et al., 2013) as a result of the quality of the available carbon

substrates and the N and P limitation at depth. On the other hand, through the analysis of 16S rRNA gene, prokaryotic communities appeared highly structured by the *Sphagnum*-species (Opelt et al., 2007; Bragina et al., 2012; Preston et al., 2012), even at large geographical scale. This *Sphagnum* specificity was further observed with FISH (Fluorescence *In-Situ* Hybridization) which revealed the transmission of a high proportion of the prokaryotic community throughout the *Sphagnum* life-cycle (Bragina et al., 2012). In parallel, analysis of the diversity of *nifH*, *pmoA*, and *mcr* genes (respectively nitrogenase for N-fixation, methane oxidation and methanogenesis) between the two dynamics stages suggested functional redundancies (Basiliko et al., 2013), and the fact the different *Sphagnum* species harbored similar functions (Juottonen et al., 2005; Bragina et al., 2013a).

While these studies contributed to give a rather good picture of who is present in the peatlands prokaryotic communities, metagenomic and metatranscriptomic have still been poorly applied to peatland prokaryotic communities. They more recently proved to be powerful tools to unveil the taxonomical and functional community structure of the prokaryotic communities associated to the degradation of organic carbon with depth (Yergeau et al., 2010; Tveit et al., 2013; Lin et al., 2014). Metatranscriptomic also emphasized the increasing expressions of methanogenic archaea with depth along two fen cores (Tveit et al., 2013). Using a microscom experiment combined with metatranscriptomic, activity of these methanogenes was shown to increase with a rise of temperature in relation with a shift in the structure of the active communities involved in the degradation of organic matter and in methanogenesis (Tveit et al., 2015). Interestingly, in the meantime, Hultman et al. (2015) combined metagenomic, metatranscriptomic and metabolic profiling that led to similar conclusions, but which emphasized high differences between the "present" organisms (metagenomics) and the active communities (metatranscriptomics). Finally, a functional comparison of the *Sphagnum magellanicum* microbiome with other vascular plants microbiome through shotgun metagenomics, suggested with only one sample that the *Sphagnum* microbiome present a specific set of functionalities that distinguishes it from other plant associated-prokaryotic communities and which is key to their functioning within peatland ecosystems (Bragina et al., 2014). This functional approach was inserted into a core of studies that argue for the vertical transmission of the prokaryotic communities throughout the *Sphagnum* life-cycle and for long-term association between the Sphagnum species and their specific prokaryotic communities over large geographical scales (Bragina et al., 2012; Bragina et al., 2013b). Gaps remain as to (i) the importance of microbial-*Sphagnum* interactions that suggest the selection of a *Sphagnum* species specific functional microbiome at the peatland surface, (ii) the necessity for the prokaryotic community to adapt functionally to deeper environmental conditions with changing available carbon substrates.

Our aim was therefore to determine which of the *Sphagnum* species (dynamic stages) or of the depth associated environmental conditions (acrotelm *vs* mesotelm) selected taxonomical and functional genes expression in two close layers of the surface of the fen (*Sphagnum fallax*) and the bog (*Sphagnum magellanicum*) of a *Sphagnum*-dominated peatland. In a second time, we aimed to determine whether the genes expression was a result from a different functional background adapted to the *Sphagnum* species or selected along a depth gradient.

Expressed taxonomical and functional diversity of the prokaryotic community was analyzed through shotgun metatranscriptomic to triplicate samples from the acrotelm (here after called surface) and the mesotelm (depth) from the fen and the bog. Genetic background between dynamics stages and these two close layers was analyzed in triplicate shotgun metagenomes.

We hypothesized an influence of the *Sphagnum* species on the microbial community structure (metagenome) and gene expression (metatranscriptomes) and an influence of depth related environmental conditions from the deeper layer.

Material and methods

Sampling site

Peat samples for the production of the metagenomes and the metatranscriptomes were collected at Les Pradeaux (3°55E; 45°32N), a *Sphagnum*-dominated peatland situated in the French Massif Central at an altitude of 1 350 m. This peatland present two main profiles or dynamic stages that correspond to the evolution of the ecosystem and that are colonized by different *Sphagnum* species and their associated vegetation (Figure 1).

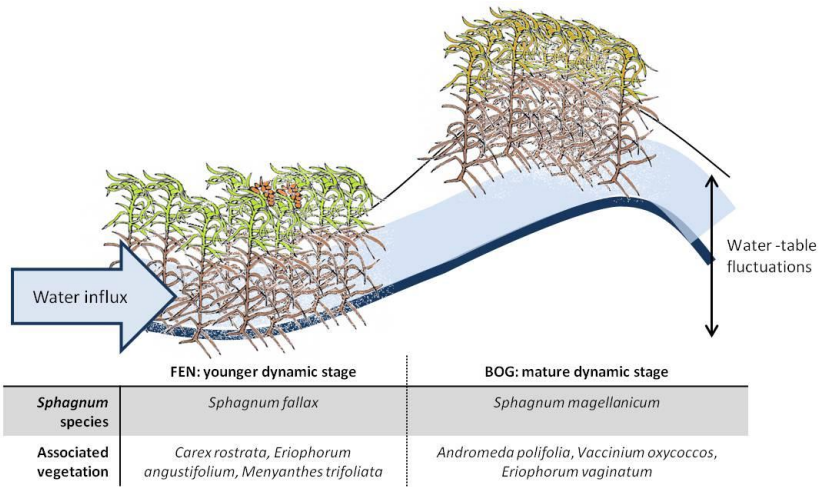


Figure 1 Sketch illustrating the differences between the fen and the bog of a *Sphagnum*-dominated peatland

Experimental design

Peat was sampled in June 2011 (metagenomes production) and September 2012 (metatranscriptomes). For each field campaign, six sampling points were selected, three in the fen and three in the bog. At each sampling point, a peat core (about 20 x 20 x 30 cm) was extracted. The active "green" capitula layer was removed, and two segments were separated into a "surface" segment (from 5 to 10 cm under the active capitula layer) and a depth segment (10 to 15 cm under the active layer) (Figure 2).

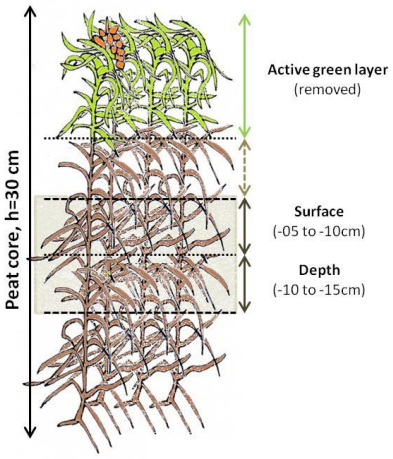


Figure 2 Sketch representing the sampling of a peat core

In each segment, a piece of peat was collected from the middle of the segment and transported in a sterile sampling tube (50mL) at 4°C until return to the laboratory. Samples were flash frozen in liquid nitrogen, and kept at -80°C until processing. In total, this represents 12 peat samples for the production of the metagenomes (2 dynamic stages x 2 depths x 3 replicates), and 12 peat samples for the production of the metatranscriptomes (2 dynamic stages x 2 depths x 3 replicates).

Metagenomes production

The 12 samples underwent DNA extraction, pyrosequencing as well as size and quality trimming (Table 1) as described previously (Quaiser et al., 2014). In short, 2g of peat matrix were lysed in 15ml of extraction buffer containing 4% cetyltrimethylam- monium bromide (CTAB), 0.5% polyvinylpyrrolidone (PVP, Sigma-Aldrich), 0.7 M NaCl, 100 mM potassium phosphate (pH 6.8), 20 mM EDTA (pH 8.0), 1% beta-mercaptoethanol, 1 M guanidin thiocyanate and incubated at 65°C for 30 min. Homogenization was obtained by vigorous vortexing every 5 min during 1min in the presence of glass beads. One volume of chloroform-isoamylalcohol (24:1) was added, vortexed for 1 min and incubated at room temperature for 5 min. The samples were centrifuged at 4000 g for 15 min at 4°C

and the aqueous phases were transferred to new tubes. Binding conditions for silica-based DNA extraction were adjusted, applied on Nucleo Spin RNA II kit columns and subsequent purification was performed following the instructions of the manufacturer (Macherey-Nagel). DNA was nebulized to obtain fragments of about 700bp. The DNA was purified with Agencourt AMPur XP magnetic beads (Beckman-Coulter). DNA fragmentation quality was checked with the High sensitivity DNA kit on a 2100 Bioanalyzer (Agilent). Subsequent library construction and pyrosequencing was performed in technical duplicates on a GS FLX system (454/Roche) at the ‘‘Functional and Environmental Genomics’’ platform (OSUR, Rennes, France). Roche/454 filtering tools and stringent filters developed locally were used to ensure the highest possible sequence quality and to suppress artificial replicates of sequences as well as sequences smaller than 250 bp.

Metatranscriptomes production

The 12 frozen peat samples for the production of the metatranscriptomes were separately and rapidly crushed using a sterilized and frozen bowl and pestle. RNA extraction was essentially performed as DNA extraction with additional DNAse treatments followed by RNA fragmentation. In short, nucleic acids applied on Nucleo Spin RNA II kit columns were purified and undergone a DNase treatment following the instructions of the manufacturer (Macherey-Nagel). Fragmentation of total RNA was performed as indicated in the cDNA Rapid Library Preparation Method Manual (454/Roche) by applying 0.1 volumes of the ZnCl2 fragmentation solution (90 mM ZnCl2; 100 mM Tris-HCl, pH 7.0) for 85 s at 70uC. The absence of remaining DNA was confirmed by PCR using pmoA gene specific primers (A189F-mb661R) and 16S rRNA gene primers (U27F-U1492R). Double-stranded cDNA generation was performed with the Universal RiboClone cDNA Synthesis System (Promega). The cDNA was purified with Agencourt AMPur XP magnetic beads (Beckman-Coulter). RNA and cDNA quality was checked with the RNA Pico 6000 Pico kit or the High sensitivity DNA kit on a 2100 Bioanalyzer (Agilent). Subsequent library construction and pyrosequencing was performed in technical duplicates on a GS FLX system (454/Roche) at the ‘‘Functional and Environmental Genomics’’ platform (OSUR, Rennes, France). Roche/454 filtering tools and stringent filters developed locally were used to ensure the highest possible sequence quality and to suppress artificial replicates of sequences as well as sequences smaller than 250 bp.

Metagenomes and metatranscriptomes analysis

After quality trimming, metagenome and metatranscriptome sequences have been uploaded on the MG-RAST server (Meyer et al., 2008), from which different outputs of the pipeline were used for further analysis. For each metagenome and metatranscriptome, estimated alpha-diversity, relative

abundance of rRNA sequences and assignments to the three main kingdoms (Archaea, Bacteria and Eukaryotes) were retrieved.

Even if questions were addressed in two different steps, the same analysis were conducted first for the metatranscriptomes (to determine the factor that selects genes expression), and then for the metagenomes (to look for a genetical background).

To determine the diversity carried in the metatranscriptomes and metagenomes, the rRNA sequences, and rRNA gene sequences from the metatranscriptomes and metagenomes have been retrieved using hmm_rRNA3.py (Huang et al., 2009), and compared by blastx against SILVA (Quast et al., 2013). rRNA gene sequences correspond to two different subunits: the small and the large (respectively SSU and LSU), that differ between prokaryotes (subunit sizes: SSU: 16S; LSU: 23S), and Eukaryotes (SSU: 18S, LSU: 28S). Taxonomical assignations/affiliations of the 4 different kinds of subunits were analyzed using MEGAN (Huson et al., 2007) for each sample. Relative abundances of the different prokaryotic phyla and eukaryotic classes were used to build a contingency matrix. For the metatranscriptomes rRNA sequences, in order to determine which factor(s) (dynamic stage/depth/type of subunit) determined the variation in the relative abundances of the different phyla/classes, a principal component analysis (PCA) was performed on the contingency matrix using FactomineR package (Husson et al., 2015), with the different factors (subunit type, dynamic stage and depth) added as qualitative information. For each component, FactoMineR indicates which quantitative and qualitative variables are correlated with the component, i.e. which variables are associated with the variance carried by each axis.

Taxonomical and functional assignations of the protein coding sequences were obtained by BLAT analysis against different databases (M5NR, Genbank…) on the MG-RAST server. Taxonomical assignations of the protein coding sequences at the phyla and class level were retrieved independently for each sample of the metatranscriptomes and metagenomes in order to study the relative abundance of the different prokaryotic phyla, and of the different classes of Proteobacteria as these classes are highly represented in soils microbiomes. Relative abundances of the prokaryotic phyla were used to build a contingency matrix. From this matrix, community structure was compared between pair-wise samples using Bray-Curtis index and Euclidean distance. These distances were used to perform hierarchical clustering for each type of dataset (metatranscriptome and metagenomes). The difference of structure of the communities based on both distances was then represented and analyzed in a two dimensions space by Non-Metric-Dimensional-Scaling (NMDS) using R vegan package (Oksanen et al., 2015). Ellipses were drawn to represent the space potentially occupied by 95 % of the replicates depending on the different characteristics associated to the

samples (metagenome *vs* metatranscriptome, fen *vs* bog, suface *vs* bog). Ellipses that are not superimposed indicate communities that can be distinguished from one another.

In an attempt to link the functional potential of the prokaryotic communities to the expressed functions (mRNA) in the metatranscriptomes, SEED assignations (level 1) were retrieved for each metatranscriptome. The relative abundance of each subsystem (based on the abundance of sequences assigned to SEED for each sample) was compared between fen and bog, and between surface and depth of each dynamic stage.

Using the functional assignations of the sequences against M5NR (evalue 10^{-5}; minimum identity 60%; minimum alignment 15), the 50 most recruited genomes were retrieved for each metagenome and metatranscriptome sample. To facilitate the analysis, genomes were ranked according to the abundance of sequences that were recruited (most recruited genome = rank 1, least recruited = rank 50) and only the 20 genomes against which the highest amounts of sequences were recruited for the total of all the samples were kept (Table 3).

Results

Using shotgun metagenomics and metatranscriptomics of triplicate samples from two close layers of the acrotelm of the fen and the bog from a *Sphagnum*-dominated peatland, our aim was to determine what factor (dynamic stage or depth) structured most the active microbial communities (metatranscriptomes) and whether this structure was associated to a different genetic background (metagenomes) associated with this factor.

Metagenomes and metatranscriptomes

After quality trimming, we obtained 893 227 metagenome sequences (Table 1) and 395 575 metatranscriptome sequences (Table 2).

Table 1 Main properties of the 12 peat metagenomes

Type	METAGENOMES											
Dynamic stage	Fen						Bog					
Depth	Surface			Depth			Surface			Depth		
Peat-column	1	2	3	1	2	3	1	2	3	1	2	3
Sample ID	pFen1_S1	pFen1_S2	pFen1_S3	pFen1_D1	pFen1_D2	pFen1_D3	pBog1_S1	pBog1_S2	pBog1_S3	pBog1_D1	pBog1_D2	pBog1_D3
MG-Rast ID	4481337.3	4481335.3	4481334.3	4481333.3	4481331.3	4481332.3	4481342.3	4481340.3	4481341.3	4481339.3	4481338.3	4481336.3
N° sequences	93863	204336	9963	176881	11987	333485	2091	19844	25841	9543	3851	1542
% rRNA	1.8	0.8	7.1	1.6	15.6	0.8	2.2	4.8	0.8	11.1	2.2	2.3
% assigned	66.7	69.4	68.1	69.7	67.7	68.5	59.3	63.3	65	65.5	67.4	66
% unknown	31.6	29.9	24.8	28.7	16.7	30.8	38.4	31.9	33.9	23.4	30.4	31.7
mean length	454	455	452	455	452	455	452	452	455	456	454	450
% GC	60	62	61	61	61	62	59	58	61	60	61	61
% Archea	0.9	0.8	0.9	0.9	0.9	1.2	0.6	1	0.9	0.9	0.9	1.1
% Prokaryotes	95.6	96.9	94.1	96.2	92.2	96.5	91.6	90.9	94.6	91.8	94.2	93
% Eukaryotes	3.3	2.1	4.8	2.7	6.8	2.1	7.3	7.9	4.4	7.1	4.7	5.8
Alpha diversity	651.682	601.113	666.081	667.914	739.308	591.758	516.51	747.04	655.424	777.645	613.376	468.465

Table 2 Main Properties of the 12 peat metatranscriptomes

Type	METATRANSCRIPTOMES											
Dynamic stage	Fen						Bog					
Depth	Surface			Depth			Surface			Depth		
Peat-column	1	2	3	1	2	3	1	2	3	1	2	3
Sample ID	pFen2_S1	pFen2_S2	pFen2_S3	pFen2_D1	pFen2_D2	pFen2_D3	pBog2_S1	pBog2_S2	pBog2_S3	pBog2_D1	pBog2_D2	pBog2_D3
MG-Rast ID	4559702.3	4559703.3	4559704.3	4559699.3	4559700.3	4559701.3	4559708.3	455979.3	4559710.3	4559705.3	4559706.3	4559707.3
N° sequences	37100	31980	35896	31609	45402	36138	33333	30454	4676	33366	39423	36198
% SSU	6	13	21	5	6	20	11	9	9	12	5	15
% LSU	7	18	36	6	6	29	17	13	12	17	6	22
% total rRNA	13	31	57	11	12	49	28	22	21	29	11	37
% assigned	42	44.9	42	41.7	45.4	40.1	43	40.5	40.9	46.6	34.7	46.2
% unknown	46.9	28.1	3.7	46.7	41	15.2	32.1	40.7	42.6	28.9	51.9	21.4
mean length	287	281	271	309	296	272	278	280	277	279	266	268
% GC	57	53	48	57	57	50	55	56	56	57	53	52
% Archea	0.7	0.6	0.2	2.9	3	0.8	0.6	0.6	0.6	1.9	5.4	1.2
% Prokaryotes	86.6	75.2	57	91.5	93.7	68.6	77.5	82.7	84.1	90.5	90	74.4
% Eukaryotes	12.5	24	42.6	5.5	3.1	30.3	21.3	16.5	14.9	7.4	4	24.1
Alpha diversity	734.925	740.443	626.44	572.634	481.416	568.009	875.95	750.577	605.644	513.483	480.169	533.83
% SEED	28	19	8	28	30	10	19	21	24	24	22	15

Metagenomic sequences per sample ranged from 1542 to 333 485 while metatranscriptomics sequences per sample ranged from 30 454 to 45 402 (one exception with 4676 sequences) and are more evenly distributed across dynamic stage and depth. On average 66 ± 3 % and 42 ± 3 % of respectively the metagenomes and metatranscriptomes sequences matched against reference databases (M5NR, SEED or Silva rRNA). There are on average more rRNA sequences in the metatranscriptomes than in the metagenomes (Mann-Whitney test, N=24, p=1.0e-4), as the proportion of rRNA gene sequences in the metagenomes ranges from 0.8 to 15.6 %, and from 11 to 49 % in the metatranscriptomes.

What factor (dynamic stage or depth) selects most the "activity" of the prokaryotic community?

Taxonomical diversity of the active prokaryotic communities observed through the analysis of transcribed rRNA sequences

rRNA sequences are highly conserved and their diversity is more represented and detailed in the databases than this of the protein-coding sequences. Therefore, rRNA gene sequences serve as standard marker to investigate the taxonomical diversity of the communities or the phylogeny of microorganisms. In the metatranscriptomes, these sequences are more abundant than in the metagenomes (Table 1 and Table 2) and can be used as a proxy to analyze the diversity of the active community. In order to gain a better view of how dynamic stage and depth may structure the taxonomical diversity of the microbial communities of peatlands, we studied the assignations of the small and large subunits rRNA genes sequences (SSU and LSU) carried for the three kingdoms in the metatranscriptomes. These taxonomical assignations were obtained by BLASTx (Altschul et al., 1997) against Silva (Quast et al., 2013) and analyzed using MEGAN (Huson et al., 2007) (Figure 3).

The communities appear dominated by the Alphaproteobacteria (15.23 ± 3.15 % (average for all subunits, dynamics stage and depth)), the Verrucomicrobia (12.60 ± 5.80 %), and the Acidobacteria (11.25 ± 6.20 %). The surface samples contain more expressed rRNA from Eukaryotes (Figure 3). However while the same main phyla are observed through the SSU and the LSU, their relative abundances vary due to the relative incompleteness of LSU rRNA databases (arb-silva.de).

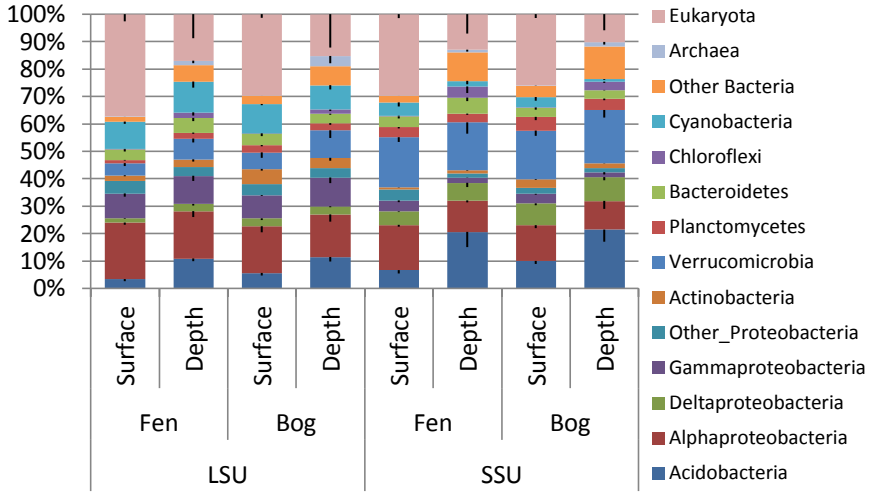


Figure 3 Average proportions of the different phyla depending on the assignation of metatranscriptomics sequences coding for the large and small subunits (LSU and SSU) of the rRNA gene sequences. Bars represent standard deviation and represented for the negative values (N=3).

In order to detect which factors (dynamic stage or depth) were associated with the variation in the abundance of the different active phyla, a principal component analysis was performed based on the taxonomical assignations of the rRNA sequences using FactoMineR (Husson et al., 2015). The two first components carried respectively 52.13 % and 20.48 % of the variance and were associated to the difference between surface and depth communities depending on the analyzed subunit, with the LSU gathered on the left hand side of the samples projection (Figure 4B), and the SSU on the right hand side. Along the first component, Eukaryota, Cyanobacteria and Alphaproteobacteria presented higher abundances in surface and were less abundant when studied with the SSU. To the contrary, Chloroflexi, Acidobacteria, Deltaproteobacteria and Verrucomicrobia were more highly represented at depth. Archaea were more abundant at depth (second component). Finally, the third component was associated to phyla which abundance varied between fen and bog samples, with more Actinobacteria in the bog, and more Bacteroidetes in the fen. This suggests that depth related environmental conditions structures more strongly the active communities than the dynamic stage (Figure 4).

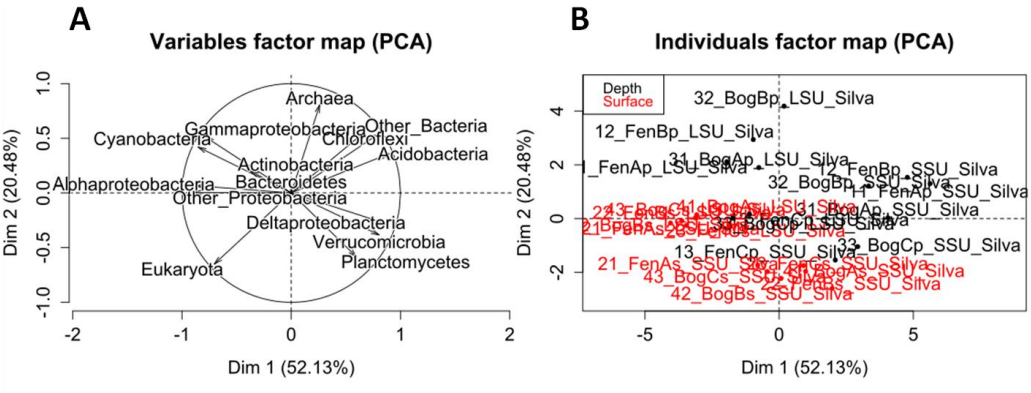


Figure 4 Variable correlation circle (A) and samples projection on the two first axis of the PCA (B) depending on the variation of the abundance of the phyla obtained through the assignations of SSU and LSU sequences against SILVA and on the sampled depth (Depth: Black, Surface:Red). Each sample is represented by two points, one for each subunit.

Spatial pattern of transcribed rRNA sequences representing ecologically differentiable groups

To test the reliability of depth-related environmental conditions on the expression of the rRNA sequences by the microbial community, we analyzed the relative proportions of rRNA sequences that matched to organisms with identified niches or functions (Figure 5).

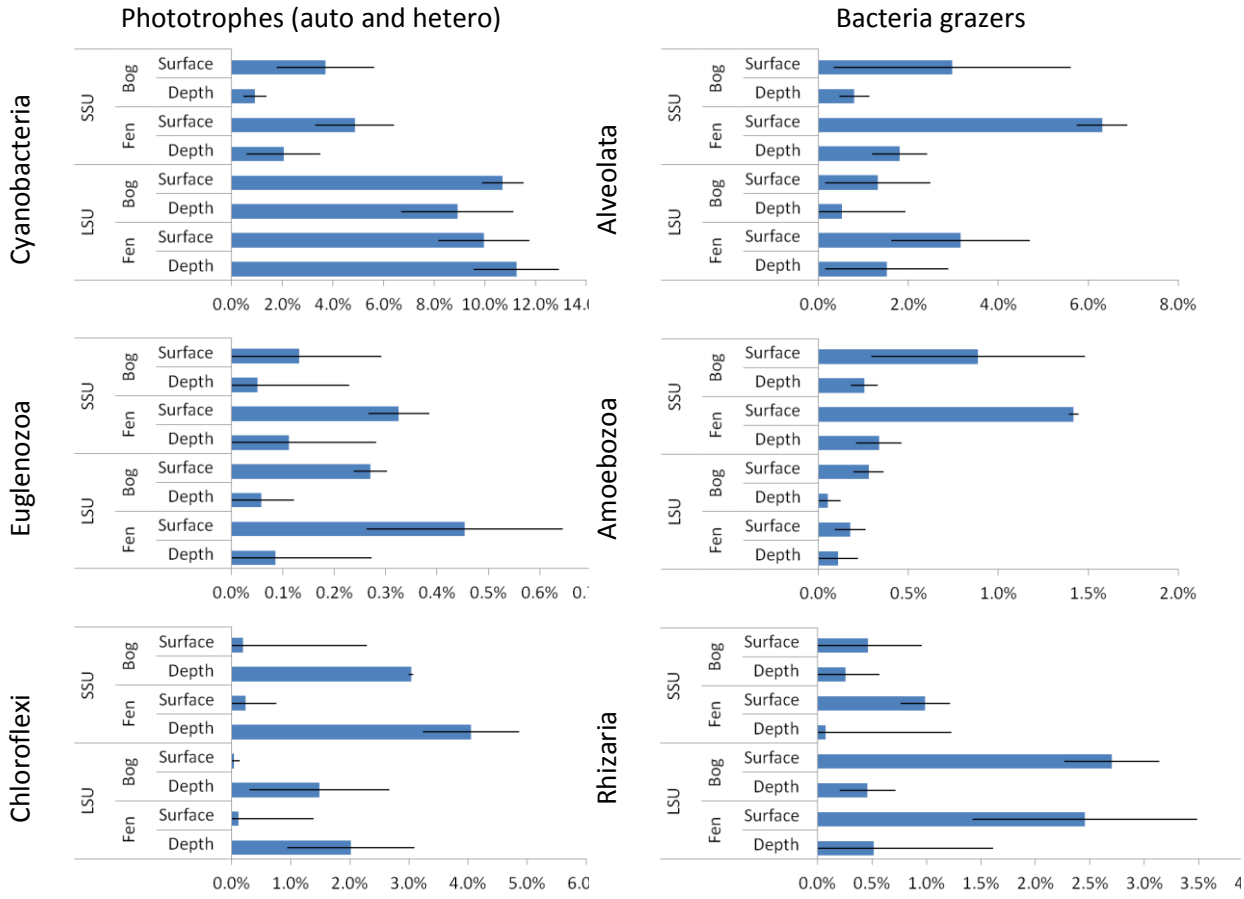


Figure 5 Proportion of sequences of rRNA gene sequences assigned to phyla/classes with identified functions (photosynthesis, bacterial grazing) and niches (aeroby/anaeroby) in the fen and in the bog, and at the surface and the depth. LSU: Large subunit; SSU: Small subunit. Bars represent standard deviation (N=3).

We put emphasis on photosynthetic organisms, which were expected to be more abundant at the surface as they need light for photosynthesis, and distinguished between aerobic and anaerobic kind of photosynthetic organisms. We also analyzed eukaryotic groups known to be bacterial grazers, as they are obligate aerobes, and are involved in the microbial loop. Aerobic primary producers were mainly abundant in the surface of the peat (Cyanobacteria, Euglenozoa), while anaerobic photosynthetic organisms (Chloroflexi) were more abundant at depth (Figure 5). The three classes of bacterial grazers were more abundant at the surface. The spatial pattern of the transcribed rRNA diversity presents a reliable depth related pattern.

Taxonomical assignations of the protein-coding sequences

In order to link the taxonomical and functional diversity, and to detect whether a similar surface-depth pattern selected the expression of functions (protein-coding genes) in the metatranscriptomes, the protein-coding sequences were compared to M5NR and assigned taxonomically and functionally using MG-RAST (Meyer et al., 2008). The taxonomic affiliations at the phyla level were retrieved for each sample, and the relative abundance of each phyla was calculated to build a contingency matrix and to perform a hierarchical clustering based on community structure (Bray-Curtis index) between each metatranscriptome.

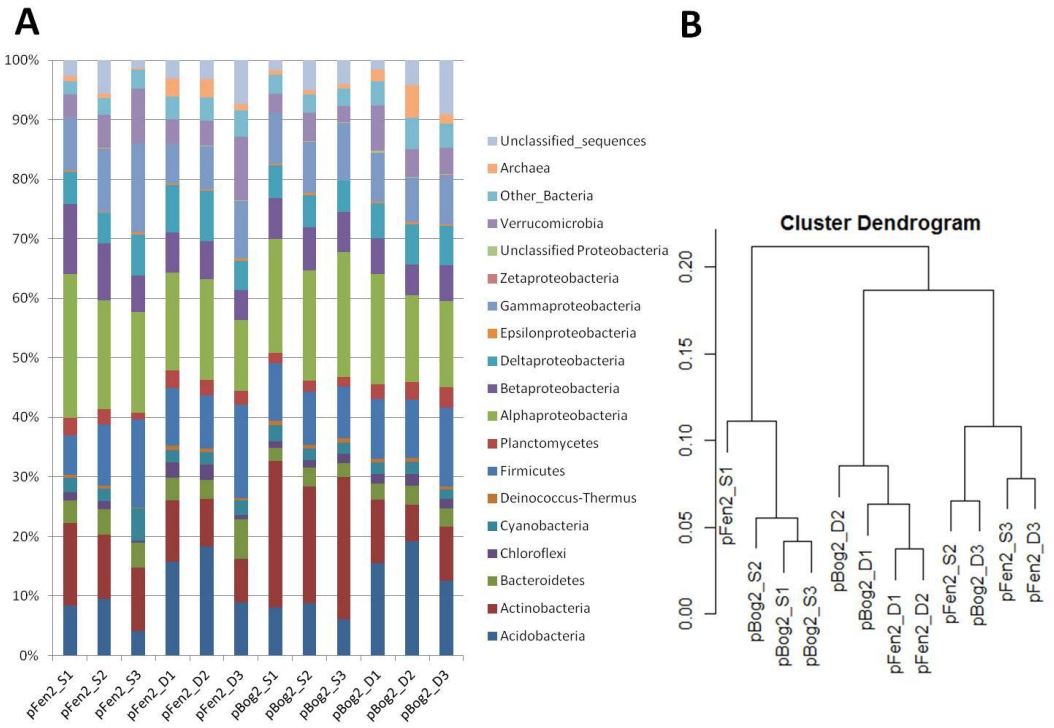


Figure 6 Relative abundance of prokaryotic phyla in the 12 metatranscriptomes depending on the taxonomical assignation of protein-coding sequences at the phyla level (A), and hierarchical clustering of the communities depending on their structure (Bray-Curtis distance) (B). Taxonomical assignations of the protein-coding sequences were obtained using MG-Rast pipeline.

The protein-coding sequences were mainly assigned to Alphaproteobacteria (14.61%) and Actinobacteria (10.65%), Acidobacteria (9.73%) and Firmicutes (8.51%) (Figure 6A). The expressed prokaryotic protein-coding sequences appear mainly structured (Bray-Curtis index) by the sampling depth (Figure 6B), with a group containing mainly surface samples from the bog (pBog2_S[1,2,3]), and another group separated into two sub-groups: one group composed of 'depth' metatranscriptomes and the second one composed of 'surface' and 'depth' metranscriptomes, including the surface and the depth of a same peat core (pFen2_S3 and pFen2_D3).

Spatial pattern of the expressed functional diversity through the SEED assignations of the metatranscriptomic sequences

The SEED subsystem is a categorized organization of homologous genes from complete genomes into gene functional groups. Our aim was to confirm the functional depth-related pattern and to identify potential structuring of the active microbial community linked to depth or dynamic stage using affiliation to functional gene groups.

On average, 20.64 ± 6.81 % of the metatranscriptomes sequences matched against the subsystem SEED database. The main subsystems represented in the peat metatranscriptomes are these dedicated as Clustering-based subsystems (14.74% of the assigned sequences), followed by Carbohydrates (11.48%), Amino-acids and derivatives (8.35%) and protein metabolism subsystems (8.33%). Photosynthesis and secondary metabolism were relatively poorly represented (respectively 0.19% and 0.55%) compared with the metabolism of aromatic compounds and Cofactors, Vitamins, Prosthetic Groups, Pigments subsystem (1.68% and 6.73%) (Figure 7A). The difference of relative abundance of subsystems tended to be higher between the different depths rather than between fen and bog (Figure 7B): among the 28 subsystems, 9 were more highly expressed at depth in both stages, and 8 more highly expressed at the surface of the two dynamic stages. Thus, 17 out of 28 subsystems presented a consistent depth-related pattern of expression.

Subsystems that were more abundant at the surface can be roughly attributed to the metabolism of plant derived organic compounds (Metabolism of aromatic compounds; Secondary metabolism; Photosynthesis; Fatty Acids, lipids and Isoprenoids) and also included the response to stress. The subsystems that were proportionally more abundant at depth are associated to cell division (DNA metabolism; Cell division and Cell Cycle), nutrients metabolism and uptake (Nitrogen; Potassium), and to the interaction with the environment (Motility and chemotaxis; Virulence, disease and defense).

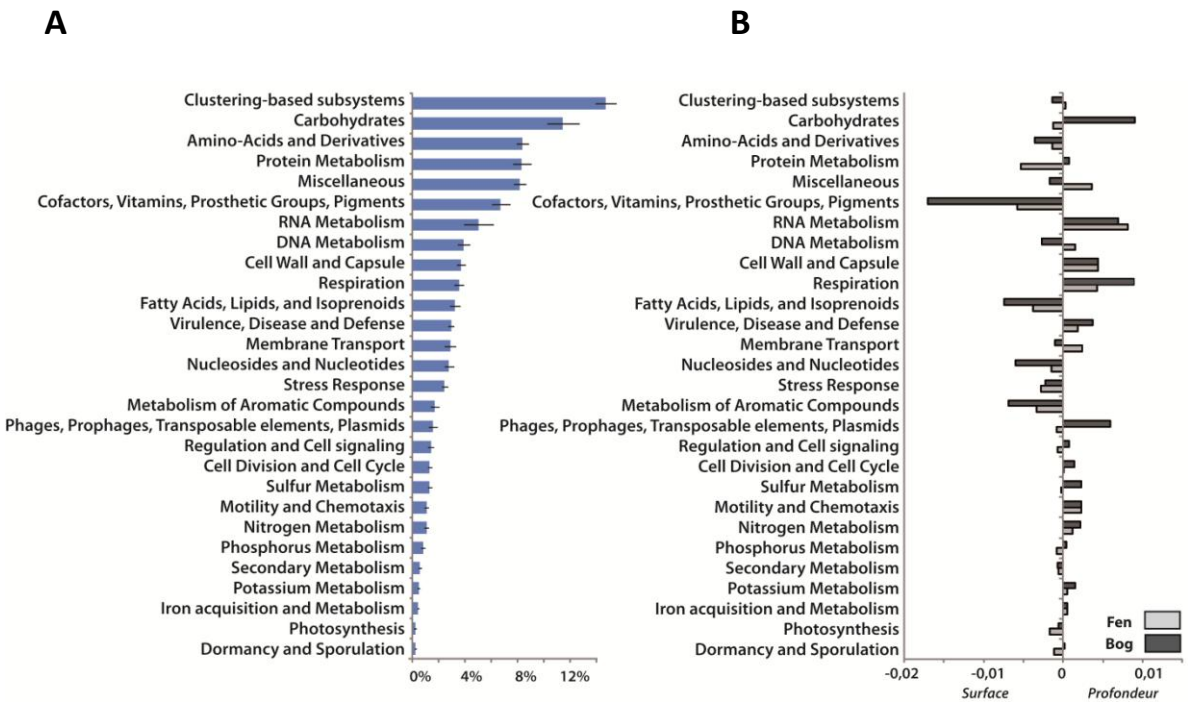


Figure 7: Mean (N=12 ± SD) abundance of the different SEED subsystems in all the metatranscriptomes (A), and difference of expression of the SEED susbsystems between surface and depth ((relative abundance at depth-relative abundance at the surface) negative value: surface, positive value: depth) depending on the dynamic stage (grey: Fen, black: Bog) (B) (N=3).

Differential repartition of taxonomic and functional diversity with depth and dynamic stage in metagenome versus metatranscriptome

In a second time, we wanted to determine whether the taxonomical diversity and the functional potential was different between the surface and depth layers as a consequence of an adaptation of the prokaryotic community to the depth-related environmental conditions that structures genes expression. We therefore compared the diversity (taxonomical and functional) of the metagenomes and the metatranscriptomes.

The comparison of the main parameters indicative of the diversity (alpha-diversity, and proportions of archaeal, prokaryotic and eukaryotic sequences) enclosed in the metagenomes and the metatranscriptomes suggest that the communities were structured differently by the depth and the dynamic stage in the two types of datasets. The diversity parameters from the metagenomes seemed structured by the dynamics stage (fen *vs* bog), with higher proportions of prokaryotes in the fen (MW, N=12, p=0.04) (Figure 8C), and higher proportions of eukaryotes in the bog (MW, N=12, p=0.04) (Figure 8E). For the metatranscriptomes, diversity parameters seemed more influenced by the depth than the dynamic stage with a higher estimated diversity in surface (MW, N=12, p=0.002) (Figure 8B), and higher proportions of sequences assigned to Archaea in depth (MW, N=12, p=0.004) (Figure 8H).

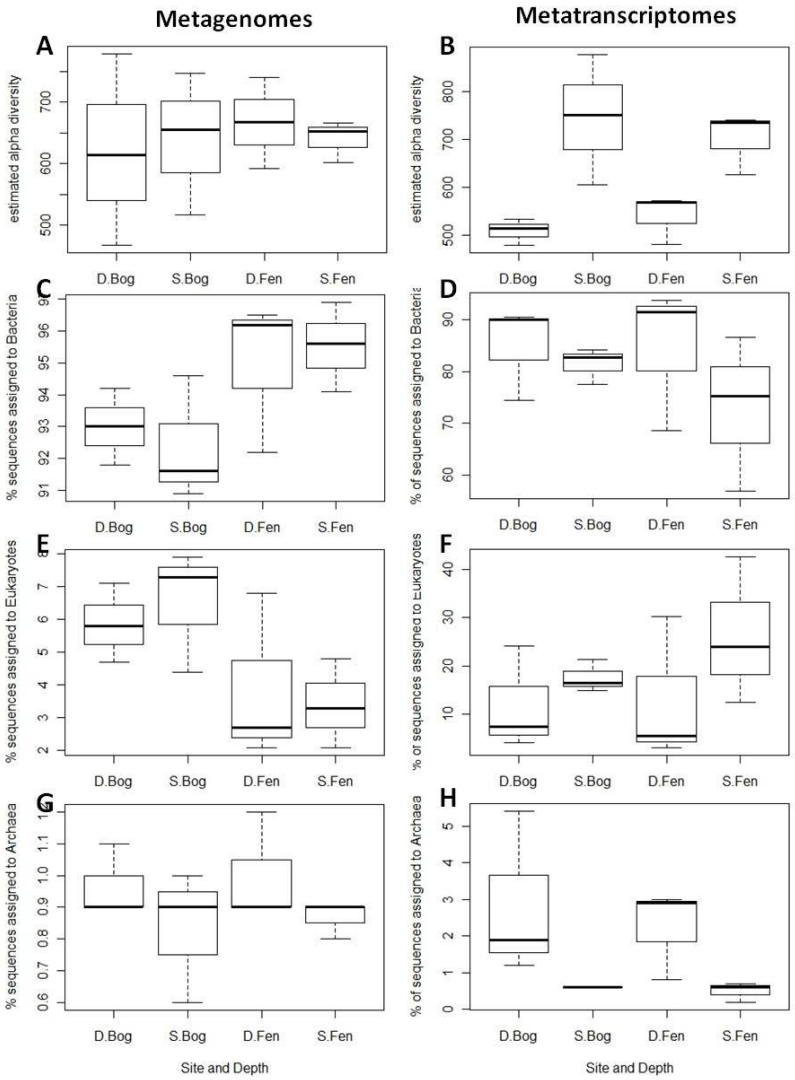


Figure 8 Estimated alpha-diversity (A-B), and percentage of sequences assigned to Bacteria (C-D), Eukaryotes (E-F) and Archaea (G-H) using MG-Rast pipelines for the 12 metagenomes (A-C-E-G) and 12 metatranscriptomes (B-D-F-H). D: Depth (-10 to -15cm under the active capitula layer), S: Surface (-5 to -10cm under the active layer). N=3.

Taxonomical diversity of the rRNA gene sequences from the metagenomes

In the same manner as done in the analysis of the metatranscriptomes, to obtain a first illustration of the taxonomical diversity in the metagenomes, we analyzed the taxonomical assignations of the rRNA gene sequences retrieved in the samples. These taxonomical assignations were obtained by BLASTx against SILVA. The amount of sequences in the metagenomes was less evenly distributed than in the metatranscriptomes: from 29 (pBog1_D3) to 2 292 sequences (pFen_D1). The rRNA gene sequences were proportionally less abundant in these datasets (Table 1). Therefore the results might be biased. However the rRNA gene sequences were as for the metatranscriptomes (Figure 3) mainly assigned to the Alphaproteobacteria (13.95 ± 7.52 %), the Acidobacteria (12.58 ± 5.19 %) and the Verrucomicrobia (12.55 ± 7.19 %) (Figure 9).

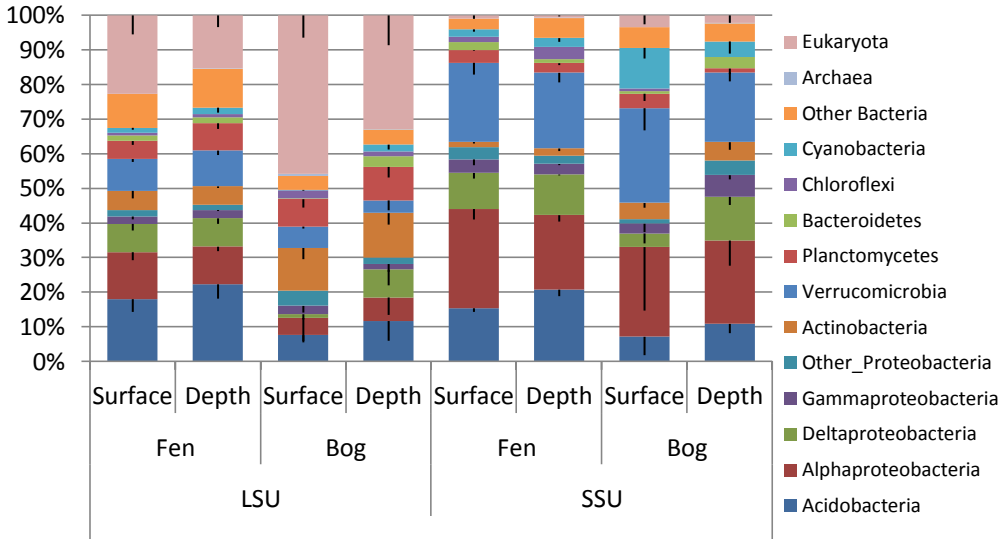


Figure 9 Average proportions of the different phyla depending on the assignation of metagenomic sequences coding for the large and small subunits (LSU and SSU) of the rRNA gene sequences. Bars represent standard deviation and represented for the negative values (N=3).

Due to the uneven distribution of the rRNA gene sequences between samples we did not perform multivariate analysis on this dataset. The diversity of the rRNA genes from the metagenomes tends to suggest that there are more differences between the fen and the bog than observed through the metatranscriptomes, and also suggests less important differences in the metagenomes from the surface and depth as could have been expected from the metatranscriptomes analysis.

Diversity through the taxonomical assignations of the protein-coding sequences from the metagenomes and the metatranscriptomes

In order to investigate the taxonomical and functional differences between the metatranscriptomes and the metagenomes, the protein-coding sequences from the metagenomes were also compared to M5NR and assigned taxonomically and functionally using MG-RAST (Meyer et al., 2008). The taxonomic affiliations at the phyla level were retrieved for each sample, and the relative abundance of each phyla was calculated to build a contingency matrix and to perform a hierarchical clustering based on community structure (Bray-Curtis index) between each sample.

In the metagenomes, the prokaryotic communities are dominated by the Alphaproteobacteria (19.58%), the Actinobacteria (12.70%), the Gammaproteobacteria (10.69%) and the Betaproteobacteria (10.03%) (Figure 10A). Based on the Bray-Curtis (Figure 10B) and Euclidean distances (data not shown), the diversity at the phyla level appears structured by the dynamic stage with fen metagenomes clustering in one group and bog metagenomes clustering into a second group.

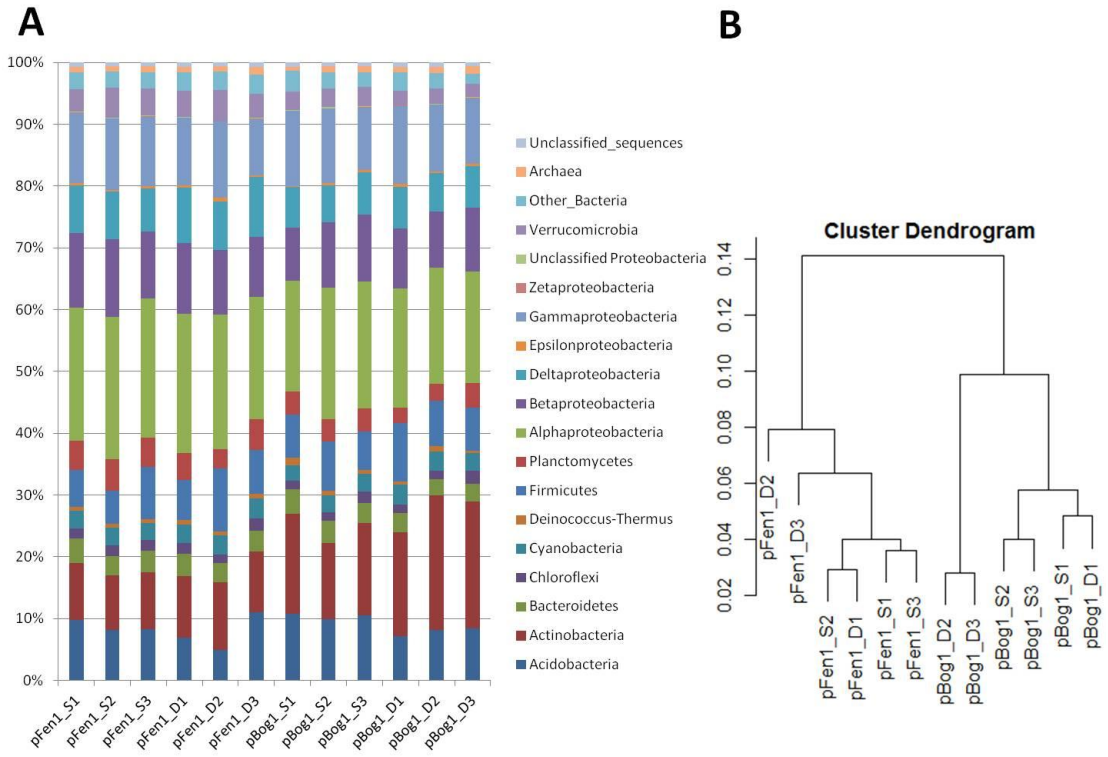


Figure 10 Relative abundance of prokaryotic phyla in the 12 metagenomes depending on the taxonomical assignation of protein-coding genes at the phyla level (A), and hierarchical clustering of the communities depending on their structure (Bray-Curtis distance) (B). Taxonomical assignations of the protein-coding sequences were obtained using MG-Rast pipeline.

In order to analyze the differences between the taxonomical assignation of the protein-coding sequences in the metagenomes and the metatranscriptomes, the variation of structure of the prokaryotic communities from all the samples was analyzed with Non-Metric Multidimensional Scaling (NMDS). NMDS was performed on the whole contingency matrix (metagenome and metatranscriptome sequences taxonomical assignations at the phyla level) depending on the Bray-Curtis distance between the samples. The stress of the representation is low (0.09) indicating a reliable representation of the samples on a two dimensions plan. Chord transformation did not decrease the stress. Ellipses were drawn to represent areas in which 95% of the communities had a probability to be represented depending on their structure and the characteristics associated to the samples (metagenomes/metatranscriptomes, fen/bog, surface/depth). The ellipses representing the metagenome and metatranscriptome areas are distinct, indicating that the two types of dataset gather samples with very different structures. The metatranscriptome ellipse is larger than the one of the metagenome suggesting more variable community structure in the metatranscriptomes. The fen and bog ellipses were superimposed in the metranscriptomes ellipse, but were distinct within the metagenome ellipse (Figure 11). The high superimposition of the fen and bog ellipses in the metatranscriptome ellipse indicated that the structure of fen and bog communities were poorly distinct. The ellipses of the surface and depth groups presented a similar area, were superimposed in

the metagenome ellipse, and were distinct in the metatranscriptomes ellipse, indicating that depth is poorly associated to differences between the metagenome samples, to the contrary of the metatranscriptomes. In the end, imbrications of the metagenome and metatranscriptome ellipses by the surface and depth ellipses suggest that the structure of the prokaryotic communities observed in the metagenomes through the taxonomical assignation of the protein coding genes become differentiated in the metatranscriptomes through a similar variation of the expressed community structure between surface and depth.

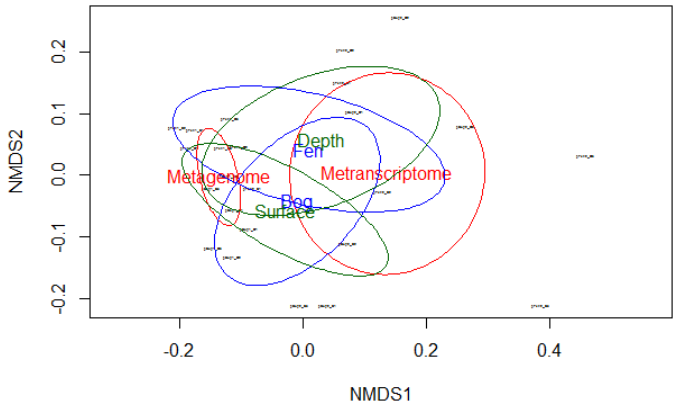


Figure 11 Non-Metric Multidimensional Scaling representing the Bray Curtis-distance between the prokaryotic communities (metagenomes and metatranscriptomes) of the peatland depending on the abundance of the prokaryotic phyla obtained from the taxonomic assignations of the protein coding genes.

As exemplified for the Acidobacteria, functional (Figure 12B) and rRNA genes (Figure 12A) were more highly expressed in depth while no significant differences can be detected in metagenome sequences. This suggested that Acidobacteria are more activated by depth-related environmental conditions, while they are equally present at the surface and in depth.

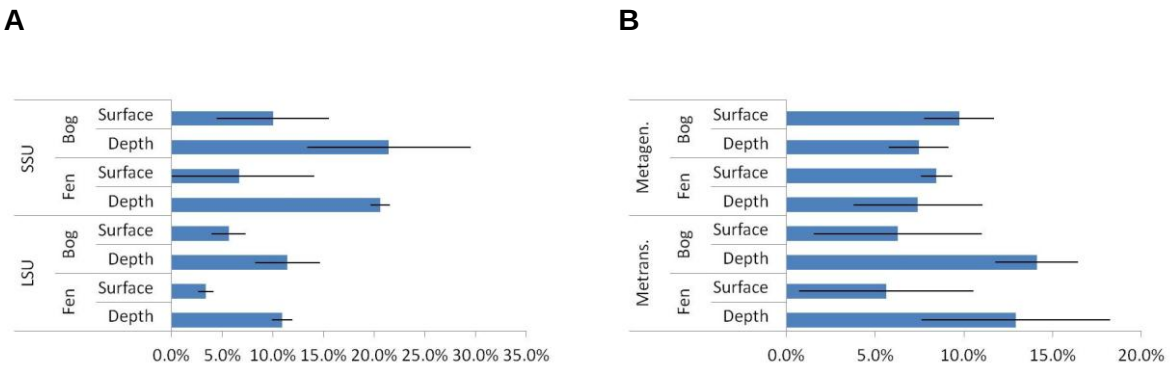


Figure 12 Abundance of the Acidobacteria depending on dynamic stage and depth and resulting from the analysis of the metatranscriptomic rRNA genes diversity (A) or on the taxonomical assignation of the gene coding protein from the metagenomes and metatranscriptomes (B). Bars represent standard error (N=3).

Functional diversity: genome recruitment of the protein-coding sequences from the metagenomes and the metatranscriptomes

In order to obtain better understanding of the differences between the potential functional diversity (metagenomes) and the expressed functional diversity (metatranscriptomes), we retrieved the 50 most recruited genomes in each sample. As the amount of sequences varied a lot between the samples, we ranked the abundance of each genome for each sample (Table 3). For example, the most recruited genome was *Candidatus Solibacter usitatus* with a total of 53 334 sequences from the metagenomes and metatranscriptomes. However, recruitment against this genome ranged from 44 (metagenome pBog1_D3) to 24 138 sequences (metagenome pFen1_D3), while it is respectively ranked 2nd and 1st most recruited genome in these samples.

Table 3 Rank of the 20 most recruited genomes (out of 220 recruited genomes) in the metagenome and the metatranscriptome (framed). "1" correspond to the most recruited genomes (most sequences affiliated to the genome for the sample), and "50" to the least recruited of the 50 genomes. Genomes are ranged by decreasing order of total sequences recruited from the metagenomes and metatranscriptomes.

Data type	Metagenomes												Metatranscriptomes											
Dynamic stage	Fen						Bog						Fen						Bog					
Layer	Surface			Depth			Surface			Depth			Surface			Depth			Surface			Depth		
Peat column	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**	**1**	**2**	**3**
Candidatus Solibacter usitatus	1	1	1		2	1	1	1	1	18	2	2	1	2	4	1	1	3	2	4	4	1	1	2
Rhodopseudomonas palustris	2	2	2	2	1	2	3	3	3	2	3	3	5	12		6	5	25	7	6	2	6	7	9
Acidobacterium capsulatum	5	6	5	5	5	3	5	4	4	3	7	6	6	8	25	3	3	8	5	5	9	4	3	7
Opitutus terrae	7	3	4	3	3	4	14	15	16	32	16		12	11	38	5	6	7	10	11	35	5	10	12
Candidatus Koribacter versatilis	3		3	1			2	2	2	1	1	1	2	3		2	2	4	1	2	5	3	2	4
Sorangium cellulosum	4	4	9	4	4	5	26	14	15	5	11		18	32		13	9	41	13	15	11	19	13	25
Bradyrhizobium sp.	9	5	8	6	12	7	31	10	13	10	8	25	9	33		11	10		19	18	12	20	21	29
Chthoniobacter flavus	6	7	10	7	7	10	11	5	10	34	19	8	7	7	50	12	12	21	32	12	25	17	12	21
Terriglobus saanensis	10	12	19	11	16	6	17	6	8	6	5	5	20	15		7	7		35	10		11	8	28
Gemmata obscuriglobus	12	9	13	8	14	14	8	13	9	25	34	21	16	21		17	23			23		34	22	31
Rhodopirellula baltica	8	10	7	13	18	9	6	12	26	47	29	7	37			36	37					42	37	46
Acidobacterium sp. MP5ACTX9	15	15	21	14	21	12	7	8	6	11	9	27	17	35		14	15		22	19	23	21	11	38
Acidobacterium sp. MP5ACTX8	13	14	12	12	20	11	4	7	5	4	20	11												
Myxococcus xanthus	14	13	31	10	37	16	10	26	20	13	25	10	31			21	18		16	35	18			
Bradyrhizobium sp. BTAi1	17	19	23	16	10	13		19	14	30	13	19	23	42		16	14		40	27	41	38	28	35
Frankia sp.	21	27	28	20	23	19	33	16	7	7	4	4	15	29		20	26		3	7	6	28		41
Burkholderia sp.	16	16	14	15	22	20	9	11	12	12	18		24			37	45			39	42		44	
Pirellula staleyi	22	21	16	18	41	17	23	46			41	49	41			26	25					32	29	33
Planctomyces limnophilus	26	18	41	26		18		20	31			17	28				36						49	
Conexibacter woesei	39			39	17	36	39	25	25	17	23	9	4	6	23	8	19	45	4	3	7	9		30

The most recruited genomes were those of *Candidatus Solibacter usitatus* (Acidobacteria), *Rhodopseudomonas palustris* (Alphaproteobacteria) (26 739 sequences), *Acidobacterium capsulatum* (Acidobacteria) (19 505 sequences), and *Opitutus terrae* (Verrucomicrobia) (18 234 sequences). The 3 first genomes appeared highly recruited in the metagenomes and the metatranscriptomes, while the rank of the genomes afterwards roughly varied more with dataset (metagenome and metatranscriptome) depth, and dynamic stage. This suggests that similar main functions are carried in the fen and bog metagenomes and at the two different depths, while they are differently expressed with depth.

This can be highlighted with the proportion of sequences recruited in each sample against the genome of Acidobacteria such as *Candidatus Solibacter usitatus*, *Candidatus Koribacter versatilis, Acidobacterium capsulatum* and *Acidobacterium sp. MP5ACTX9* (Figure 13). In the metagenomes, to the exception *Candidatus Solibacter usitatus*, the abundance of these genomes varied more with the dynamic stage and little with depth. In the metatranscriptomes, the genomes of these Acidobacteria were more abundantly expressed at depth (Figure 13). This observation is in accordance with the analysis of the taxonomical assignations of the SSU and LSU from the metatranscriptomes, and the taxonomical assignations of the protein coding genes from in the metagenomes and metatranscriptomes (Figure 11).

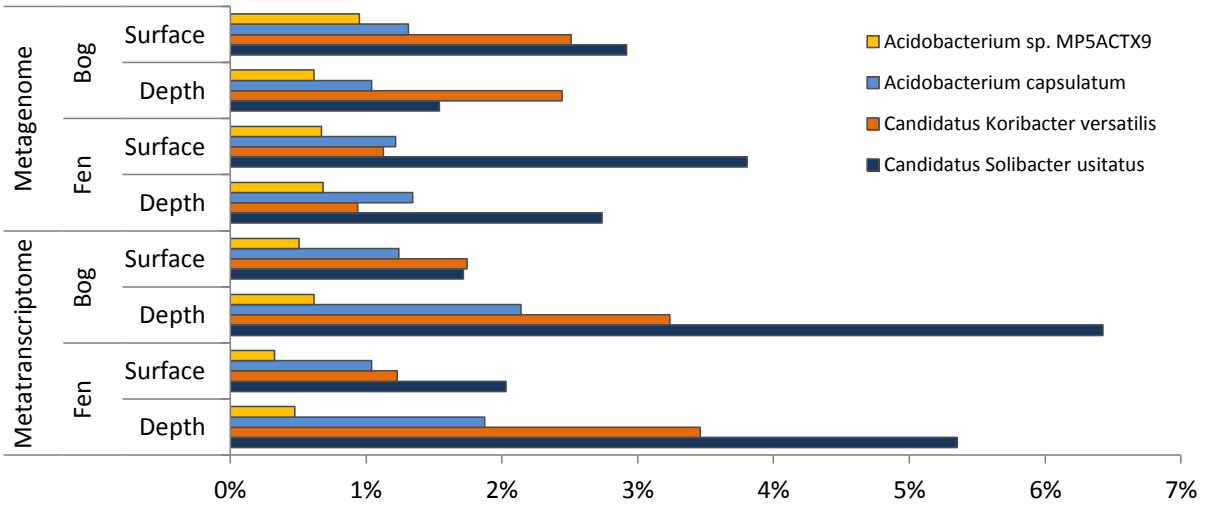


Figure 13 Proportion of sequences recruited against the genomes of 4 Acidobacteria among the genomes recruited in the metagenomes and the metatranscriptomes, depending on depth and dynamic stage.

Discussion

The aim of this study was to address the influence of dynamic stage and depth on the structure of the microbial communities from peatlands. To our knowledge, this is the first study that combines

shotgun metagenomics and metatranscriptomic to triplicate samples from the acrotelm (surface) and catotelm (depth) of a fen and a bog of a *Sphagnum*-dominated peatland. The gene expression level (metatranscriptomics) was analyzed to determine the effect of depth and dynamic stage on the community structure. We combined the results with the analysis of the potential (metagenome) genetic background of these communities in order to detect whether the background diversity behind gene expression corresponded to an adaptation of the microbial community.

Main findings

Between the acrotelm and the mesotelm of a *Sphagnum*-dominated peatland, *Sphagnum*-microbiome gene expression is mainly driven by depth-related environmental conditions. The metagenome and metatranscriptome analysis hint for the existence of a *Sphagnum* functional core-microbiome, with functions not carried by the same organisms in the fen and the bog.

Community genes expression is mainly driven by depth-related environmental conditions

The diversity of identified metabolically active organisms reveals a surface/depth pattern

Thanks to the presence of rRNA sequences covering the three kingdoms and to the relatively similar proportions of these sequences in the metatranscriptomes, we obtained a rather comprehensive vision of the diversity of the metabolically active microorganisms.

The composition of the fen and bog prokaryotic communities in the metatranscriptomes was dominated by Alphaproteobacteria, Acidobacteria, and Verrucomicrobia (Figure 3), which at the phyla level is similar to those determined in other fen and bogs analyzing the diversity of 16s rRNA gene or using FISH counts (Dedysh al et al., 2006; Kulichevskaya et al., 2011; Lin et al., 2012; Tveit et al., 2013; Lin et al., 2014). Prevalence of these taxa is generally associated to their functional fitness in the peatland ecosystem. Analysis of the abundance of the rRNA sequences according to an identified niche (aerobic/anaerobic) indicate that the surface layer is more aerobic as it supports a higher rRNA expression from obligate aerobes (bacterial grazers, primary producers: Cyanobacteria, Euglenozoa) (Figure 5). The higher abundance of Eukaryotes rRNA sequences and more especially protist grazers at the surface, also observed by Tveit et al (2013) is associated to a lower bacterial genes expression in this layer (Figure 8D), suggesting a potential top-down control of the Protozoa on the bacterial biomass (Figures 3, 5, 8) in accordance with a previous study (Morales et al., 2006). The extent to which protist grazing may limit prokaryotic activity cannot be addressed presently as eukaryotic cells produce less rRNA genes copy than the prokaryotes (Fierer et al., 2012).

In contrast, we observe through the metatatranscriptomes that microorganisms from the depth experienced at least partially anaerobic conditions. The main evidence for these anoxic conditions is the higher abundance of Archaea obtained through the analysis of SSU and LSU rRNA (Figure 3), protein-coding sequences (Figure 6), and the recruitment the genomes of methanogene Archaea. This hypothesis is also strengthened by the higher proportion of the anaerobic phototrophes Chloroflexi at depth (Figure 5).

From a different angle, the possibility for fine distinction between depth and surface in such complex microbial communities witnesses for the quality of the metagenomes and metatranscriptomes, and emphasizes the importance of replication of samples to analyze the spatial patterns of prokaryotic communities in peat and therefore soils matrix.

Protein coding sequences expression is also structured by depth-related environmental conditions

Protein-coding genes are associated to the same main phyla as the rRNA sequences (Figure 3 and 6), suggesting a reliable link between the taxonomical and functional diversity. We thus looked for a depth-related pattern of expressed functional (protein-coding) genes in relation to the environmental conditions suggested by the diversity of the identified microorganisms between the two layers.

At the surface, we observed a higher expressed alpha-diversity (Figure 8B), which combined to the higher abundance of the functional category “stress response” (SEED assignations) in the same layer, suggests that more desiccated conditions at the surface created more spatial heterogeneity as already suggested previously (Bragina et al., 2014; Parry et al., 2014). Functional categories of genes assignations suggests a higher autotrophic activty at the surface (in relation to the higher abundance of rRNA from Cyanobacteria for example), with more sequences assigned to “Metabolism of aromatic compounds”, “Secondary metabolism, and Photosynthesis”, as observed in previous study that compared fen and bog at two depths (Lin et al., 2014). In contrast functional categories analysis also suggested that at depth, microorganisms transcribe more RNA for genes encoding functions involved in respiration and nitrogen and phosphorus metabolism (Figure 7), suggesting more heterotrophic activity. In the meantime, higher assignations to the functional category “chemotaxis/motility” and to “Virulence Disease and defense” at depth suggest competition and thus interactions among the microbial communities as suggested previously (Bragina et al., 2014).

Metagenomic analysis suggests that the genetic diversity is more structured by the dynamic stage than by depth

rRNA genes from the metagenomes differ more between fen and bog

At the phyla level assignations of the rRNA genes from the metagenomes are similar to these of the metatranscriptomes (Figure 9). However, to the contrary of the metatranscriptomes, the structure of the prokaryotic communities seems more dependent on the dynamic stage, and therefore on the *Sphagnum* species. This distinction between fen and bog communities in the metagenomes is also observed through the taxonomical assignations of the protein-coding sequences at the phyla level (Figure 10).

The different microbial community structure in the fen and the bog is in accordance with previous studies that focused on the 16S rRNA gene diversity (Opelt at al., 2007, Bragina et al., 2012; Lin et al., 2012). This difference can be explained through the different trophic status of the two dynamic stages, associated with different physico-chemical features, such as pH, that shape differently the communities (Preston et al., 2012).

No apparent functional difference in the metagenome from the fen and the bog

Interestingly, the abundance of some phyla varied consistently between the two dynamics stages, Bacteroidetes being found to be more abundant in the fen, and Actinobacteria in the bog (Figure 9, Figure 10). These two phyla have been shown to be involved in the degradation of cellulose, hemicelluloses and the production of debranching enzymes (Pankratov et al., 2011; Tveit et al., 2013). It can therefore be hypothesized that different phyla carry similar functions in the two *Sphagnum* species from the fen and the bog, as already suggested for N-fixing genes or methanogene/methanotrophes communities (Bragina et al., 2012; Basiliko et al., 2013). In the same manner, while the studies conducted on the diversity of the rRNA genes observed in general a twice higher diversity in the fen compared with the bog, in consequence to the higher nutrient diversity, we did not observe such a difference for the protein-coding genes (Figure 8A). This suggests that there might be some functional redundancy in the fen, but that the functional potential as a whole does not differ between fen and bog. This is emphasized by the variation of the structure of the prokaryotic communities depending on the taxonomy of the protein-coding genes (Figure 11). Fen and bog community structures were distinct but rather similar compared to the variation of the community structure in the metatranscriptomes. And the structure of the communities of the fen and the bog varied in the same way between surface and depth from the metagenomes to the

metatranscriptomes (Figure 11). It is therefore likely that these similar sets of functions are adapted to the environment shaped by the interactions between the microbiome and the *Sphagnum,* as suggested by the functional categories abundances (SEED assignations) (Figure 7) (Bragina et al., 2014). However, these functions are likely carried by different taxa between the *Sphagnum*-species, as a consequence of the influence of the *Sphagnum* on its environment (pH, water-logging) (Preston et al., 2012), and of the conservation of an adapted microbiome throughout the *Sphagnum* life-cycle (Bragina et al., 2012; Bragina et al., 2013b) that sustains *Sphagnum* growth (Vile et al., 2014).

Towards a Sphagnum core-microbiome?

There is a growing interest about the microbial communities associated with plants (Zilber-Rosenberg and Rosenberg, 2008; Vandenkoornhuyse et al., 2014). In the case of the *Sphagnum* mosses and the sphagnosphere, research has mainly focused on the effect of the *Sphagnum* on the microbial communities (Jassey, 2011). However, we still have a poor understanding of the impact of the prokaryotic community on the *Sphagnum* fitness (Vile et al., 2014), and of the interactions among the prokaryotes. Vertical transmission of the prokaryotic communities along the *Sphagnum* life-cycle (Bragina et al., 2012; Bragina et al., 2013b), and adapted (Bragina et al., 2014) and conserved prokaryotic community structures in relation to a *Sphagnum* species over long geographical distances hint for a long-term association between the prokaryotic community and the *Sphagnum*.

While the taxonomical diversity of the fen and bog prokaryotic communities differed, the similar selection of functional gene expression in relation to depth suggests a similar functioning of the prokaryotic communities from *Sphagnum fallax* and *Sphagnum magellanicum*. The common functional attributes of the fen and bog communities were however emphasized when we observed the high sequence recruitment to four main genomes that are dominant in all the samples.

The high recruitment of sequences to the genomes of *Candidatus Koribacter versatilis* and *Candidatus Solibacter usitatus* seems coherent, as sequences of rRNA genes close to their rRNA genes have been retrieved from the peat (Pankratov et al., 2011). These genomes contain genes that allow their growth from diverse and complex carbon substrates, and genes that allows their resistance to a high moisture variation (Ward et al., 2009), which allows a good fitness in the *Sphagnum* environment.

The general dominance of these genomes suggests that a core of functional genes is widely carried and expressed similarly by the prokaryotes from the two depths of the fen and the bog. However, the question remains as to the main functions that compose this microbiome and as to the diversity

of the organisms that are involved in it. It is possible that this functional potential and its expression is carried by a group of syntrophic/symbiotic (tighly interacting) organisms that share the metabolism of the *Sphagnum* compounds, as could be hypothesized by horizontal gene transfer between Acidobacteria and *Rhodopseudomonas palustris* (the second most dominant gemone recruited) (Quaiser et al., 2003), which suggests possible co-adaptation between two of the main phyla present in the fen and the bog, or by the impossibility to obtain pure cultures of peatland Acidobacteria, while these grew more easily in a biofilm-mediated co-culture (Dedysh et al., 2006).

Références du Chapitre 2

Artz R.E.E. 2009. Microbial community structure and carbon substrate use in northern peatlands. In: Baird A.J., Belyea L.R., Comas X., Reeve A.S., Slater L.D. (eds), Carbon cycling in northern peatlands. *Geophysical Monograph Series* 184: 111-129.

Basiliko N., Henry K., Gupta V., Moore T.R., Driscoll B.T., Dunfield P.F. 2013. Controls on bacterial and archaeal community structure and greenhouse gas production in natural, mined, and restored Canadian peatlands. *Frontiers in Microbiology* 4, article 215, doi: 10.3389/fmicb.2013.00215

Bragina, A., Berg, C., Cardinale, M., Shcherbakov, A., Chebotar, V., and Berg, G. 2012. Sphagnum mosses harbour highly specific bacterial diversity during their whole lifecycle. *ISME J*. 6 : 802-813. doi:10.1038/ismej.2011.151.

Bragina A., Berg C., Müller H., Moser D., and Berg G. 2013a. Insights into Functional Bacterial Diversity and Its Effects on Alpine Bog Ecosystem Functioning. *Scientific Reports* 3 (January). Nature Publishing Group: 1955. doi:10.1038/srep01955.

Bragina, A., Cardinale M., Berg C., and Berg. G. 2013b. Vertical Transmission Explains the Specific Burkholderia Pattern in Sphagnum Mosses at Multi-Geographic Scale. *Frontiers in microbiology* 4:394.

Bragina, A., Oberauner-Wappis L., Zachow C., Halwachs B., Thallinger G. G., Müller H., and Berg G. 2014. The Sphagnum Microbiome Supports Bog Ecosystem Functioning under Extreme Conditions. *Molecular Ecology*. doi:10.1111/mec.12885.

Dedysh, S. N., Pankratov, T. A., Belova, S. E., Kulichevskaya, I. S., and Liesack, W. 2006. Phylogenetic Analysis and In Situ Identification of Bacteria Community Composition in an Acidic Sphagnum Peat Bog. Appl. Environ. Microbiol. 72, 2110-2117. doi:10.1128/AEM.72.3.2110-2117.2006.

Fierer, Noah et al. 2012. Cross-Biome Metagenomic Analyses of Soil Microbial Communities and Their Functional Attributes. *Proceedings of the National Academy of Sciences of the United States of America* 109(52):21390–95. Huang, Y., Gilna, P., and Li, W. (2009). Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* 25, 1338-1340. doi:10.1093/bioinformatics/btp161

Hultman, J., Waldrop, M. P., Mackelprang, R., David, M. M., McFarland, J., Blazewicz, S. J., Harden, J., Turetsky, M. R., McGuire, A. D., Shah, M. B., et al. (2015). Multi-omics of permafrost, active layer and thermokarst bog soil microbiomes. *Nature* 521, 208-212.

Huson, D. H., Auch, A. F., Qi, J., and Schuster, S. C. (2007). MEGAN analysis of metagenomic data. *Genome Res.* 17, 377-386. doi:10.1101/gr.5969107.

Husson F., Josse J., Le S.and Mazet J. 2015. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining. *R package version 1.29*. http://CRAN.R-project.org/package=FactoMineR

Jassey V. 2011. *Impact d'un réchauffement climatique sur le fonctionnement de la sphagnosphère : relations polyphenols-communautés microbiennes.* Thèse de l'Université de Franche-Comté, Besançon (France): 238 p.

Juottonen, H., Galand, P. E., Tuittila, E.-S., Laine, J., Fritze, H., and Yrjälä, K. (2005). Methanogen communities and Bacteria along an ecohydrological gradient in a northern raised bog complex. *Environ. Microbiol*. 7, 1547-1557.

Kulichevskaya, I. S., Belova S. E., Komov V. T., Dedysh S. N., and Zavarzin. G. A. 2011. Phylogenetic Composition of Bacterial Communities in Small Boreal Lakes and Ombrotrophic Bogs of the Upper Volga Basin. *Microbiology* 80(4):549–57.

Lin X., Green S., Tfaily M.M., Prakash O., Konstantinidis K.T., Corbett J.E., Chanton J.P., Cooper W.T., and Kostka J.E. 2012. Microbial Community Structure and Activity Linked to Contrasting Biogeochemical Gradients in Bog and Fen Environments of the Glacial Lake Agassiz Peatland. *Applied and Environmental Microbiology* 78 (19): 7023–31. doi:10.1128/AEM.01750-12.

Lin X., Tfaily M.M, Steinweg J.M., Chanton P., Esson K., Yang Z.K., Chanton J.P., Cooper W., Schadt C.W., Kostka J.E. 2014. Microbial Community Stratification Linked to Utilization of Carbohydrates and Phosphorus Limitation in a Boreal Peatland at Marcell Experimental Forest, Minnesota, USA. *Applied and Environmental Microbiology* 80 (11). American Society for Microbiology: 3518–30. doi:10.1128/AEM.00205-14.

Meyer F., D. Paarmann, M. D'Souza, R. Olson , E. M. Glass, M. Kubal, T. Paczian , A. Rodriguez , R. Stevens, A. Wilke, J. Wilkening, R. A. Edwards. 2008. The Metagenomics RAST server – A public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC Bioinformatic 9:386.*

Morales, Sergio E. et al. 2006. Comparison of Bacterial Communities in New England Sphagnum Bogs Using Terminal Restriction Fragment Length Polymorphism (T-RFLP). *Microbial Ecology* 52(1):34–44. (http://www.ncbi.nlm.nih.gov/pubmed/16729225).

Oksanen J., F. Blanchet G, Kindt R., Legendre P., Minchin P. R. , O'Hara R. B., Simpson G. L., Solymos P., Henry M., Stevens H. and Wagner H. 2015. vegan: Community Ecology Package. *R package version 2.3-0*. http://CRAN.R-project.org/package=vegan

Opelt, K., Berg, C., Schönmann, S., Eberl, L., and Berg, G. 2007. High specificity but contrasting biodiversity of *Sphagnum*-associated bacterial and plant communities in bog ecosystems independent of the geographical region. *ISME J*. 1, 502-516. doi:10.1038/ismej.2007.58.

Pankratov, T. A., Ivanova, A. O., Dedysh, S. N., and Liesack, W. (2011). Bacterial populations and environmental factors controlling cellulose degradation in an acidic Sphagnum peat. Environ. *Microbiol*. 13, 1800-1814

Parry, L. E., Holden J., and Chapman.P.J. 2014. Restoration of Blanket Peatlands. *Journal of environmental management* 133:193–205.

Preston M.D., Smemo K.A., McLaughlin J.W., Basiliko N. 2012. Peatland microbial communities and decomposition processes in the James Bay Lowlands, Canada. *Front. Microbiol.* 3:70 doi: 10.3389/fmicb.2012.00070

Quaiser, A. et al. 2003. Acidobacteria Form a Coherent but Highly Diverse Group within the Bacterial Domain: Evidence from Environmental Genomics. *Molecular microbiology* 50(2):563–75.

Quaiser, A., Bodi, X., Dufresne, A., Naquin, D., Francez, A. J., Dheilly, A., Coudouel, S., Pedrot, M., and Vandenkoornhuyse, P. (2014). Unraveling the stratification of an iron-oxidizing microbial mat by metatranscriptomics. *PLoS One* 9. doi:10.1371/journal.pone.0102561

Quast, C. et al. 2013. The SILVA Ribosomal RNA Gene Database Project: Improved Data Processing and Web-Based Tools." *Nucleic acids research* (41):590–96.

Rydin H., Jerglum J. 2006. *The Biology of Peatlands.* Biology of Habitats Series. Oxford University Press, Oxford & New-York

Serkebaeva, Y. M., Kim, Y., Liesack, W., and Dedysh, S. N. 2013. Pyrosequencing-based assessment of the bacteria diversity in surface and subsurface peat layers of a northern wetland, with focus on poorly studied phyla and candidate divisions. *PLoS One* 8, e63994.

Tveit, A., Schwacke, R., Svenning, M. M., and Urich, T. 2013. Organic carbon transformations in high-Arctic peat soils: key functions and microorganisms. *ISME J*. 7, 299-311. doi:10.1038/ismej.2012.99.

Tveit A., Urich T, Frenzel P., and Svenning M. M. 2015. Metabolic and Trophic Interactions Modulate Methane Production by Arctic Peat Microbiota in Response to Warming. *Proceedings of the National Academy of Sciences of the United States of America* 112 (19): E2507–16. doi:10.1073/pnas.1420797112.

Vandenkoornhuyse, P., Quaiser, A., Duhamel, M., Le Van, A., & Dufresne, A. (2015). The importance of the microbiome of the plant holobiont. *The New Phytologist*, *206*(4), 1196–206. doi:10.1111/nph.13312

Vile M.A., Wieder R.K., Zivkovic T., Scott K.D., Vitt D.H., Hartsock J.A., Iosue C.L., Quinn J.C., Petix M., Fillingim H.M., Popma J.M.A., Dynarski K.A., Jackman T.R., Albright C.M., Wykoff D.D. 2014. N2-fixation by methanotrophs sustains carbon and nitrogen accumulation in pristine peatlands. *Biogeochemistry* 121: 317-328.

Ward, N. L. et al. 2009. Three Genomes from the Phylum Acidobacteria Provide Insight into the Lifestyles of These Microorganisms in Soils. *Applied and Environmental Microbiology* 75(7):2046–56.

Yergeau, E., Hogues H., Whyte L. G. , and Greer C. W. 2010. The Functional Potential of High Arctic Permafrost Revealed by Metagenomic Sequencing, qPCR and Microarray Analyses.*The ISME journal* 4(9):1206–14.

Zilber-Rosenberg, I., & Rosenberg, E. (2008). Role of microorganisms in the evolution of animals and plants: the hologenome theory of evolution. *FEMS Microbiology Reviews*, *32*(5), 723–35. doi:10.1111/j.1574-6976.2008.00123.x

Chapitre 3:

Large scale cross-peatlands comparisons of viral abundance and diversity

Chapitre 3: Large scale cross-peatland comparison of viral abundance and diversity

Résumé du Chapitre 3

L'étude de la dynamique spatio-temporelle des communautés virales d'une tourbière a montré un fort effet structurant des variations temporelles et des conditions environnementales associées sur la diversité virale, ainsi qu'un passage de la stratégie lysogénique vers la stratégie lytique au cours du printemps. En parallèle, nous avions aussi montré une spécificité des communautés de procaryotes pour deux espèces Sphaignes dominantes dans cette tourbière, et une similarité fonctionnelle de ces communautés dont l'expression est sélectionnée par les conditions environnementales liées à la profondeur.

Afin de mieux caractériser le fonctionnement du compartiment viral des tourbières à *Sphagnum* et de déterminer si ce fonctionnement est similaire entre des sites éloignés géographiquement, nous avons comparé l'abondance et la diversité virale de 5 tourbières, situées en Finlande, au Canada (Québec), en France (2 sites) et sur l'île subantarctique d'Amsterdam. Afin de mettre l'accent sur les patrons spatiaux de l'abondance et la diversité virale, une méthode d'extraction des virus du sol a été adaptée pour extraire les virus de la matrice de tourbe. La diversité virale a ainsi pu être comparée à une échelle spatiale très fine des différents sites, c'est-à-dire entre deux profondeurs proches de la surface des tourbières (acrotelme et mésotelme), mais aussi entre les communautés virales de la matrice de tourbe et de l'eau des Sphaignes.

L'abondance des particules virales de la matrice de tourbe dans les 5 tourbières étudiées s'étend de $4.34 \pm 1.47\ 10^{6}$ à $1.42 \pm 0.73\ 10^{10}$ particules virales par gramme de tourbe sèche indiquant une très forte variabilité géographique. En Finlande où l'abondance virale a été comparée entre le fen et le bog, on observe une faible variabilité intra-site. L'analyse conjointe de l'abondance des particules virales et des procaryotes et de la variation des paramètres physico-chimiques indique une covariation positive des entités biologiques, du carbone organique dissout et du sulfate, confirmant l'hypothèse qu'une plus forte activité procaryote en lien avec l'augmentation du carbone organique dissous est associée à une plus forte production virale.

L'analyse de 19 métaviromes produits pour les communautés virales de l'eau des Sphaignes et de la matrice à deux profondeurs de l'acrotelme des tourbières du Canada et de Finlande montre un très fort effet du site sur la diversité géographique virale totale ou taxonomiquement identifiée. Le deuxième effet important est celui de l'origine de la communauté virale : de l'eau ou de la matrice.

En Finlande, la diversité virale totale contenue dans les communautés de la matrice de tourbe varie d'abord en fonction du stade dynamique, puis en fonction de la profondeur, confirmant l'hypothèse de communautés d'hôtes procaryotes taxonomiquement distinctes et une sélection similaire de leur activité en fonction de la profondeur. Les communautés virales de l'eau de Sphaigne du fen et du bog présentaient une plus forte ressemblance avec les communautés virales de la matrice de tourbe du fen, confirmant aussi le rôle de la nappe d'eau dans la structuration des communautés virales de l'eau de Sphaigne : cet effet s'associe à une homogénéisation de la diversité virale quand la nappe d'eau est haute.

Au Canada où la nappe était plus basse que la profondeur de tourbe échantillonnée, on observe une forte distinction entre les communautés de la matrice de tourbe et celles de l'eau de Sphaigne. Cependant, les deux types de communautés présentent une très forte structuration en fonction du point d'échantillonnage et non de la profondeur, confirmant là-encore le rôle de l'imbibition de la nappe dans l'homogénéisation des communautés virales, et dans la structuration de l'activité des communautés d'hôtes. Cette forte hétérogénéité spatiale en conditions sèches renforce aussi l'importance des Sphaignes comme micro-habitat pour les communautés de l'eau de Sphaigne et de la matrice de tourbe, renforçant ici le concept de Sphagnosphère.

La comparaison des communautés virales de l'eau de Sphaigne du Canada et de Finlande avec la diversité virale observée dans le site échantillonnée dans le chapitre 1, qui présente les mêmes espèces de Sphaignes que la Finlande suggère que les communautés virales issues d'espèces de Sphaignes similaires présentent une plus forte ressemblance, liées aux relations ténues et entretenues entre la Sphaigne et son microbiome. Enfin, la spécificité des 31 communautés virales de tourbe produites à ce jour et comparée à la diversité virale actuellement connue dans d'autres écosystèmes couvrant les eaux marines, douces, les milieux hyper-salins et les sols des déserts froids renforce l'idée d'un fonctionnement similaire des communautés virales et de leurs hôtes dans des tourbières déconnectées et éloignées.

Large scale cross-peatlands comparisons of viral abundance and diversity

Introduction

Viruses remained for a long time ignored in ecosystem functioning as their small size, and their host dependence have necessitated the development of adequate tools to allow their identification (Wommack et al., 2015). Yet, increasing body of studies about these biological acellular entities acknowledged their high abundance and their crucial role in the structuration and regulation of microbial communities (Fuhrman 1999; Suttle 2005; Suttle 2007). These observations were obtained through two main approaches: viral counts and sequencing of viral nucleic acids that is yet called "the third age of phage ecology" (Comeau et al., 2008; Wommack et al., 2015). In parallel to technical approaches, most knowledge about viral ecology was gained with either parallel comparisons of viral abundance and diversity between different ecosystems (Williamson et al., 2005; Angly et al., 2006 ; Srinivasiah et al., 2008; Fancello et al., 2012; Roux et al., 2012; Brum et al., 2015a), or with analysis of the short term variations and seasonal fluctuations (Jackson-Farnell and Ward 2003; Lopez-Bueno et al., 2009; Rodriguez-Brito et al., 2010; Lauro et al., 2011; Parsons et al., 2012; Chow and Fuhrman., 2012; Pagarete et al., 2013). In the end, it is the cross-combination of the technical approaches, counts and sequencing, and of spatial and temporal sampling strategies that generates a more comprehensive view of viral ecology in a given ecosystem (Payet and Suttle 2013; Brum et al., 2015a). While viral ecology of marine and freshwater ecosystems have received increasing attention since more than two decades, further technical limitations have postponed the study of viral ecology in soils and wetlands by a dozen of year (Williamson et al., 2003; Williamson et al., 2005; Srinivasiah et al., 2008). Yet, apart from a seasonal dynamics of viral abundance in a riverine wetland (Jackson-Farnell and Ward, 2003), a *Sphagnum*-dominated peatland (Chapter 1.A.), and a few comparisons of viral abundances (Williamson et al., 2005; Williamson et al., 2007; Srinivasiah et al., 2008) and diversity (Fierer et al., 2007; Zablocki et al., 2014; Adriaenssens et al., 2015), knowledge about viral ecology in soils and wetlands is still scarce.

In the spatio-temporal analysis carried in a temperate *Sphagnum*-dominated peatland, seasonality appeared as the factor structuring most the diversity and abundance of the viral communities (Chapter 1.A.) since the genetic diversity from the winter/spring pore water viral communities was distinct from the diversity of summer and autumn communities. This diversity shift was hypothesized to be a consequence of a seasonal shift in the viral strategies from lysogenic to lytic interactions linked to higher host activity in summer. This is in accordance with the hypothesis that prophages act as regulatory switches of bacteria in dependence on stress conditions (Feiner et al., 2015). In parallel,

the same temporal fluctuation of viral diversity was observed in the Antarctic ocean, and was similarly explained by a low host activity associated to lysogenic viral strategies in winter and a switch to lytic strategies in summer when hosts become metabolically more active (Brum et al., 2015a). Considering that most peatlands are situated under boreal latitudes, this similar seasonality of virus/host interactions between a polar ecosystem and a temperate peatland stresses the need to compare the viral compartment in different peatlands under higher latitudes to confirm the functioning of the viral compartment in these ecosystems. Furthermore, the most concerning conclusions of the study led in the Antarctic ocean was the potential disruption of virus/hosts interactions in favor of lytic over lysogenic strategies in response to increased prokaryotic productivity with water warming, and its consequences on the carbon cycling in the polar regions (Paul et al., 2008; Brum et al., 2015a). In soils, the prevalence of lysogeny is estimated to increase with latitude were global changes are also expected to occur faster and stronger (Serreze et al., 2000). Considering the putative role of lysogeny in the seasonality of peatland viral ecology, the importance of peatlands as major soil carbon stocks, and the importace of the lowered prokaryotic activity in this function, the assessment of viral ecology in peatlands by comparing different sites with different environmental conditions is necessary to document their role in these ecosystems.

From the comparison of viral diversity along latitudinal gradients of marine plankton emerged the issue of the ubiquity of viral types in correlation to the distribution range of their hosts and the extent of viral diversity at local and global scales (Breitbart and Rohwer., 2005; Mojika et al., 2015, Brum et al., 2015b). Such analyses question both the way viruses migrate (Hewson et al., 2012; Winter et al., 2013) and evolve (Krupovic et al., 2011). Interestingly, compared with the viral diversity from other ecosystems, the pore-water viruses from the surface of the peatland appeared seasonally highly variable, but also very specific for this ecosystem and distinct from the viral communities of the surface of two geographically close freshwater lakes (Chapter 1.A.). In contrast, the viral communities from these two lakes appeared genetically close (Roux et al., 2012), which questions the importance of allochtonous viruses in the viral communities in the surface of ecosystems (Hewson et al., 2012), and the specificity of peatland viral communities at large geographical scales. In addition, prokaryotic genetic diversity studied through rRNA gene profiling has been shown to be specific for a *Sphagnum*-species over large geographical distances (Opelt et al., 2007; Bragina et al., 2012) while the gene expression in these communities is shaped and selected by the surrounding environmental conditions (Tveit et al., 2013; Lin et al., 2014; Chapter 2). As viral diversity results from the diversity and the activity of their host community, cross-peatland comparison of the viral diversity represents an interesting starting point to address the specificity and the potential of viral evolution or migration in similar but distant biotopes.

Finally, the prokaryotic diversity and gene expression observed through the DNA and RNA pool extracted directly from the peat appeared spatially structured (Chapter 2). In contrast, the viral diversity was studied in the *Sphagnum* pore-water and exhibited no spatial pattern (Chapter 1.A.).

Consistence of viral ecology between geographically distant peatlands was investigated through the comparison of viral abundance and diversity in five different peatlands along a climatic gradient using an adapted extraction method of the viral particles from the peat matrix previously developed on soil samples (Williamson et al., 2005) at two different depths of the peat surface (the acrotelm and the superior part of the mesotelm). In addition, we compared the metaviromes of the viral communities from the peat matrix and from the *Sphagnum* pore-water. We expected an influence of site and depth on viral abundance as observed in Chapter 1.A. in relation to the differences in the physico-chemical features and climatic conditions of the different sites. For the viral diversity, we expected differences in the peat matrix and pore-water and an effect of the site and depth. Considering the potential functional specificity of *Sphagnum* prokaryotic communities (Chapter 2; Bragina et al., 2014), and the taxonomical specificity of the prokaryotic communities for a *Sphagnum* species (Opelt et al., 2007; Bragina et al., 2012), we also looked for the influence of *Sphagnum* species on the viral diversity, and for the specificity of viral communities for peatland ecosystems even at a broad geographical range.

Material and methods

Site description and experimental design

Five *Sphagnum*-dominated peatlands were selected in order to cover a maximum latitudinal gradient, i.e boreal (Finland and Canada), temperate (Frasne and La Guette in France) and Southern Hemisphere (Amsterdam Isle) (Fig. 1.). The main climatic and floristic characterics are summarized in Table 1.

Figure 1 Geographical situation of the peatlands (A-E, C=La Guette, D=Frasne) for which viral abundance and/or diversity is analyzed. Precise GPS coordinates and main features of the peatlands are detailed in Table 1.

Table 1 Peatlands sampled for the comparison of viral particles abundance and virus diversity

Point	Country	Peatland	Date	Latitude	Longitude	Altitude	Mean annual Temperature and precipitations	Sampled stages	*Sphagnum* type and associated vegetation
A	**Finland**	**Lakkasuo**	11/09/14	61°47' N	24°18' E	150m	+3.0°C ; 650 mm	fen bog	*S.fallax / Carex rostrata S. magellanicum / Eriophorum*
B	**Canada**	**Bois-des-Bel**	24/09/14	47° 58' N	69° 26' W	28m	+4.3°C ; 1070 mm	bog	*S. fuscum*
C	**France**	**La Guette**	14/10/14	47° 19' N	02° 16' E	160m	+11°C ; 880 mm	bog	*S. rubellum Calluna*
D	**France**	**Frasne**	15/10/14	46°49' N	06°10' E	836m	+7°C ; 1400 mm	fen	*S. fallax, S. rubellum, Carex limosa, C. rostrata*
E	**Amsterdam Isle**	**Caldeira**	13/12/14	37° 50' S	77° 33' E	~500m	+14°C ; 1100 mm	fen	*Sphagnum* spp.

In order to strictly compare the viral and prokaryotic communities, a common sampling protocol was designed ahead of the field campaigns. In each site, peat cores were extracted (N= 2 to 4) (Table 2) and peat was sampled at the two depth to analyze viral particles and prokaryotes abundances and diversity. In addition, pore-water was expressed from the peat to study the viral diversity from the pore-water (Table 2). Metaviromes were produced for both pore-water and peat matrix.

Table 2 Peat and pore-water samples collected for the comparison of viral ecology between the different peatlands.

	VPA and PA (surface/depth)	"Peat" metaviromes*	"Water" metaviromes*
Lakkasuo (Fen)	N=3 / layer	vFin_Fen2_Sp vFin_Fen2_Dp vFin_Fen3_Sp vFin_Fen3_Dp	vFin_Fen1w vFin_Fen3w *in progress*
Lakkasuo (Bog)	N=3 / layer	vFin_Bog1_Sp vFin_Bog2_Sp vFin_Bog3_Sp	vFin_Bog3w *in progress*

		vFin_Bog2_Dp vFin_Bog3_Dp	
Bois-des-Bel	N=2 / layer	vCan_Nat1_Sp vCan_Nat1_Dp vCan_Nat2_Sp vCan_Nat2_Dp	vCan_Nat1_Dw vCan_Nat2_Sw vCan_Nat2_Dw
La Guette	N=3 / layer	*In progress*	*In progress*
Frasne	N=3 / layer	*In progress*	*In progress*
Caldeira	N=4 / layer	*In progress*	*In progress*
TOTAL	36 samples	12 peat metaviromes	6 pore water metaviromes

***In the metaviromes names: p (last letter) corresponds to viruses extracted from the peat matrix, w (last letter) corresponds to viruses extracted from the pore-water, number (1, 2 or 3) corresponds to the peat column (or sampling point), S corresponds to surface metaviromes, and D corresponds to depth metaviromes. Surface and depth were not sampled separately for the pore-water samples from Lakkasuo. Surface: 5 to 10cm, and Depth: 10 to 15cm below the active capitula layer, except for Canada, Surface: 5 to 15cm, and Depth: 15 to 25cm under the capitula layer.**

Sampling protocol

For each site, N (2 to 4) peat pads were cut up (Table 2), and the active capitula layer was removed. Peat fractions were harvested before the extraction and prefiltration (125 µm) of pore-water from the *Sphagnum* (about 500mL) for both the surface and depth layers (respectively 5 to 10cm, and 10 to 15cm below the active capitula layer). In Canada, "surface" corresponds to 5 to 15cm, and "depth" to 15 to 25cm. Samples from Canada, Finland, la Guette and Frasne were kept at 4°C and sent within 4 days in Rennes, and processed immediately upon arrival. Samples from the Amsterdam Isle were frozen on the field due to the length of the field campaign and kept at -20°C until processing.

Physico-chemical parameters

Conductivity, pH, and O_2 were measured for each depth during the field sessions at la Guette, Frasne and Amsterdam Isle. For each depth, 250 mL of pore water was prefilterd at 125µm and kept frozen at -20°C for DOC and anions (nitrate, sulfate and acetate) analyses at the OSUR and Ecobio analytic platforms, using a Bioritech DOC Analyser and a Dionex Analyzer respectively.

In addition, peat Fresh/Dry matter ratio was estimated as the difference between the fresh and the oven-dried peat mass at 105°C for 72 hours.

Peat samples processing for the extraction of the viruses from the peat matrix

For each peat sample, roots were removed, and a define weight of wet peat was extracted (Table 3), dived into potassium citrate buffer, vortexed and left over-night at 4°C. The potassium citrate and peat mix was three times sonicated for 1min (bath with ice), and shaked for 30s (Williamson et al., 2005). The mix was filtered on a folded standard fiber filter (Fisherbrand) and a 2mL aliquot was collected.

Table 3 Fresh weight of wet peat matrix and volume of citrate buffer used for the extraction of the viral-like particles and of the prokaryotes for each sample

	Weight of wet peat (g)	Volume of buffer (mL)	Fresh:Dry matter ratio
Finlande (fen)	3	40	13.66
Finlande (bog)	3	40	16.58
Canada	20	200	7.31
La Guette	20	200	20.39
Frasne	20	200	24.52
Amstersam Isle	22	200	21.72

Glutaraldehyde was added to the aliquot (2% final concentration), which was then flash frozen in liquid nitrogen, and kept at -80°C until used for cytometry. PEGylation procedure was applied to the rest of the filtrate (Colombet et al., 2007) for the production of viral concentrates to extract the viral DNA (Chapter 1.A). Briefly, after a one week incubation in PEG, the mix PEG+supernatant was centriguged at 10000g and 4°C for 30min. The supernatant was carefully discarded and the PEG pellet was washed with KCl 1M and centrifuged at 12000g and 4°C for 10 minutes. When a too high volume of KCl was required to wash the PEG pellet, viral particles were further concentrated on Vivacon 500 columns (Sartorius) to a volume of ~ 2mL.

Pore-water samples processing for the extraction of viruses from the pore-water

Pore-water samples pre-filtered during the field session (125µm) were centrifuged (4°C, 20min, 10 000g, Beckman Coulter), and the supernatant was filtered on folded standard fiber filters (Fisherbrand). For cytometry, a 2mL aliquot was collected after this step and treated as described previously (Chapter 1). Viral concentrates were produced using the PEGylation procedure (Colombet et al., 2007) and further concentrated on a Vivacon 500 column (Sartorius).

Prokaryote and viral-like particles abundances

Aliquot stored in Glutaraldehyde (2ml, 2%) were quickly thawed and diluted between 40 and 80x in Tris-EDTA buffer (pH=8, Sigma-Aldricht). SYBR-green was added to a 1X final concentration. Samples were either incubated at room temperature for 30 min (prokaryotes), or heated at 80°C for 15min (viral particles), and then cooled at 4°C during 4 minutes. Counts were obtained by flow cytometry (BD Accuri C6). For each sample, one unlabeled (MilliQ water) and three labeled (SYBR-green) subsamples were counted. Unlabelled samples were used to determine the background noise.

Abundance of viral particles and prokaryotes per gram of dry peat were calculated using the equation: VPA or PA g^{-1} of dry peat= (VPA or PA mL-1)x(volume of buffer / g of wet peat)x(gram of wet peat / gram of dry peat).

Metaviromes production

After the production of the viral concentrates, peat (identified with a "p" at the end of their name) and pore-water (identified with a "w") viral concentrates were treated in the same manner (Table 2). Viral concentrates were filtered through a 0.22 µm filter (Minisart, Sartorius). The filtrate was diluted 5x in ultrapure H_2O (Sigma). Remaining non-viral DNA was digested with 10 U DNAse RQ1 RNAse free (Promega) at 37 °C for 1h. Viral DNA was extracted as previously described (Quaiser et al., 2015). DNA quality was checked with the High Sensitivity DNA kit on a 2100 Bioanalyzer (Agilent). Whole genome amplification (WGA) was applied in triplicate for each sample using GenomiPhi V2 (GE Healthcare) for 90min following manufacturers' instructions and the triplicates were pooled. Illumina sequencing was performed at the "Functional and Environmental Genomics" platform (OSU, Rennes, France). Filtering tools and filters were used to ensure the highest sequence quality and to remove artificial replicates of sequences. The quality of the metaviromes was assessed and ensured by respectively determining and removing the amount of rRNA and tRNA sequences using respectively Meta_RNA that identifies SSU and LSU rRNAs from the three kingdoms (Huang et al., 2009) and tRNAscan-SE (Lowe and Eddy, 1997). Due to technical failures, some metaviromes are still not sequenced (Table 2, "*in progress*" samples).

Metaviromes analysis

Metaviromes sequences were uploaded on the metavir webserver (Roux et al., 2011; Roux et al., 2014) that performed taxonomical assignation of the viral sequences against RefSeqVirus by tblastx (bit-score 50) (Altschul et al., 1997). Assignations of the sequences to the different viral types were retrieved using MEGAN (Huson et al., 2007), and used to build a contingency matrix to study the effect of site and extraction method on the taxonomical diversity of the identified sequences.

Effects of extraction method, depth, dynamic stage and site were tested on the metaviromes from the peat matrix and the pore-water from Finland and Canada using commet (k=33, t=4) (Maillet et al., 2012). As 4 x 33kmers means that the similarity between two sequences is based on 132 out of ~250 nucleotides, preliminary tests had been done with t=2 and t=3 (data not shown), to make sure that increasing t would not be too stringent according to the sequences length. The effect of site on the viral diversity with emphasis on the importance of the *Sphagnum* species was studied through the comparison of the metaviromes from pore-water originating from Canada (3 metaviromes), Finland (3 metaviromes) and 12 metaviromes produced from a French *Sphagnum*-dominated peatland (Chapter 1.A.) which present similar *Sphagnum* species and associated vegetation as peatland from Finland. Sequences from these 18 metaviromes were compared using commet (k=33, t=4). All these analyses were performed on the Galaxy platform (Le Bras et al., 2013).

We used Metavir blast-based comparison (Roux et al., 2011; Roux et al., 2014) (tblastx on metaviromes subsamples composed of 50 000 reads of a 100 nucleotides length) to compare the 31 peatland metaviromes with 67 available metaviromes originating from freshwater, seawater, hypersaline and polar desert ecosystems.

Statistical analyses

The single and combined effects of site and depth on viral-like particles (VPA) and prokaryotes abundances (PA) and on VPR (virus to prokaryotes ratio), were respectively analyzed through Mann-Whitney and Kruskal-Wallis tests. A principal component analysis was applied on the physico-chemical dataset obtained for the peatlands from France (la Guette, Frasne) and Amsterdam Isle to study the variation of the physico-chemistry between sites, and its potential links with the variation of abundance of the biological entities. Site and depth were added as supplementary qualitative data and PA, VPA and VPR were added as supplementary quantitative data. Effects of qualitative data and correlations between variables and supplementary quantitative data and the components of the PCA were obtained using the FactoMineR package (Husson et al., 2015). These analyses were performed using the R software (version 2.14) (R Development Core Team, 2013).

Results

The aim of the study was to determine whether important features of viral ecology, such as viral abundance, VPR and the link between the viral particles abundance and the physico-chemistry were consistent across peatlands representing a gradient of climatic conditions (Table1). Therefore we adapted a method for the extraction of the viral particles from the peat matrix allowing the determination of finer spatial patterns. We also wanted to test the influence of *Sphagnum* species and site on the viral diversity carried in pore-water and peat matrix metaviromes, and to investigate whether peatland viral communities appear specific for this type of ecosystem.

Viral particles and prokaryote abundances

Prokaryote abundance (PA) ranged from 2.30 ± 0.82 10^6 prokaryotes per gram of dry peat at the surface of Amsterdam Isle's peatland, to 3.63 ± 1.35 10^9 at the surface of the Finnish fen (Figure 2). PA differed from site to site and we observed a combined effect of site x depth layer (surface *vs* depth) (Table 4): PA generally increased with depth to the exception of the fen and the bog in Finland.

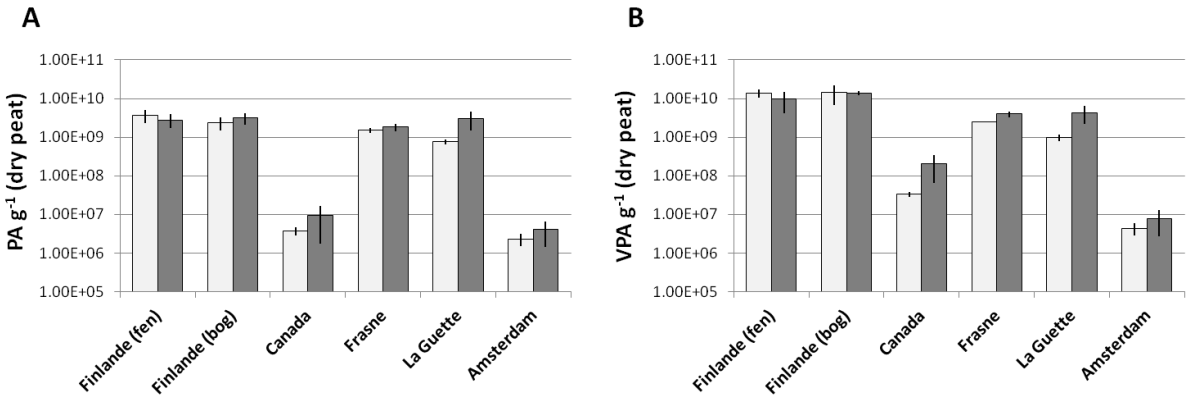


Figure 2 PA (A) and VPA (B) at the surface (white) and depth (B) layer of the 5 studied *Sphagnum*-dominated peatlands.

The same trend was observed for the abundance of the viral-like particles (VPA) (Figure 2). The minimum abundance was observed at the surface of the peatland in Amsterdam Isle ($4.34 \pm 1.47\ 10^{6}$ viral-like particles per gram of dry peat), and the higher values at the surface of the bog in Finland ($1.42 \pm 0.73\ 10^{10}$ viral-like particles per gram of dry peat). In the same way, VPA was significantly influenced by both the site, and the combined effect of site x depth layer (Table 4), with higher VPA at depth, to the exception of the Finnish peatland.

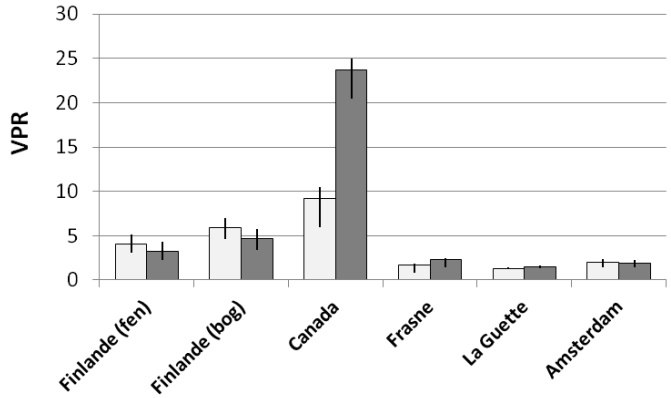


Figure 3 Virus to Prokaryotes Ratio (VPR) at the surface (white) and at depth (grey) of the different peatlands. Error bars represent standard deviation.

Table 4 Effects (p-values) of site and depth on VPA, PA and VPR. Abbreviations: KW: Kruskal-Wallis test, MW: Mann-Whitney test

	KW (site x layer)	KW (site)	MW (layer)
PA g^{-1} dry peat	**1.76E-05**	**2.63E-07**	0.2317
VPA g^{-1} dry peat	**5.97E-06**	**3.13E-08**	0.4833
VPR	**3.00E-05**	**1.99E-07**	0.9179

The virus to prokaryote ratio (VPR) varied from 1.29 ± 0.16 at the surface of La Guette, to 23.65 ± 3.20 at depth of the Canadian peatland. The second highest VPR was recorded at the surface of the bog in Finland (5.90 ± 1.06) (Figure 3). As well as the abundance of biological entities, VPR was strongly influenced by the site and the combined effect of site x depth layer (Table 4).

Link between biological entities abundance and the physico-chemical parameters

In order to detect whether physico-chemical differences drove the distinction between sites, and whether they varied with the abundance of the viral particles and of the prokaryotes, a principal component analysis was performed on the matrix containing values for the physico-chemical parameters measured in the surface and depth samples from La Guette, Frasne and Amsterdam Isle (Table 5). Physico-chemical parameters from Finland samples are not available, and DOC and anions concentrations were not determined for Canada. These samples were selected as all their physico-chemical features for surface and depth were measured on the same devices, allowing a fine spatial pattern and avoiding technical bias.

Table 5 Mean (± SD) values for the physico-chemical parameters measured in the peatlands from Amsterdam Isle, Frasne and La Guette (N=3).

	Layer	pH	Conductivity (µS cm-1)	DOC (mg.L-1)	Chlorure (mg.L-1)	Nitrate (mg.L-1)	Sulfate (mg.L-1)	Fresh :Dry matter
Amsterdam Isle	***Surface***	4.7 (0.2)	68.0 (26.0)	20.7 (2.9)	17.6 (4.0)	2.0 (0.5)	3.2 (1.2)	25.4 (6.4)
	Depth	4.6 (0.2)	73.1 (15.9)	22.5 (7.0)	19.9 (3.4)	1.4 (0.3)	3.6 (2.1)	21.7 (7.4)
Frasne	***Surface***	4.2 (0.2)	41.2 (3.5)	51.2 (1.2)	8.9 (1.1)	1.2 (0.1)	3.0 (0.2)	25.0 (3.1)
	Depth	4.1 (0.1)	39.8 (3.8)	48.4 (1.0)	7.7 (0.4)	1.1 (0.1)	2.7 (0.1)	25.1 (3.6)
La Guette	***Surface***	5.5 (0.0)	69.1	63.4 (21.0)	16.1 (6.4)	0.4 (0.2)	2.1 (0.4)	19.4 (3.2)
	Depth	5.5 (0.0)	69.1	48.3 (18.3)	8.4 (0.7)	0.1 (0.1)	1.3 (0.1)	22.2 (2.0)

Surface: 5 to 10 cm below the green capitula layer, Depth: 10 to 15cm below the green capitula layer. Abbreviation: DOC=Dissolved Organic Carbon.

A PCA was performed on the matrix containing the physico-chemical values of each sample. This matrix was completed with the addition of site and layer (surface *vs* depth) as supplementary qualitative information, and VPA, PA and VPR as supplementary quantitative data using the R package FactoMineR (Husson et al., 2015).

The first component of the PCA carried 42.07% of the variation of the physico-chemical features. Nitrate, sulfate and chloride concentrations were positively correlated to this first component, while DOC concentration and pH were negatively correlated. VPR appeared positively correlated to the first component, to the contrary of PA and VPA which increased with DOC. La Guette and Amsterdam Isle samples were significantly distinct along this first component (Figure 4), with higher DOC in La Guette samples, and higher anion concentrations in the samples from Amsterdam Island. Depth did not seem to be associated to significant changes in the physico-chemistry of the samples along the 3 first components. Hierarchical clustering of the samples based on the 2 first components of the PCA (73.03 % of the dataset variation) indicates that the surface and depth samples from the same peat column tend to be more similar in Frasne and Amsterdam (Figure 5).

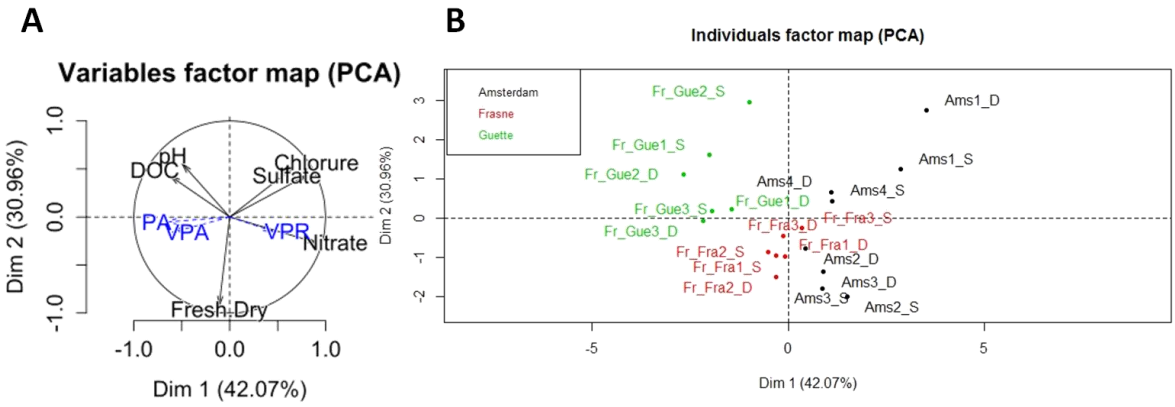


Figure 4 Projection of the variables (A) and of the samples (B) from Amsterdam Isle, la Guette and Frasne peatlands depending on the variation (principal component analysis) of their physico-chemistry to determine what factors influenced most the variation of the physico-chemistry and to detect potential co-variation with viral particles and prokaryote abundance. Number and letter next to the name of the peatland represent the peat column number (sampling point) and the sampled layer; S: surface, D: depth.

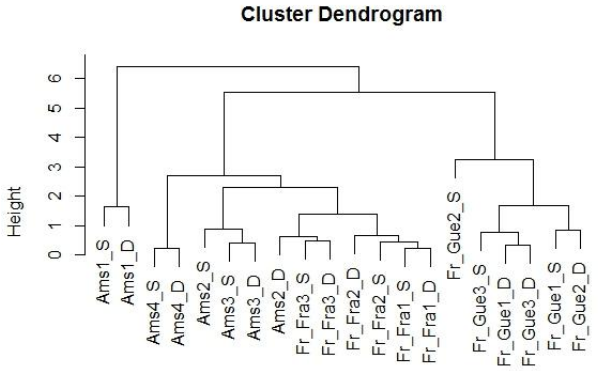


Figure 5 Hierarchical clustering of the peat samples from Amsterdam Isle (Ams), Frasne (Fra) and la Guette (Gue) based on the principal component analysis performed on their physico-chemical attributes. Numbers represent the peat column (sampling point) from which the sample originates, S and D represent the sampled layer (S: surface, D: depth).

Viral Diversity in the pore-water metaviromes

In order to compare the diversity of the viral communities from the different peatlands, we produced viral concentrates from the pore-water from Finland and Canada and sequenced the viral DNA using Illumina technology. We obtained six new peatland metaviromes that include 3 metaviromes from the Finnish peatland (2 fen and 1 bog) (Table 6) and 3 from the Canadian peatland (2 metaviromes from depth and one from the surface) (Table 7). This represent 2 150 789 sequences for the pore-water metaviromes from Finland, and 2 044 939 for these from Canada. Sequences from these six metaviromes were compared to RefSeqVirus by tBLASTx (bit-score 50) (Altschul et al., 1997) on Metavir (Roux et al., 2011; Roux et al., 2014). Only very low proportions of the sequences matched to referenced viral genomes representing respectively 5.22 ± 1.00 % of the viral sequences from Finland and 5.49 ± 0.45 % for the viral sequences from Canada.

Table 6 Informations about the metaviromes from the pore-water from Finland

Country	**Finland**		
Type	**Pore-water metaviromes**		
Origin	***Fen***	***Fen***	***Bog***
Name	**vFin_Fen1w**	**vFin_Fen3w**	**vFin_Bog3w**
Sequencing type	Illumina	Illumina	Illumina
Metavir ID	6079	6080	6081
N° of sequences	727 150	733 963	689 685
% assigned sequences	6.1	4.69	4.88

Table 7 Informations about the metaviromes from the pore-water from Canada

Country	**Canada**		
Type	**Pore-water metaviromes**		
Area	**Natural**		
Origin	***Depth***	***Surface***	***Depth***
Name	**vCan_Nat1_Dw**	**vCan_Nat2_Sw**	**vCan_Nat2_Dw**
Sequencing type	Illumina	Illumina	Illumina
Metavir ID	6094	6095	6096
N° of sequences	885 598	606 800	552 541
% assigned sequences	4.96	5.94	5.89

Taxonomic assignations of the sequences (representing 112 343 sequences from Finland and 114 448 sequences from Canada) were analysed with MEGAN (Huson et al., 2007). Among the sequences assigned, some viral types were highly represented, such as ssDNA viruses (Microviridae, Circoviridae and Other ssDNA viruses) representing 94.59 % of the assigned sequences from Finland and 32.88% of the total assignation of the sequences from Canada (Figure 6). Other types were poorly represented, such as the environmental viruses, mainly represented by the genome of the "chimera" from the Boiling Spring Lake (0.10 % and 0.05% of the assigned sequences from Finland and Canada respectively) (Figure 6). Considering the affiliation to known viral types, pore-water viromes from Finland and Canada appear highly distinct, with more dsDNA viruses in Canada (total of Caudovirales, Mimiviridae, Phycodnaviridae and Other dsDNA viruses representing 63.86% of the assigned sequences) and a large dominance of referenced ssDNA sequences in Finland. For the dsDNA viruses in Canada, metaviromes from the deeper layer seem to present a different pattern than this from the surface, with more Caudovirales (prokaryotic hosts), and less Phycodnaviridae (eukaryotic hosts). For the assigned viral sequences from Finland, ssDNA viruses Circoviridae (eukaryotic hosts) appear to be important in the fen and the bog while the Microviridae (bacterial hosts) seem more abundant in the fen.

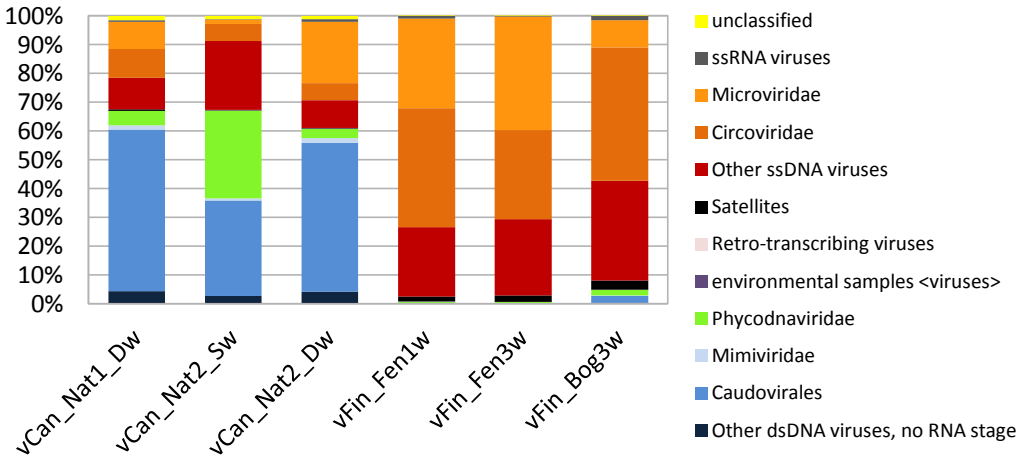


Figure 6 Proportions of the different viral types represented in the peatland metaviromes from the pore-water from Canada and Finland (tblastx against RefSeqVirus, bit-score 50). In the metaviromes name: Can= Canada, Fin=Finland, number represent the peat column number.

Viral diversity in the peat matrix metaviromes

In order to analyze the spatial patterns of viral diversity extracted of the peat matrix, 13 metaviromes were produced from surface and depth layers from Finland and Canada samples (Table 8, Table 9). In total 7 683 821 sequences from Finland, and 3 254 306 from Canada were included in the analysis.

Table 8 Main properties of the metaviromes produced from the viruses from the peat matrix from Finland.

Country	**Finland**								
Type	**Peat metaviromes**								
Dynamic stage	**Fen**				**Bog**				
Depth	**Surface**		**Depth**		**Surface**			**Depth**	
Name	**vFin_Fen2_Sp**	**vFin_Fen3_Sp**	**vFin_Fen2_Dp**	**vFin_Fen3_Dp**	**vFin_Bog1_Sp**	**vFin_Bog2_Sp**	**vFin_Bog3_Sp**	**vFin_Bog2_Dp**	**vFin_Bog3_Dp**
Sequencing	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina
Metavir ID	6088	6089	6092	6093	6082	6084	6087	6090	6091
N° of sequences	600 210	709 317	692 460	635 275	659 996	774 239	1 902 013	930 535	779 776
% assigned	9.99	4.11	7	12.62	11.13	11.09	8.99	4.26	7.13

The taxonomical assignations of the sequences of the peat matrix viruses were analyzed by tblastx against the RefSeqVirus database (Metavir) (Roux et al., 2011; Roux et al., 2014) (bit-score 50) (Figure 7). The viral sequences from the peat matrix from Finland were proportionally more similar to referenced viral genomes than previously observed for the pore-water viral diversity (8.48 ± 2.87 % of the sequences matched to a reference) whereas the proportion of assigned sequences from Canada remained low (5.79 ± 0.94 %). On the whole, about 651 588 and 188 424 sequences from the peat-matrix viruses from respectively Finland and Canada were assigned to a referenced viral genome.

Table 9 Main properties of the metaviromes produced from the peat matrix from Canada.

Country	Canada			
Type	Peat metaviromes			
Area	Natural			
Peat column	1		2	
Depth	Surface	Depth	Surface	Depth
Name	vCan_Nat1_Sp	vCan_Nat1_Dp	vCan_Nat2_Sp	vCan_Nat2_Dp
Sequencing	Illumina	Illumina	Illumina	Illumina
Metavir ID	6100	6101	6102	6103
N° of sequences	784 862	813 320	825 206	830 918
% assigned	7.24	5.99	5.11	4.82

As for the metaviromes from the pore-water, the diversity of the viral communities from the peat matrix observed through the assignation to referenced viral genomes varied between samples (Figure 7).

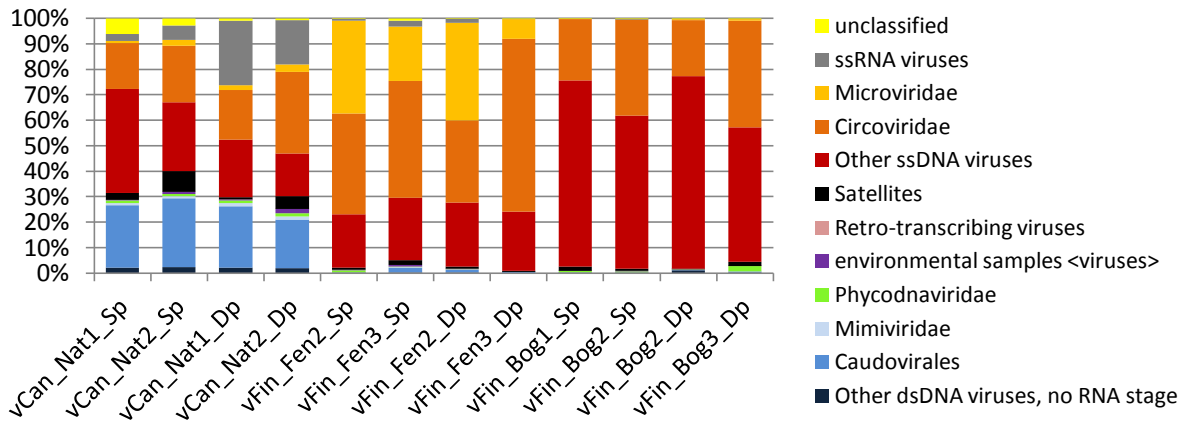


Figure 7 Proportions of the different viral types represented in the metaviromes produced from the peat matrix in the surface and in depth from the peatlands from Canada and Finland (tblastx against RefSeqVirus, bit-score 50).

Higher proportions of ssRNA viruses were observed in Canada (12.72 %), more especially at depth (21.30 %). In peat matrix as in pore-water metaviromes from Finland, we observed again a higher dominance of ssDNA viruses (96.44%) compared with Canada (51.78 %), and a difference between fen and bog, with more Microviridae in the fen (25.95 %) compared with the bog (2.16 %). However, in Finland, no surface/depth pattern was observed when looking at the taxonomical assignations.

Integrated analysis of the metaviromes based on taxonomical assignations

Whereas the proportions of assigned sequences were inferior to 10% of the sequences from each metaviromes, this already represented a high number of matches to known viral genomes as a consequence from the high amounts of sequence produced for each metavirome. Our aim was to

determine the variation of the taxonomic affiliations depending on site (Finland *vs* Canada), sampled depth (surface *vs* depth), and between pore-water *vs* peat matrix.

A contingency matrix was built from the assignations to the most abundant viral types for all the peat matrix and pore-water metaviromes from Finland and Canada. These viral types include, Caudovirales (dsDNA, prokaryotic hosts), Mimiviridae (dsDNA, protists hosts) and Phycodnaviridae (dsDNA, eukaryotic hosts), Circoviridae (ssDNA, eukaryotic hosts) and Microviridae (ssDNA, bacterial hosts). A principal component analysis was performed on the dataset. Site, depth and origin (peat matrix vs pore-water) were added as supplementary information (Figure 8). The first component carries 42.09 % of the datasets variation, and was associated to the difference between Canada and Finland, with more referenced Mimiviridae, Other dsDNA viruses, Caudovirales and retro-transcribing viruses in Canada, and more referenced Circoviridae and Other ssDNA viruses in Finland (Figure 8). The second component carries 18.73 % of the dataset variation and is associated to the difference between pore-water and peat-matrix viral communities, with more Phycodnaviridae in the pore-water communities, and more referenced environmental viruses, satellites and ssRNA viruses (Scleropthora macrospora virus A) in the peat-matrix samples. The third component was associated to the difference between the fen and bog samples from Finland, with more Microviridae matches in the fen, and more "other ssDNA" viruses in the bog.

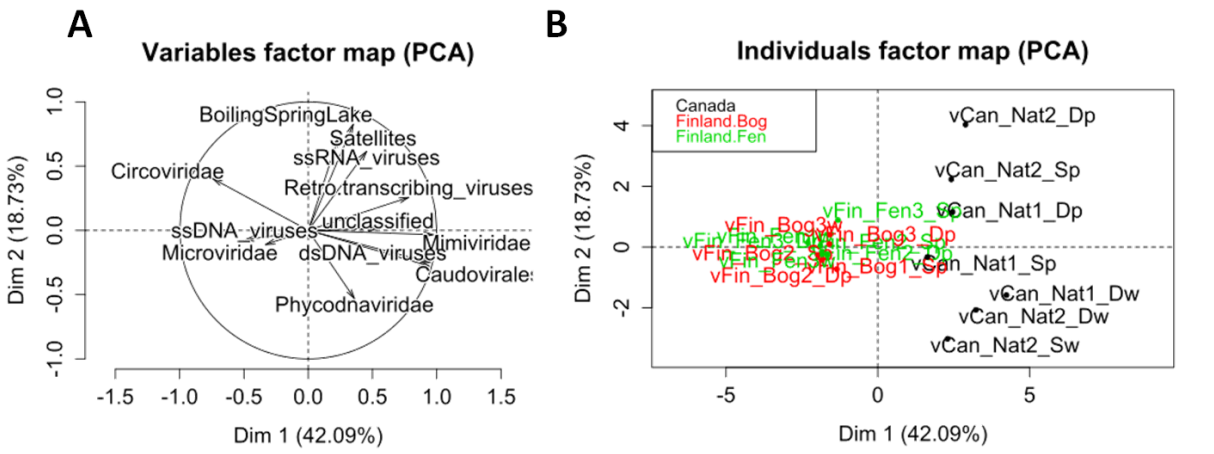


Figure 8 Variable factor map (A) (variation of abundance of the viral types along the two first components of the PCA) and metaviromes projection (B) depending on the site.

This analysis revealed that the site (Canada/Finland), followed by peat-matrix/pore-water differentiation had the strongest impact on the metaviromes similarities/differences. However, no difference in the proportions of referenced viral genomes associated with the sampled depth was detected.

Metavirome similarity analysis independent of taxonomic affiliation

As mentioned previously, results obtained from the taxonomical assignations correspond to a low proportion of the sequences of the pore-water and peat matrix metaviromes. Further comparisons of the sequences, independent of taxonomical assignations, and taking into account the whole diversity carried in the metaviromes were performed to investigate the effect of depth, *Sphagnum* species (fen *vs* bog), site origin (Finland *vs* Canada), sample origin (peat-matrix *vs* pore-water), and ecosystem on the diversity of the viral communities.

Effect of sample origin, depth and site origin on the diversity of the metaviromes

In order to determine metaviromes similarities depending on origin and spatial variables, we compared the 19 metaviromes from the pore-water and the peat matrix from Finland and Canada using commet (commet, t=4, k=33) (Maillet et al., 2012). Commet output is a distance matrix composed of the normalized percentage of similarity between pairwise metaviromes. This similarity is based on the amount of shared "similar" sequences between each metaviromes and normalized to the total number of sequences in the respective metaviromes. Two sequences are considered as similar if they present at least 4 identical kmers of 33 nucleotides.

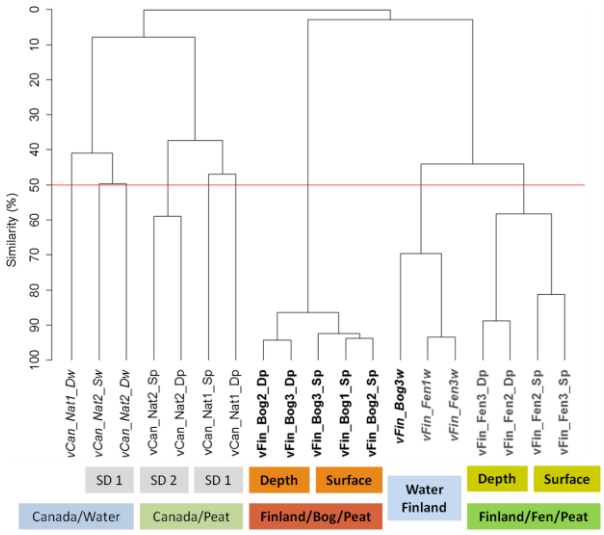


Figure 9 Hierarchical clustering based on the normalized average similarity between the 19 metaviromes from Finland (Bold) and Canada (normal) depending on the extraction method (pore-water in italic), depth layer, and *Sphagnum* species (in Finland, *Sphagnum fallax* in the fen (bold grey), *Sphagnum magellanicum* in the bog(bold black)) (commet, t=4, k=33). For Canada: SD1, SD2 indicates Surface (S) and Depth (D) and sampling point (1,2).

According to the hierarchical clustering based on the distance matrix (normalized percentage of similarity), the viral communities appear grouped depending on the site they originate (Finland or Canada) (Figure 9). In Canada, metaviromes cluster in dependence on sample origin, with pore-water metaviromes on one-hand (average similarity between pore-water metaviromes from Canada: 44.06 ± 3.94 %) and peat matrix metaviromes on the other hand (average peat matrix group similarity:

47.67 ± 7.46 %). In the pore-water and peat matrix groups, the surface and depth samples from a same sampling point share higher similarities.

In Finland, the metaviromes from the peat matrix clustered distinctly according to the spatial scale, and present clear differences between bog (average group similarity: 91.54 ± 2.78 %) and fen (77.24 ± 9.46 %), with differences within each group between surface and depth communities (Figure 9). The fen peat-matrix group clusters, albeit with lower similarity, with the pore-water metaviromes group (both groups similarity 70.42 ± 12.92 %) (Figure 9).

This pattern of similarity hints for an effect of site (Finland/Canada, fen/bog), habitat (fen/bog), depth and sample origin (peat-matrix/pore-water) on the structure of the viral community.

Effect of Sphagnum species and site on the diversity of pore-water metaviromes

One important aim of this study was to investigate the influence of the *Sphagnum* species on the viral community. In this end, we added 12 pore-water metaviromes from a French *Sphagnum*-dominated peatland (Chapter 1.A.) to the analysis. These 12 metaviromes originate from the fen (N=6) and the bog (N=6), which are represented by the same *Sphagnum* species as in Finland (respectively *Sphagnum fallax* in the fen and *Sphagnum magellanicum* in the bog). These viral communities were sampled during different seasons in 2011 and 2012 and presented a yearly cycle, allowing a good coverage of the viral diversity present at this site. As viral diversity appears to be influenced by site and sample origin (Figure 9), we compared the viral diversity of this 12 pore-water metaviromes from France with the diversity carried in the pore-water metaviromes from Finland and Canada. The 18 pore-water metaviromes were compared with commet (k=33, t=4) independently from taxonomical assignations.

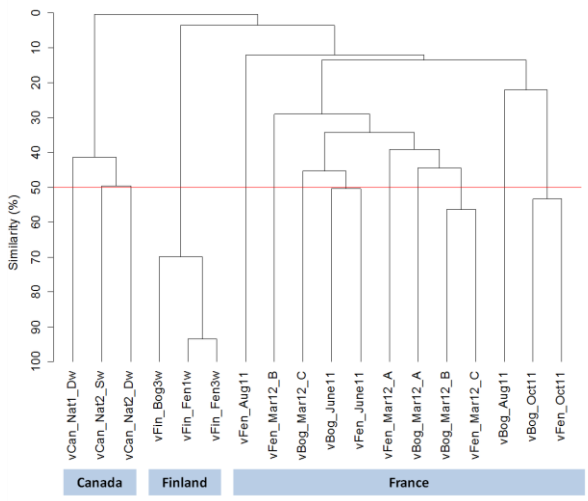


Figure 10 Hierarchical clustering of the metaviromes based on normalized similarity of the pore-water from Finland, Canada and France (commet, t=4, k=33).

The viral communities from the three sites cluster into distinct groups (Figure 10). Average normalized intra-site similarity is 77.73 ± 13.64 % for the metaviromes from Finland, 44.06 ± 3.94 % for the metaviromes from Canada, and 23.86 ± 13.81 % for the metaviromes from France. The metaviromes from France share 3.63 ± 3.01 % of similarity with the metaviromes from Finland, and 0.09 ± 0.06 % with the metaviromes from Canada, while the metaviromes from Canada and Finland share 1.68 ± 1.73 % of similarity (Figure 9).

As the 12 metaviromes from France present a temporal variation, we then looked for more precision in the similarity between the three sites. The similarity between the viral communities from Finland and Canada ranges from 0.37 % to 5.10 %. The similarity between the metavirome from the bog of Finland (vFin_Bog3w) and the metaviromes from France ranges from 0.75 % (vBog_Oct11) to 5.90 % (v_Bog_Mar12_B). For the fen metaviromes from Finland (vFin_Fen1w, vFin_Fen3w) similarity with the metaviromes from France (vBog_Mar12_B, vFen_Mar12_A and vFen_Aug11) reaches more than 8 %. In contrast, similarity between the metaviromes from France and Canada never exceeded 0.22%, to one exception reaching 1.4 % (vBog_Aug11; vCan_Nat2_Sw). This result suggests a *Sphagnum* species specificity of the viral communities between disconnected biotopes.

Are peatland viral communities specific for their ecosystem?

Finally, the production of metaviromes originating from geographically distant *Sphagnum*-dominated peatlands allowed the investigation of the specificity of the viral communities for these ecosystems.

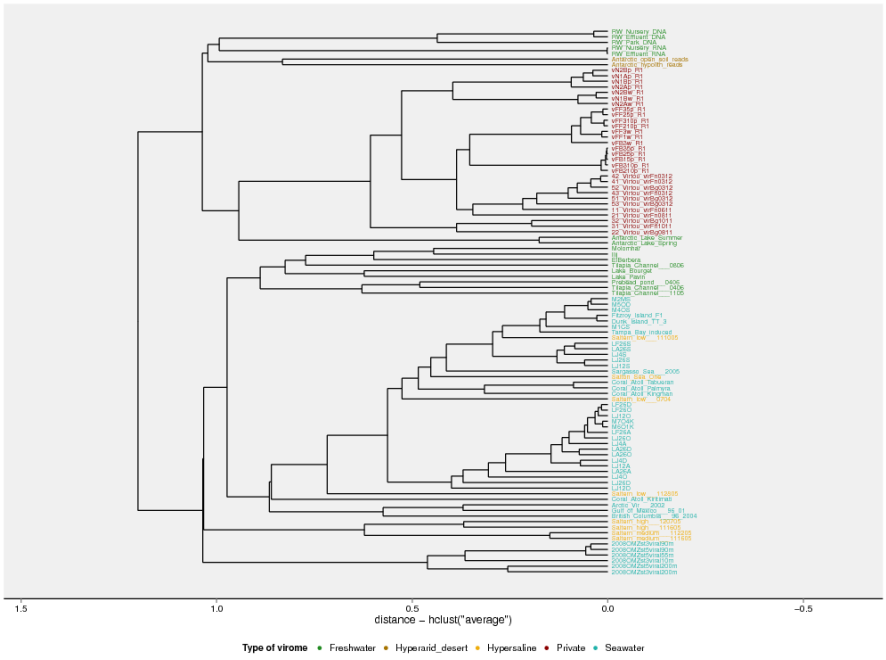


Figure 11: Cross-ecosystems comparisons of the metaviromes from peatlands (red) with 67 metaviromes originating from freshwater (green), seawater (blue), hypersaline (yellow) and polar soils (brown) ecosystems. Comparisons are based on tblastx between subsamples composed of 50 000 sequences of 100 nucleotides.

In this end, the 31 viral communities from the peat matrix and pore-water samples from Finland, Canada and France were compared by tblastx on Metavir with 67 metaviromes representing different ecosystems spanning from freshwater, seawater, polar desert to hypersaline biotopes (Roux et al., 2011; Roux et al., 2014).The viral communities from peatlands form a distinct group (Figure 11) indicating a specific viral diversity for these ecosystems.

Discussion

The aim of this work was to improve the knowledge of viral ecology in *Sphagnum*-dominated peatlands through the comparison of different peatlands at a large geographical scale of latitudes and longitudes. Comparing different peatlands (sites), dynamic stages, depth layers (acrotelm and surface of mesotelm) and pore-water versus peat-matrix viral communities, we searched to identify fine spatial differences in viral abundance and diversity as well as similar patterns across disconnected peatlands as a consequence of an expected similar functioning of the virus/hosts interactions in the Sphagnosphere.

The comparison of the viral communities from the pore-water and the peat matrix sheds light onto the functioning of the viral compartment in peatlands, and highlights the structuration of peat matrix metaviromes from a fine (depth) to a large (site) spatial scale. This progress in our understanding of viral communities was possible thanks to the production of 31 metaviromes from *Sphagnum*-dominated peatlands. While only few soils metaviromes have been published to date (Zablocki et al., 2014; Adriaenssens et al., 2015), this effort represents the second largest collection of environmental metaviromes after the large 43 marine metaviromes collection produced following *Tara Ocean* sampling campaign (Hurwitz and Sullivan, 2013; Brum et al., 2015b).

Abundances of viral-like particles, and link with the physico-chemistry

Counts represent a good approach to compare viral and prokaryote abundances (respectively VPA and PA) in different ecosystems. Given the variability of VPA and PA in the spatio-temporal dynamics of the temperate *Sphagnum*-dominated peatlands (Chapter 1.A.) and across soils (Williamson et al., 2005), we expected a high variability of these parameters between the 5 peatlands. Peat-matrix counts indicate that VPA between sites is highly variable with a 3271-fold range between the minimal and the maximal VPA. PA appears highly variable too with a 1572-fold range between the minimum PA recorded in Amsterdam Isle and the maximum in the Finnish peatland (Figure 2). This variability is higher than presently acknowledged in soils (Srinivasiah et al., 2008) and confirms the influence of site on peat matrix VPA and PA. Within the peatland of Finland, VPA and PA varied little between the fen and the bog (Figure 2), in accordance with the trend observed in the temporal dynamics of the pore-water VPA and PA from the temperate peatland (Chapter 1.A.).

We also observe higher VPA and PA at depth in nearly all sites, and using a finer measure of the physico-chemical parameters at the two depths of three peatlands, we observe a positive relation between the concentration of DOC and VPA and PA (Figure 4). This tends to confirm that depth is associated to a higher prokaryotic activity (Chapter 2) and subsequent higher viral activity and abundance in peatlands (Chapter 1.A.). This also confirms that the functioning of viral ecology is different at depth compared with the surface, probably as a consequence of different environmental conditions and host activity shaped by the water-table fluctuations (Chapter 1.A.; Chapter 2).

In parallel, the range of VPR observed in this chapter is by one order of magnitude larger than this observed in the temporal dynamics (Chapter 1.A.) and while VPR never exceeded 6 in the seasonal dynamics (Chapter 1.A.), it ranges from 1 to 23 between the different peatlands (that is using the extraction method from the peat) (Figure 3). This indicates that virions can be thighly attached to *Sphagnum* cells and not observed in the pore water, and suggests that the pore-water and peat matrix viral diversity could present differences.

Taxonomical assignations suggest that site and sample origin (pore-water vs peat matrix) affect viral diversity

In Finland and Canada, the comparison of the pore-water and peat matrix viral communities, dependant or not of the taxonomical assignations of the viral sequences yielded coherent and interesting results in regard to the functioning of the viruses and their host activity in relation to environmental conditions.

Using the variation of taxonomical assignations between the pore-water and peat matrix metaviromes from Finland and Canada, even if low proportions of metaviromes sequences matched to referenced viral genomes, we observed that site (Canada vs Finland) was the factor that affected most the proportion of the viral types, followed by the metavirome origin (pore-water vs peat matrix) (Figure 8). In Finland, we also distinguished a low difference between fen and bog metaviromes, but we did not detect any difference in the referenced diversity of the surface and depth metaviromes. These results tend to confirm the spatial variability between peatlands viral communities, and the effect of sample origin as was suggested by the range of VPR. Distinction between fen and bog viral communities was interesting as it was not observed in the pore-water samples from the French temperate peatland (Chapter 1.A.). This indicates that peat-matrix metaviromes harbor a viral diversity that emphasizes more the spatial differences.

The comparison of the whole diversity carried in the metaviromes from Finland and Canada sheds light onto the functioning of the viral and prokaryote compartment depending on the effect of site, dynamic stage, depth and origin

Further comparison of the viral communities from Finland and Canada, based on the whole diversity carried in the metaviromes confirmed the distinction between the diversity from the two sites (Figure 9).

In Finland, the bog and fen viral communities from the peat matrix clustered into two distinct groups (Figure 9). As the viral diversity depends on host diversity, this tends to confirm that fen and bog harbor taxonomically different host community (Opelt et al., 2007; Bragina et al., 2012, Chapter 2). Then, within each group, we observe a difference of diversity between the peat matrix communities from surface and depth. Considering that taxonomical assignations of the peat matrix metaviromes from Finland varied with the dynamic stage and not with depth (Figure 8), this strongly suggests that within fen and bog, the difference between surface and depth communities results from a similar genetical background which activity is selected by depth-related environmental conditions (Lin et al., 2014; Tveit et al., 2013; Chapter 2). This diversity pattern in Finland represents probably the major result of this project, as it confirms the propositions made in Chapter 1.A., and in Chapter 2 about the functioning of the viral and prokaryotic communities in *Sphagnum*-dominated peatlands, and indicates that a fine sampling of viral diversity in soil ecosystems represent a powerful tool to explore the spatial patterns structuring hosts diversity and transcriptional activity in soil matrixes.

In addition, while the peat matrix metaviromes from the bog in Finland were highly similar and presented a surface/depth pattern, the pore-water metaviromes from the fen and the bog were more similar to the peat matrix metaviromes from the fen (Figure 9). This observation is coherent with the nature of the dynamic stages. The fen receives water inputs from the site water influxes and from the precipitations and is therefore more influenced by the water-table level than the bog which is only rain-fed (Rydin and Jeglum, 2006). The higher similarity between fen peat-matrix communities and the fen and bog pore-water (Figure 8) communities emphasizes the role of the water inputs, water logging and water table level on the diversity of the pore-water viral communities suggested in Chapter 1.A.

Influence of the water-table level on the structure of the viral and active microbial communities can further explain the pattern of the viral diversity in Canada. To the contrary of Finland, the pore-water and peat matrix viral diversity were clearly distinct in Canada, and for both types of metaviromes, the viral communities from the same sampling point were more similar than the viral communities from the same depth (Figure 9). In Canada, at the time of sampling, the water-table level was much below

the "depth" of the analyzed peat columns (-43 cm). In addition, the variation of the physico-chemical parameters in the peatlands from Frasne (were the water table was above the surface layer) and Amsterdam Isle (water-table below the depth layer) was more important between the different sampling points than between the surface and depth samples (Figure 5). Thus, when the water-table is too low to imbibate the peat column, hosts activity sheltered in *Sphagnum* mosses (either attached to the peat matrix or in the pore water of the hyalocytes) seems to be less structured by depth (oxic/anoxic gradient), and more structured by micro-scale environmental conditions and spatial (horizontal) heterogeneity (Parry et al., 2014). This high similarity between the different layers of a same sampling point under dry conditions tends to confirm the role of the *Sphagnum* mosses as a micro-habitat.

It is indeed acknowledged that *Sphagnum* mosses harbours prokaryotic communities at the surface of their leaves as well as inside their hyalocytes (Bragina et al., 2012). The constraint to enter the hyalocytes is the possibility for a biological entity to enter the pore of the *Sphagnum* cells (Gilbert and Mitchell, 2006). Interestingly, in the pore-water metaviromes from Canada and Finland, we observed higher proportions of assigned sequences to Phycodnaviridae (Figure 6), a viral family which is represented by large viruses (such as Emiliana Huxleyi virus 86) that contain large genomes (King et al., 2011). Futhermore, Phycodnaviridae hosts are eukaryotes (algae and animals) that are bigger entities than prokaryotes. We can hypothesize that the difference of viral diversity between water-pore and peat-matrix communities is influenced by the possibility for the hosts, and the viruses to enter or not in the *Sphagnum* pores.

Finally, similarity between pore-water and peat matrix viral diversity seems to vary with the *Sphagnum* species. Indeed, the pore-water and peat-matrix metaviromes were distinct in the bog profiles from Finland (*Sphagnum magellanicum*) and Canada (*Sphagnum fuscum*) (Figure 9). The very high VPR recorded in Canada and the second highest VPR observed in the bog of Finland compared with the lower VPR from France and Amsterdam Isle (Figure 3) could be explained by a higher attachement of the virions to *Sphagnum fuscum* or *magellanicum* representing bog profiles.

Comparison of the peatlands metaviromes from Finland, Canada and France and the functioning of the Sphagnosphere at large geographical scales

The role of the *Sphagnum* as micro-habitat further hints for specific interactions within the *Sphagnum* mosses as in the context of the "Sphagnosphere" (Jassey, 2011). Each *Sphagnum* species has a specific niche (pH, vegetation…) (Rydin and Jeglum, 2006) and metagenomic analyses combined with FISH observation proposed that the structure of the prokaryote communities was specific for respectively S*phagnum magellanicum* and *Sphagnum fallax* (Opelt et al., 2007; Bragina et al., 2012)

even when comparing distant peatlands. One hypothesis was thus that viral communities from similar *Sphagnum* species but distant *Sphagnum* populations would be more similar than viral communities originating from peatlands dominated by different *Sphagnum* species: i.e. that we could observe related viral types between distant ecosystems as observed in Roux et al. (2015). However, it was also possible, due to the rapid and complex evolution of viral genomes (Krupovic 2011), and to the rapid adaptation of prokaryote communities that such similarity was not maintained since these ecosystems are disconnected.

The higher similarity (up to 11 %) of viral communities from the peatlands of France and Finland (Figure 10) that present similar *Sphagnum* species and associated plants communities tends to validate the first hypothesis, but questions the paradigm of rapid viral evolution, in the same manner as do the increasing body of evidence of annual recurrence of viral communities in the oceans (Tucker et al., 2011; Chow and Fuhrman, 2012; Pagarete et al., 2013; Brum et al., 2015a). For now, we lack replicates of very distant peatlands (Amsterdam Isle) or of sites presenting similar *Sphagnum* species (Frasne: *Sphagnum fallax*) to discuss more precisely the dissimilarity between the viral communities from Canada, and these from Finland and France.

None-the-less, the specificity of the viral communities from geographically distant peatlands compared with the viral communities from other ecosystems (Figure 11) reinforces the idea that the viral and therefore microbial communities from *Sphagnum*-dominated peatlands present common features adapted to biotic and abiotic interactions within the *Sphagnum* mosses (Chapter 2).

Conclusion

The present study highlights a higher variation of peat matrix viral abundance in peatlands than presently acknowledged in soils. Pore-water and peat matrix viral diversity also varies greatly and the viral communities appear structured at scales ranging from the sampled layer and *Sphagnum* species to the sampled peatlands. Analysis of the viral diversity in two different stages and at two different depths of a single peatland tends to confirm that each *Sphagnum* type harbors a specific host community, and that surface and depth communities share a similar genetical background which expression differs with depth. This depth gradient is very likely associated with the oxic-anoxic conditions generated by the water-table level fluctuations, and the draw-down of the water-table seems to increase the spatial heterogeneity of the viral diversity. Finally, peatlands viral communities are specific for their ecosystem hinting for a specific functioning of the *Sphagnum* host communities, adapted to the Sphagnosphere.

Acknowledgements

We thank Line Rochefort, Claire Boismenu, Nathalie Van der Putten and Harri Vasander for the peat sampling in Bois-des-Bel, Amsterdam Isle and Lakkasuo. We thank Fatima Laggoun and Daniel Gilbert for their assistance in the sampling of the peatlands from La Guette and Frasne. We thank the members of the 'Environmental and Functional Genomics Platform' (Biogenouest Genomics) for sequencing and the molecular ecology facility (UMR 6553 ECOBIO) for support. We thank Patricia Madec, Nathalie Le Bris and Sophie Coudouel (OSUR Rennes) for technical assistance.

Références du Chapitre 3

Adriaenssens, E. M., Van Zyl, L., De Maayer, P., Rubagotti, E., Rybicki, E., Tuffin, M., and Cowan, D. A. 2015. Metagenomic analysis of the viral community in Namib Desert hypoliths. *Environ. Microbiol*. 17: 480-95.

Altschul, S. F., Madden, T. L., Schäffer, A. a., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res*. 25: 3389-3402. doi:10.1093/nar/25.17.3389.

Angly, F. E., Felts, B., Breitbart, M., Salamon, P., Edwards, R. A., Carlson, C., Chan, A. M., Haynes, M., Kelley, S., Liu, H., et al. 2006. The marine viromes of four oceanic regions. *PLoS Biol*. 4: e368. doi:10.1371/journal.pbio.0040368.

Bragina, A., Berg, C., Cardinale, M., Shcherbakov, A., Chebotar, V., and Berg, G. 2012. Sphagnum mosses harbour highly specific bacterial diversity during their whole lifecycle. *ISME J*. 6: 802-813. doi:10.1038/ismej.2011.151.

Bragina, A., Oberauner-Wappis, L., Zachow, C., Halwachs, B., Thallinger, G. G., Müller, H. and Berg, G. 2014. The Sphagnum Microbiome Supports Bog Ecosystem Functioning under Extreme Conditions. *Molecular Ecology* 23 (18): 4498-4510. doi:10.1111/mec.12885.

Breitbart, M., and Rohwer, F. 2005. Here a Virus, There a Virus, Everywhere the Same Virus? *Trends in Microbiology* 13 (6): 278–84. doi:10.1016/j.tim.2005.04.003.

Brum, J. R., Hurwitz, B. L., Schofield, O., Ducklow, H. W. and Sullivan, M. B. 2015. Seasonal Time Bombs: Dominant Temperate Viruses Affect Southern Ocean Microbial Dynamics. *The ISME Journal*, August. Nature Publishing Group. doi:10.1038/ismej.2015.125.

Brum, J. R., Ignacio-Espinoza, J. C., Roux, S., Doulcier, G., Acinas, S. G., Alberti, A., Chaffron, S. et al. 2015. Ocean Plankton. Patterns and Ecological Drivers of Ocean Viral Communities. *Science (New York, N.Y.)* 348 (6237): 1261498. doi:10.1126/science.1261498.

Chow, C-E. T., and Fuhrman J. A. 2012a. Seasonality and Monthly Dynamics of Marine Myovirus Communities. *Environmental Microbiology* 14 (8): 2171–83. doi:10.1111/j.1462-2920.2012.02744.x.

Colombet, J., Robin, A., Lavie, L., Bettarel, Y., Cauchie, H. M. and Sime-Ngando, T.. 2007. Virioplankton 'Pegylation': Use of PEG (polyethylene Glycol) to Concentrate and Purify Viruses in Pelagic Ecosystems. *Journal of Microbiological Methods* 71 (3): 212–19. doi:10.1016/j.mimet.2007.08.012.

Comeau, A. M., Hatfull, G. F., Krisch, H. M., Lindell, D., Mann, N. H., and Prangishvili, D. 2008. Exploring the prokaryotic virosphere. *Res. Microbiol*. 159, 306-313. doi:10.1016/j.resmic.2008.05.001.

Fancello, L., Trape, S., Robert, C., Boyer, M., Popgeorgiev, N., Raoult, D. and Desnues, C. 2012. Viruses in the Desert: A Metagenomic Survey of Viral Communities in Four Perennial Ponds of the Mauritanian Sahara. *The ISME Journal*, April. doi:10.1038/ismej.2012.101.

Farnell-Jackson, E. A., and Ward, A. K. 2003. Seasonal patterns of viruses, bacteria and dissolved organic carbon in a riverine wetland. *Freshw. Biol*. 48: 841-851. doi:10.1046/j.1365-2427.2003.01052.x.

Feiner, R., Argov, T., Rabinovich, L., Sigal, N., Borovok, I. and Herskovits, A. A. 2015. A New Perspective on Lysogeny: Prophages as Active Regulatory Switches of Bacteria. *Nature Reviews Microbiology* 13 (10): 641–50. doi:10.1038/nrmicro3527.

Fierer, N., Breitbart, M., Nulton, J., Salamon, P., Lozupone, C., Jones, R., Robeson, M. et al. 2007. Metagenomic and Small-Subunit rRNA Analyses Reveal the Genetic Diversity of Bacteria, Archaea, Fungi, and Viruses in Soil. *Applied and Environmental Microbiology* 73 (21): 7059–66. doi:10.1128/AEM.00358-07.

Fuhrman, J A. 1999. Marine Viruses and Their Biogeochemical and Ecological Effects. *Nature* 399 (6736): 541–48. doi:10.1038/21119.

Hewson, I., J. G. Barbosa, J. M. Brown, R. P. Donelan, J. B. Eaglesham, E. M. Eggleston, and B. A. LaBarre. 2012.

Temporal Dynamics and Decay of Putatively Allochthonous and Autochthonous Viral Genotypes in Contrasting Freshwater Lakes. *Applied and Environmental Microbiology* 78 (18): 6583–91. doi:10.1128/AEM.01705-12.

Huang, Y., Gilna, P., and Li, W. 2009. Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* 25: 1338-1340. doi:10.1093/bioinformatics/btp161.

Hurwitz, B. L., and Sullivan, M. B. 2013. The Pacific Ocean Virome (POV): A Marine Viral Metagenomic Dataset and Associated Protein Clusters for Quantitative Viral Ecology. *PloS One* 8 (2): e57355. doi:10.1371/journal.pone.0057355.

Husson F., Josse J., Le S.and Mazet J. 2015. FactoMineR: Multivariate Exploratory Data Analysis and Data Mining. *R package version 1.29*. http://CRAN.R-project.org/package=FactoMineR

Huson, D. H., Auch, A. F., Qi, J. and Schuster, S. C. 2007. MEGAN Analysis of Metagenomic Data. *Genome Research* 17 (3): 377–86. doi:10.1101/gr.5969107.

Jassey V. 2011. *Impact d'un réchauffement climatique sur le fonctionnement de la sphagnosphère : relations polyphenols-communautés microbiennes.* Thèse de l'Université de Franche-Comté, Besançon (France): 238 p.

King A.M.Q., Adams M.J., and Lefkowitz E.J. 2011. Virus Taxonomy: Classification and Nomenclature of Viruses : Ninth Report of the International Committee on Taxonomy of Viruses. *Elsevier*. https://books.google.com/books?id=KXRCYay3pH4C&pgis=1.

Krupovic, M., Prangishvili, D., Hendrix, R. W. and Bamford, D. H.. 2011. Genomics of Bacterial and Archaeal Viruses: Dynamics within the Prokaryotic Virosphere. *Microbiology and Molecular Biology Reviews : MMBR* 75 (4): 610–35. doi:10.1128/MMBR.00011-11.

Lauro, F. M, DeMaere, M. Z., Yau, S., Brown, M. V., Ng, C., Wilkins, D., Raftery, M. J. et al. 2011. An Integrative Study of a Meromictic Lake Ecosystem in Antarctica. *The ISME Journal* 5 (5). International Society for Microbial Ecology: 879–95. http://dx.doi.org/10.1038/ismej.2010.185.

Le Bras, Y., Roult, A., Monjeaud, C., Bahin, M., and Quénez, O. 2013. Towards a Life Sciences Virtual Research Environment. *e-biogenouest.org*. Available at: https://www.ebiogenouest.org/resources/129/download/jobim_YLeBras_2013.pdf.

Lin, X., Tfaily, M. M., Steinweg, J. M., Chanton, P., Esson, K., Yang, Z. K., Chanton, J. P., Cooper, W., Schadt, C. W. and Kostka, J. E. 2014. Microbial Community Stratification Linked to Utilization of Carbohydrates and Phosphorus Limitation in a Boreal Peatland at Marcell Experimental Forest, Minnesota, USA. *Applied and Environmental Microbiology* 80 (11): 3518–30. doi:10.1128/AEM.00205-14.

López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A., and Alcamí, A. 2009. High diversity of the viral community from an Antarctic lake. *Science* 326: 858-861. doi:10.1126/science.1179287.

Lowe, T. M., and Eddy, S. R. 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res*. 25: 955-964. doi:10.1093/nar/25.5.955.

Maillet, N., Lemaitre, C., Chikhi, R., Lavenier, D., and Peterlongo, P. 2012. Compareads: comparing huge metagenomic experiments. *BMC Bioinformatics* 13: S10. doi:10.1186/1471-2105-13-S19-S10.

Mojica, K. D. A., and Brussaard, C. P. D. 2014. Factors Affecting Virus Dynamics and Microbial Host-Virus Interactions in Marine Environments. *FEMS Microbiology Ecology* 89 (3): 495–515. doi:10.1111/1574-6941.12343.

Opelt, K., Berg, C., Schönmann, S., Eberl, L., and Berg, G. 2007. High specificity but contrasting biodiversity of Sphagnum-associated bacterial and plant communities in bog ecosystems independent of the geographical region. *ISME J*. 1: 502-516. doi:10.1038/ismej.2007.58.

Pagarete, A., Chow, C., Johannessen, T., Fuhrman, J. A., Thingstad, T. F., and Sandaa, R. A. 2013. Strong seasonality and interannual recurrence in marine myovirus communities. Appl. Environ. *Microbiol*. 79: 6253-6259. doi:10.1128/AEM.01075-13.

Parry, L. E., Holden, J. and Chapman, P. J. 2014. Restoration of Blanket Peatlands. *Journal of Environmental Management* 133 (January): 193–205. doi:10.1016/j.jenvman.2013.11.033.

Paul, J. H. 2008. Prophages in Marine Bacteria: Dangerous Molecular Time Bombs or the Key to Survival in the Seas? *The ISME Journal* 2 (6): 579–89. doi:10.1038/ismej.2008.35.

Quaiser, A., Dufresne, A., Ballaud, F., Roux, S., Zivanovic, Y., Colombet, J., Sime-Ngando, T., and Francez, A.-J. 2015. Diversity and comparative genomics of Microviridae in *Sphagnum*- dominated peatlands. *Front. Microbiol*. 6: 375. doi:10.3389/fmicb.2015.00375.

Rodriguez-Brito, B., Li, L., Wegley, L., Furlan, M., Angly, F., Breitbart, M., Buchanan, J. et al. 2010. Viral and Microbial Community Dynamics in Four Aquatic Environments. *The ISME Journal* 4 (6): 739–51. doi:10.1038/ismej.2010.1.

Roux, S., Faubladier, M., Mahul, A., Paulhe, N., Bernard, A., Debroas, D., and Enault, F. 2011. Metavir: a web server dedicated to virome analysis. *Bioinformatics* 27: 3074-3075. doi:10.1093/bioinformatics/btr519.

Roux S., Enault F. 2012. Assessing the Diversity and Specificity of Two Freshwater Viral Communities through Metagenomics. *PloS One* 7 (3): e33641. doi:10.1371/journal.pone.0033641.

Roux, S., Tournayre, J., Mahul, A., Debroas, D., and Enault, F. 2014. Metavir 2: new tools for viral metagenome comparison and assembled virome analysis. *BMC Bioinformatics* 15: 76. doi:10.1186/1471-2105-15-76.

Roux, S., F. Enault, V. Ravet, J. Colombet, Y. Bettarel, J.C. Auguet, T. Bouvier, et al. 2015. Analysis of Metagenomic Data Reveals Common Features of Halophilic Viral Communities across Continents. *Environmental Microbiology*, October. doi:10.1111/1462-2920.13084.

Serreze, M. C., J. E. Walsh, F. S. Chapin III, T. Osterkamp, M. Dyurgerov, V. Romanovsky, W. C. Oechel, J. Morison, T. Zhang, and R. G. Barry. 2000. Observational Evidence of Recent Change in the Northern High-Latitude Environment. *Climatic Change* 46 (1-2): 159–207. doi:10.1023/A:1005504031923.

Srinivasiah, S., Bhavsar, J., Thapar, K., Liles, M., Schoenfeld, T. and Wommack, K. E. 2008. Phages across the Biosphere: Contrasts of Viruses in Soil and Aquatic Environments. *Research in Microbiology* 159 (5): 349–57. doi:10.1016/j.resmic.2008.04.010.

Suttle, C. A. 2005. Viruses in the Sea. *Nature* 437 (7057): 356–61.

Suttle C. A. 2007. Marine Viruses--Major Players in the Global Ecosystem. *Nature Reviews Microbiology* 5 (10): 801–12. doi:10.1038/nrmicro1750.

Tucker, K. P., Parsons, R., Symonds, E. M. and Breitbart, M. 2011. Diversity and Distribution of Single-Stranded DNA Phages in the North Atlantic Ocean. *The ISME Journal* 5 (5): 822–30. doi:10.1038/ismej.2010.188.

Tveit, A., Schwacke, R., Svenning, M. M., and Urich, T. 2013. Organic carbon transformations in high-Arctic peat soils: key functions and microorganisms. *ISME J*. 7: 299-311. doi:10.1038/ismej.2012.99.

Williamson, K. E., Wommack, K. E. and Radosevich, M. 2003. Sampling Natural Viral Communities from Soil for Culture-Independent Analyses. *Applied and Environmental Microbiology* 69 (11): 6628–33. doi:10.1128/AEM.69.11.6628-6633.2003.

Williamson, K. E., Radosevich, M., and Wommack, K. E. 2005. Abundance and diversity of viruses in six Delaware soils. *Appl. Environ. Microbiol*. 71: 3119-3125. doi:10.1128/AEM.71.6.3119-3125.2005

Winter, C., Matthews, B. and Suttle, C. A. 2013. Effects of Environmental Variation and Spatial Distance on Bacteria, Archaea and Viruses in Sub-Polar and Arctic Waters. *The ISME Journal* 7 (8): 1507–18. doi:10.1038/ismej.2013.56.

Wommack, K. E., Nasko, D. J., Chopyk, J. and Sakowski, E. G. 2015. Counts and Sequences, Observations That Continue to Change Our Understanding of Viruses in Nature. *Journal of Microbiology (Seoul, Korea)* 53 (3): 181–92.

Zablocki, O., Zyl, L. V., Adriaenssens, E. M., Rubagotti, E., Tuffin, M., Cary, S. C. and Cowan, D. 2014. High-Level Diversity of Tailed Phages, Eukaryote-Associated Viruses, and Virophage-like Elements in the Metaviromes of Antarctic Soils. *Applied and Environmental Microbiology* 80 (22): 6888–97. doi:10.1128/AEM.01525-14.

Chapitre 4:

Effects of human-derived disturbances on viral abundance and diversity in *Sphagnum*-dominated peatlands

Chapitre 4: Do viral abundance and diversity reflect the effects of human-derived disturbances on peatland functioning?

Résumé du Chapitre 4

La fonction de puits de carbone des tourbières est associée à un service écosystémique de régulation du climat. Cette fonction est menacée par les perturbations d'origine anthropique. Nous avons donc cherché à déterminer, pour deux de ces perturbations, si elles affectaient le fonctionnement du compartiment viral.

Nous avons étudié l'impact du réchauffement climatique sur l'abondance des particules virales dans deux stations expérimentales situées en France et en Sibérie. Ces stations sont équipées de dispositifs d'Open-Top-Chambers (OTC) pour simuler une augmentation de la température annuelle moyenne de 1 à 2°C. En Sibérie, à Muhkrino, ce dispositif est couplé à un dispositif de manipulation de la colonne de tourbe qui affecte le niveau de la nappe d'eau. Nous n'avons pas détecté de changement significatif de l'abondance des particules virales et des procaryotes dans les échantillons issus des placettes équipées d'OTC. En revanche, la combinaison du traitement OTC avec la manipulation de la nappe semble associée à une diminution de l'abondance virale, et à une augmentation du carbone organique dissous (COD), suggérant une inversion de la relation entre le COD, l'activité et l'abondance des procaryotes, et l'activité virale.

D'autre part, nous avons étudié l'effet de mesures de restauration écologique menées depuis 15 ans et 1 an sur des tourbières affectées par les changements d'occupation des terres (exploitation de tourbe, drainage). Le but de ces mesures de restauration est de retrouver la fonction de puits de carbone des tourbières. Leur succès est souvent attesté par le recouvrement de la végétation et l'équilibre des flux de carbone. En revanche, le recouvrement de la diversité et de l'activité microbienne est rarement vérifié. A la Guette (France), site restauré depuis un an, l'abondance des virus et des procaryotes est plus faible en profondeur qu'en surface, comme cela avait été observé au début du programme de restauration du site suivi à quinze ans. Ces plus faibles abondances en profondeur, comparées à la zone naturelle, suggérent que dans la zone en cours de restauration l'activité microbienne reste perturbée.

A Bois-des-Bel au Québec, où le programme de restauration écologique a été lancé il y a 15 ans, la comparaison de l'abondance des particules virales et des procaryotes entre la zone restaurée et la zone naturelle indique un potentiel recouvrement d'une plus forte abondance virale et procaryote en profondeur du site restauré, suggérant un fonctionnement proche des tourbières non perturbées. En revanche, la diversité virale de la matrice de tourbe et de l'eau de Sphaigne reste distincte entre les

deux zones (naturelle et restaurée), indiquant que la diversité et l'activité microbienne sous-jacentes des deux zones sont différentes. Cette différence est notamment liée à une distinction entre la surface et la profondeur de la colonne de tourbe de la zone restaurée, qui suggère comme à Mukhrino qu'une perte de la continuité de la colonne de tourbe joue un rôle impacte l'activité et la diversité des procaryotes.

Do viral abundance and diversity reflect the effects of human-derived disturbances on peatland functioning?

Introduction

Many studies that focus on peatland functioning are recurrently justified by the importance of these ecosystems as a carbon-sink, accumulating organic matter since the beginning of the Holocene and currently storing around 25-30 % of the terrestrial C-stock (Mitra et al., 2005; Frolking et al., 2011). However, this functional process, recognized as a climate regulating ecosystem service (MEA, 2005) is threatened because of the increasing pressure of human activities such as land use change or mining, on both tropical and boreal peatlands (Hooijer et al., 2010 ; .Junk et al., 2012).

To-date most societal issues about the loss of ecosystem services surround our ability to model and predict ecosystems functioning in response to global change. The large panel of environmental changes expected in response to anthropic activity spans from increased mean annual temperatures, longer summer droughts or thaws, to disrupted biogeochemical cycles and trophic networks. Monitoring of climate indicators already suggested that these changes occur faster and stronger at higher latitudes where most of *Sphagnum*-dominated peatlands are situated (Serreze et al., 2000 ; IPCC, 2013). Whether peatlands can maintain their carbon storage capacity under such changes is a major issue (Yergeau et al., 2010). Increased temperatures, longer snow periods and longer summers affect the seasonality of northern peat soils, and lead to strong changes in C-fluxes, plant-microbial interactions, microbial activity, subsequent decomposition processes and net ecosystem exchange (Dorrepaal et al., 2009 ; Mackelprang et al., 2011; Jassey et al., 2013). For instance, longer and more intense permafrost soil thaws are expected to increase methane effluxes by activating initially unactive micro-organisms in deeper peat layers (Tveit et al., 2013). Direct warming effects on peatland functioning also induce changes inside the ecosystem that modulate such effects. Ward et al. (2013) showed that change of vegetation with warming concurrenced the direct effect and then exerted a significant feedback on the GHG (Greenhouse-Gaz) fluxes. Under natural peatland functioning, methane produced at depth can be oxidized at the surface by methanotrophic bacteria and the released carbon dioxide sustains *Sphagnum* growth and production (Vile et al., 2014). In addition to the disruption of the production/decomposition imbalance and the microbial communities implicated in the carbon cycle, longer thawing periods in Siberian peatlands have allowed the discovery of new viruses (Abergel and Claverie, 2014; Legendre et al., 2015)[1], the

[1] While viral ecologists were enthusiastic about these new giant viruses and their incredibly large genomes, virologists were more concerned about sanitary consequences of the possible re-emergene of presently uninfecting viruses. But that issue is not addressed here.

response of these ecosystems to upcoming changes is critical in regard of the evolution of Earth climate through the potential modification of their carbon cycle.

These concerns are not limited to the future functioning of pristine peatlands as these ecosystems have also long been exploited to sustain the production of human goods. Peat has traditionally been extracted for fuel or reclaimed for agriculture and forestry purposes (Francez, 2000; Mitsch and Gosselink, 2007). It is now acknowledged that these activities have strongly modified the structural, hydrological and biological peatland functioning and properties, turning exploited sites from carbon sinks to carbon sources (Francez and Vasander 1995; Strack, 2008).

Since 30 years, as awareness rose about the trade-offs between ecosystem exploitation, ecosystem services and human-derived global changes, restoration programs have been set-up in order to restore peatland structure and functioning. Restoration of ecosystems became a major strategy to enhance both the recovery of biodiversity losses and the provision of ecosystem services (Bullock et al., 2011), even if the concept of ecosystem services is not explicitly mentioned in peatland restoration manipulations (Kimmel and Mander, 2010). Once these restoration programs set up, peatland dynamic remains to be monitored in order to evaluate the success of the restoration practices. The success of peatland restoration programs is mainly assessed using vegetation records (Poulin et al., 2013), followed later by the recovery of carbon fluxes (Waddington et al., 2001) and microbial activity (Andersen et al., 2006 and 2010; Artz et al., 2008) but how the microbial functional diversity is recovered remains yet poorly addressed. In addition, how the viral loop can interact in the recovery of microbial functions is completely unknown. Given viruses specificity for their hosts, and their dependence on their host activity, their diversity represent a tool as it could be an indicator that mirror changes in the microbial compartment diversity and activity that has been shown to be crucial for peatlands functioning.

. The aim of the present chapter was to test whether viral ecology could give clues about the way peatland microbial compartment reacts in response to (i) experimental climatic changes (T°C and hydrology) induced with Open Top Chambers devices (OTC) and water-table level manipulations, and (ii) cutover peatland restoration. The questions were addressed using viral-particles and prokaryotic counts by cytometry in relation to the physico-chemistry (environmental changes). This approach was combined to the analysis of the viral diversity in the natural and restored area of a formerly exploited peatland in Québec.

Material and methods

Geographical repartition of the experimental stations

Effect of global changes on the viral and prokaryotic compartments was studied in two experimental stations situated in France and Russia. The effect of peat exploitation and peatland restoration was studied in two other peatlands, in France and Canada (Figure 1).

Figure 1 Erreur ! Pas de séquence spécifié. **Geographical repartition of the peatlands for the monitoring of the effect of global changes (Stars) and of peatlands exploitation and restoration (Black circles). A: Frasne (France); B: Mukhrino (Russia); C: La Guette (France); D: Bois-des-Bel (Canada).**

Effect of climatic changes: increased temperature (OTC) and water-table manipulation at the experimental stations of Mukhrino (Russia) and Frasne (France).

In France (Frasne) and Russia (Mukhrino) two experimental stations have been created in a fen and a bog profile to mimic and monitor the effects of climatic changes on the functioning of *Sphagnum*-dominated peatlands. Different plots from these two stations have been equipped with open-top chamber devices (OTC) that were designed according the ITEX (International Tundra Experiment) protocol in order to increase the air temperature by 2 to 3 degrees maximum (annual average) over the scale of a decade. Both Frasne and Mukhrino experimental stations are composed of 1m² plots, with or without OTC (controls) and each plot is equipped with a central piezometer to allow peat water sampling.

Frasne experimental station is composed of 12 plots: 6 plots in a fen profile (*Sphagnum fallax, Carex limosa* and *C. rostrata* dominant*)* (3 controls and 3 OTC), and 6 plots (same repartition) in a bog

profile (*Sphagnum magellanicum and S. capillifolium, Eriophorum vaginatum,* and *Andromeda polifolia* dominant). The device installed in 2008 lead to an average increase of 1,2 °C of the air temperature in the OTCs after 16 months of incubation (Delarue et al., 2010; Jassey et al., 2013).

Mukhrino bog profile (*Sphagnum fuscum, Eriophorum vaginatum, Andromeda polifolia* dominant) also consists in 6 plots, 3 controls and 3 OTC, and the fen profile (mainly *Sphagnum balticum, Scheuchzeria palustris, Carex limosa, Vaccinium oxycoccos* dominant) of the experimental station is combined to an experimental modification of the water-table level by raising or lowering peat columns (0.01 m^3) . For this water-table modification, 10cm of peat under the active *Sphagnum* layer have been cut-off and harvested (1m² of 10cm thick peat) to create the -10cm treatment plots that mimics an increased water-table level. The harvested peat was added under the active *Sphagnum* layer of the +10cm treatment to decrease the water-table level of these plots. For the "0" treatment, peat was cut-off in the same way, but left on the plot to monitor the effect of the sole peat matrix disruption. Control plots were left uncut. The 4 water-table treatments (Control (water-table control), 0, -10cm and +10cm) were completed with the OTC devices (C (OTC control) and OTC), and each cross-treatment was applied in triplicate. The fen station of Mukhrino was thus composed of 24 plots (4 (water table treatments) x2 (warming treatments) x3 (replicates)).

Restoration of exploited peatlands: Bois-des-Bel (Québec) and La Guette (France).

La Guette peatland is situated in France and has been subjected to afforestation by *Pinus silvestris* and *Betula* spp as a result from drainage and changes in land-use. In parallel, subsequent modification of the peat ecosystem (increased nutrients inputs) has lead to its colonization by *Molinia caerula*. Drainage and disturbances have approximately divided the site into an upstream area, that remained unaffected, and a downstream area in which peat became degraded. A restoration program of the peatland, mainly focusing on the restoration of the water regime started February in 2014.

The Canadian peatland of Bois-des-Bel is an experimental station used for the evaluation of a large peat restoration program (Rochefort et al., 2003). This station is composed of a natural area, a formerly exploited unrestored area, and a restored area. Extraction of peat in Bois-des-Bel took place from 1973 to 1980, and the site was then abandoned for 20 years before a restoration program was initialized. At the program set up, vegetation and especially *Sphagnum* had not re-colonized the cutover area which still represented a carbon source. To evaluate the effect of the restoration on the vegetation, carbon budget and microbial activity, only 8 of the 11 exploited hectares have been restored in 2000: soil was leveled, *Sphagnum* were sowed, drainage channels (ditches) were blocked, and the restored area was fertilized with phosphorus (Rochefort et al., 2003).

Sampling strategies

Experimental modification of the temperature and water-table

In the two experimental stations, water from the piezometers was collected to measure VPA, PA, VPR. Water from the 30 piezometers of Mukhrino station was sampled on 07 July 2014. In Frasne, the 12 piezometers were sampled on 15 October 2014. Water samples were kept at 4°C until return to the laboratory, filtered at 5µm (Minisart, Sartorius), and 2mL were aliquoted for counting. Glutaraldehyde was added to the aliquots to a final concentration of 2% (Sigma, TEM grade). Samples were flash frozen in liquid nitrogen and kept at -20°C before counting.

Effect of peat exploitation and restoration

Briefly, for each sampling point, the active capitula layer was removed, 'surface' peat was harvested for the 'peat' surface sample and put in sterile tubes (Falcon 50mL), pore-water was expressed from the peat for the 'pore-water samples' and for the physico-chemistry, prefiltered at 125µm and stored at 4°C in sterile bottles. The same sampling method was applied to the 'depth' layer.

In la Guette, three peat columns (N=3) were sampled in each of the undisturbed "upstream" and recently restored "downstream" areas. Pore-water samples from the depth could not be retrieved in the restored area as a consequence of too low water content. In Bois-des-Bel, peat columns from the natural and the restored area have been collected in replicate (N=2). The peat columns were divided between surface (5 cm to 15cm under the active capitula layer) and depth (15 to 25cm).

Physico-chemical features

Effect of climatic changes

In Mukhrino and Frasne, temperature, pH, O_2 and conductivity were measured directly in the piezometers during the sampling sessions. For the samples from Mukhrino, Dissolved Organic Carbon (DOC) and Soluble Organic Nitrogen (SON) were quantified at the ISTO (Orléans) in water sampled on the same day and conserved at 4°C. For the samples from Frasne, DOC and anions (nitrate, sulfate and acetate) were analyzed in the prefiltered (125µm) water from the piezometers at the OSUR and Ecobio analytic platforms, using a Bioritech DOC Analyser and a Dionex Analyzer respectively.

Effect of restoration

Physico-chemistry for the samples from La Guette was obtained in the same way as described in chapter 3.

Viral particles and prokaryotes abundances

Samples processing and counting

Peat and pore-water samples were processed immediately after return to the lab. Peat matrix and pore-water samples were processed separately and viral-particles and prokaryote abundances per mL of pore-water and per gram of dry peat were obtained by cytometry as described in Chapter 3.

Metaviromes production

Metaviromes from the peat matrix and the pore-water were produced for the two depths from the natural and restored area of Bois-des-Bel. Metaviromes production was done as described in chapter 3. In total, 14 metaviromes were produced from Bois-des-Bel.

Table 1 Metaviromes produced for the comparison of the viral diversity from the peat matrix and the pore-water samples in the natural and restored area of Bois-des-Bel (Canada)

	"Peat" metaviromes	"Water" metaviromes
Natural	vCan_Nat1_Sp vCan_Nat1_Dp vCan_Nat2_Sp vCan_Nat2_Dp	vCan_Nat1_Dw vCan_Nat2_Sw vCan_Nat2_Dw
Restored	vCan_Rest1_Dp vCan_Rest1_Sp vCan_Rest2_Dp vCan_Rest2_Sp	vCan_Rest1_Sw in *progress* vCan_Rest2_Dw vCan_Rest2_Sw

In the metaviromes names: p (last letter) corresponds to peat matrix viral communities, w (last letter) corresponds to pore-water viral communities, number (1 or 2) corresponds to the peat column (or sampling point), S corresponds to surface metaviromes, and D corresponds to depth metaviromes.

Metaviromes analysis

The metavirome sequences were uploaded on Metavir (Roux et al., 2011; Roux et al., 2014) and subjected to tblastx (bit score: 50) (Altschul et al., 1997) against RefSeqVirus. Matches to the different viral types were analyzed using MEGAN (Huson et al., 2007) and used to build a contingency matrix that represents the relative proportions of the sequences assigned to each viral type. A principal component analysis was performed on this contingency matrix representing the different viral types in the 14 metaviromes to determine what factor (area, type of metavirome and depth) affected the referenced viral diversity (R FactorMineR package, Husson et al., 2015). The similarity of the peat matrix and pore-water metaviromes was compared using commet (k=33, t=4) on the galaxy server (Maillet et al., 2012; Le Bras et al., 2013).

Statistical analysis

For the samples from the piezometers, single and combined effect of OTC and water-table modification on VPA, PA and VPR were tested using Kruskal-Wallis tests. A principal component analysis was performed using the R FactoMineR package (Husson et al., 2015) on the physico-chemical features measured in the piezometers water. Biological counts were added as supplementary quantitative information, and experimental station (Frasne, Mukhrino), dynamic

stage (fen, bog), warming treatment and water-table treatment were added as supplementary qualitative information.

For the samples from La Guette and Bois-des-Bel, effect of area (natural *vs* restored), depth (surface *vs* depth) and combined effect of area and depth on VPA, PA and VPR were compared separately for each peatland. A principal component analysis was performed on the physico-chemical features measured in the pore-water samples from La Guette, with biological counts added as supplementary quantitative information, and area, site and depth added as supplementary qualitative information.

Statistical analyses were performed using the open-source statistical software R (version 2.14) (R Development Core Team, 2013)

Results

Effect of OTC and water-table modification on viral-particles and prokaryotes abundances

In order to determine how increased temperature and water-table modification could (i) affect viral-particles and prokaryote abundance, (ii) peat-water physico-chemistry and the potential relationships between the abundance of the biological entities and abiotic features, water samples were collected in the piezometers from two experimental stations situated in Frasne (France) and Mukhrino (Russia).

Due to the high variability of the abundances, we did not detect any significant effect of the warming and water-table manipulation experiments. Some trends are somehow interesting. First VPA and PA seem higher in the fen than in the bog in both Frasne and Mukhrino (Figure 2A and 2C), and the VPR seem to evolve in a different manner in the two dynamic stages with the warming experiment. There is a potentially increasing VPR in the bog with the OTC, and an inverse trend in the fen (Figure 1E). Secondly, any modification of the peat matrix (Cut, -10cm and +10cm) in the fen in Mukhrino seems to be associated to a decreased VPA (Figure 2D). In this water-table modification experiment, any modification of the water-table combined to the OTC treatment seems to lead to a decrease in PA (Figure 2B), VPA (Figure 2D) and VPR (Figure 2F).

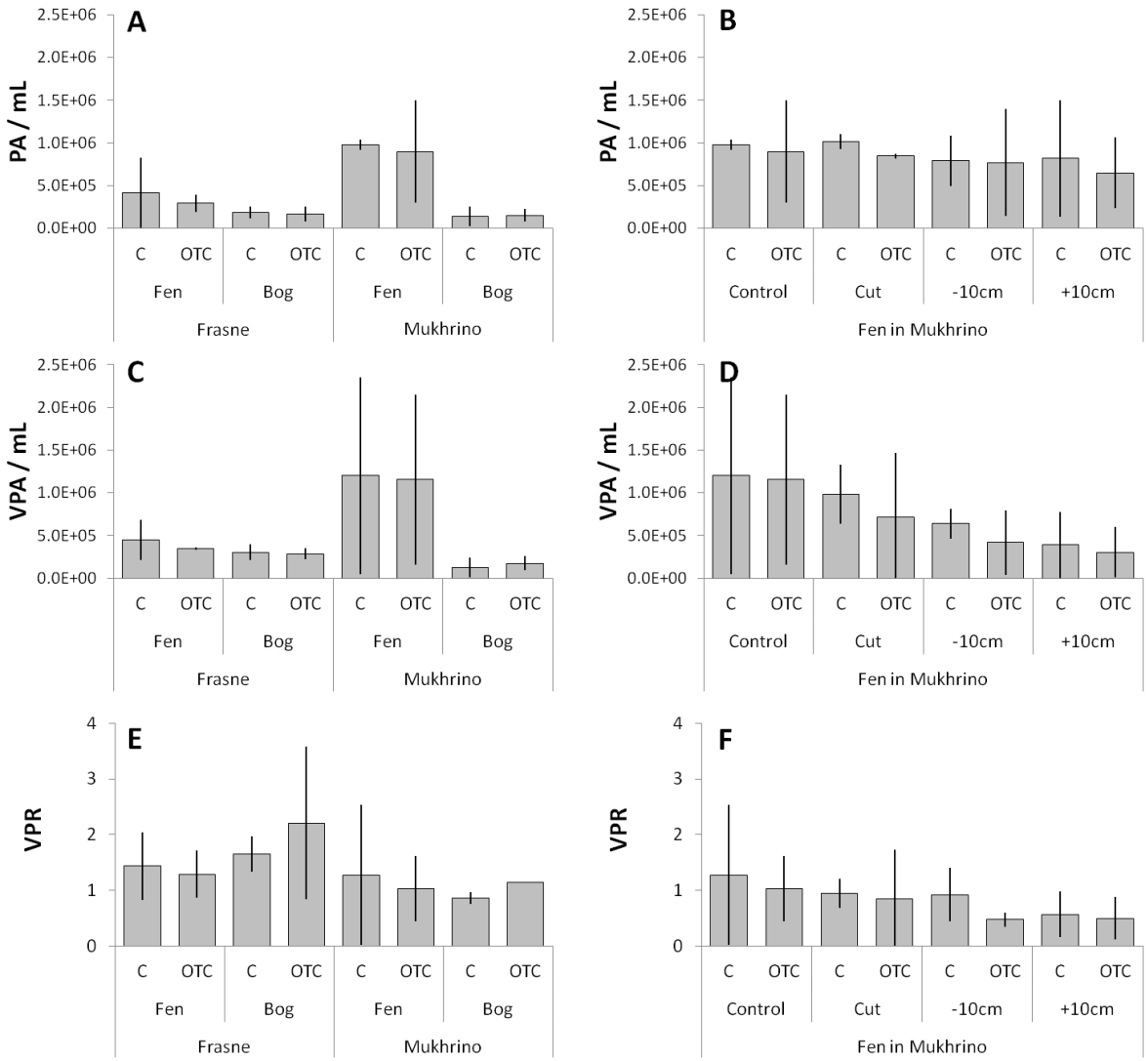


Figure 2 Prokaryotes abundance (A-B), viral-particles abundance (C-D) and VPR (E-F) measured in the water sampled from the piezometers in the experimental devices from Frasne (France) and Mukhrino (Russia), depending on the dynamic stage and experimental warming (A-C-E) or on the cross-effect of the water-table modification and experimental warming in the fen profile from Mukhrino (B-D-F). Abbreviations: VPA: Viral particles abundance; PA: prokaryotes abundance: VPR: Virus to prokaryote ratio; OTC: open top chamber; C: warming control (no OTC); Control (water-table modification control); Cut: cut-peat; -10cm: peat removed (increased water table); +10cm: peat added (decreased water table). N=3, bars represent standard deviation. For A-C-E, values from the fen in Mukhrino are given for the "Control" piezometers.

Effect of OTC and water-table modification on the physico-chemistry

In order to analyze the variation of the physico-chemical features between the treatments, and to observe potential relationships between the variation of the physico-chemistry and the abundance of the biological entities, a principal component analysis (PCA) was performed on the physico-chemical features measured in all the piezometers from Frasne and Mukhrino. Biological variables (VPA, PA and VPR) were added as quantitative information; site (Frasne/Mukhrino), dynamic stage (fen/bog), warming treatment and water-table manipulation were added as qualitative information.

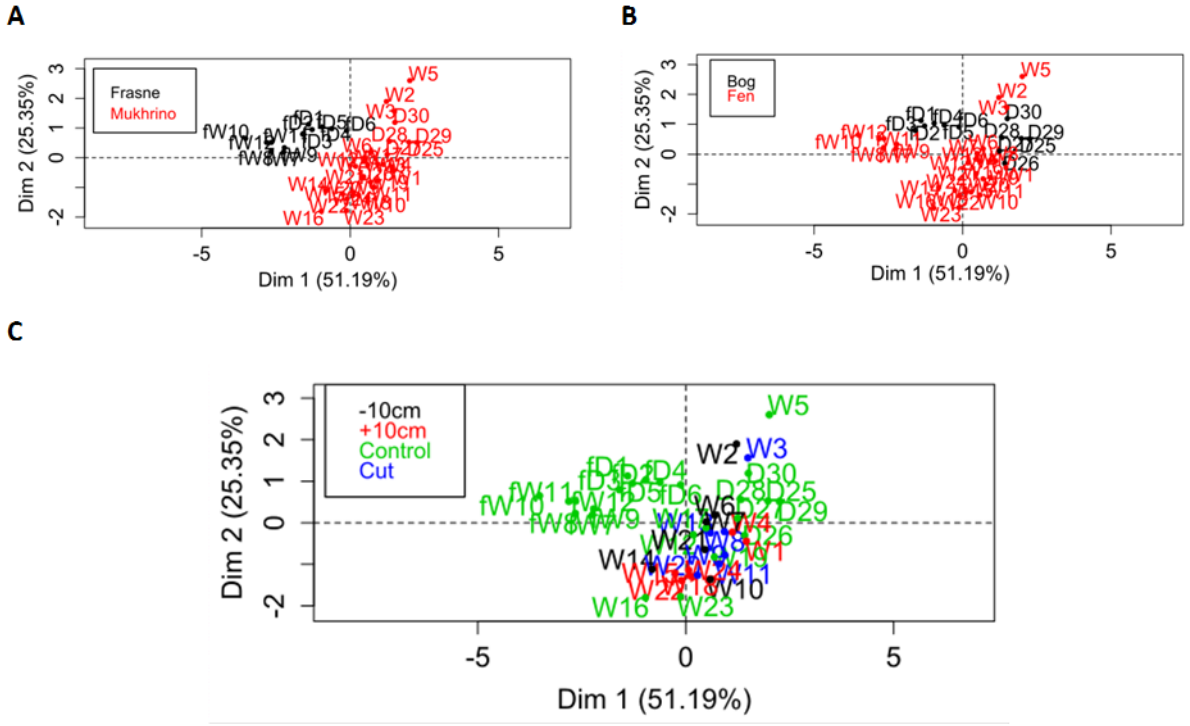


Figure 3 Projections of the samples from the principal component analysis performed on the physico-chemical features depending on experimental site (Frasne or Mukhrino) (A), fen or bog profile (B) or water-table experiment (C).

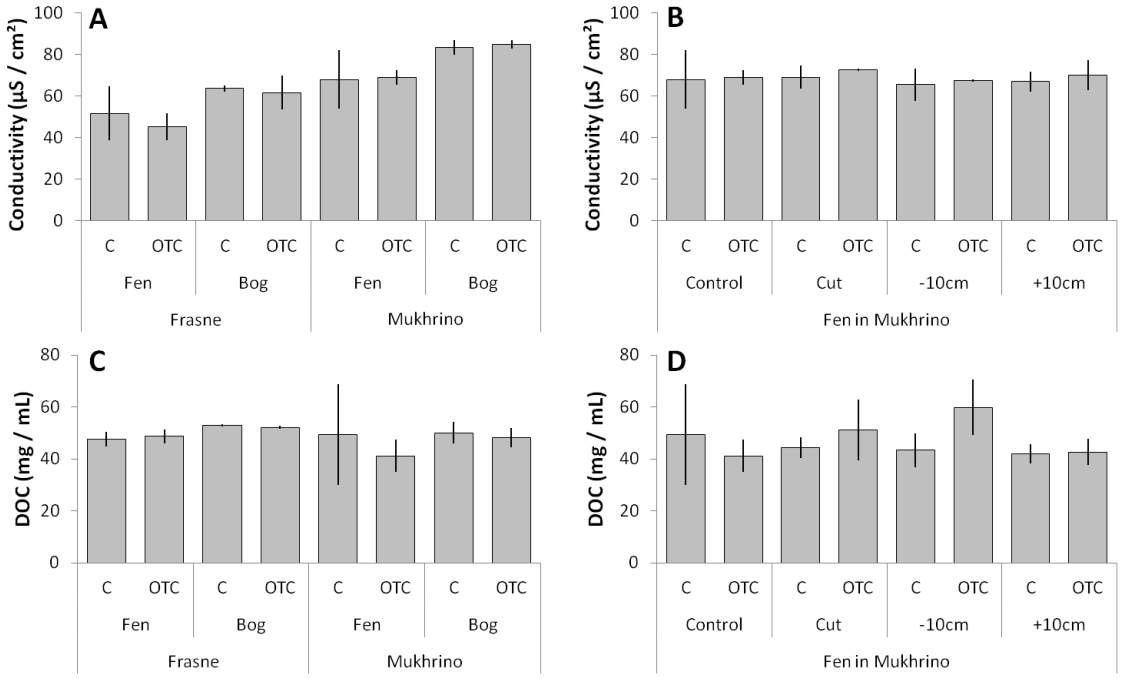


Figure 4 Conductivity (A-B) and dissolved organic carbon (DOC) (C-D) measured in the water sampled from the piezometers in the experimental devices from Frasne (France) and Mukhrino (Russia), depending on the dynamic stage and experimental warming (A-C) or on the cross-effect of the water-table modification and experimental warming in the fen profile from Mukhrino (B-D). Abbreviations: OTC: open top chamber; C: warming control (no OTC); Control (water-table modification control); Cut: cut-peat; -10cm: peat removed (increased water table); +10cm: peat added (decreased water table). N=3, bars represent standard deviation. For A-C, values from the fen in Mukhrino are given for the "Control" piezometers.

The first component of the PCA carried 51.2 % of the information and was mainly driven by conductivity (positive values) and pH (negative values), which were significantly associated to the difference between Mukhrino (positive values) and Frasne (negative values) (Figure 3A, Figure 4A). The VPR was weakly but significantly correlated to the first component and thus higher in Frasne were the conductivity was lower than in Mukhrino. The second component carried 25.3 % of the variance, and was mainly driven by dissolved organic carbon (DOC) (positive values). VPA was negatively correlated to this component and to DOC. Bog was positively correlated to this component (positive values), to the contrary of fen (negative values) (Figure 3B, Figure 4C), suggesting a higher VPA in the fen were the DOC was lower. Finally, to the contrary of "control" water-table, +10cm treatment (decreased water table level) was negatively correlated to this component, and associated to low VPA and high DOC (Figure 3C, Figure 4D).

Effect of peatland restoration

Prokaryotes and viral-particles abundances were obtained from surface and depth peat samples in natural and restored areas from La Guette peatland (France), where restoration procedure started a year ago, and from Bois-des-Bel (Canada) were restoration of the exploited area started 15 years ago.

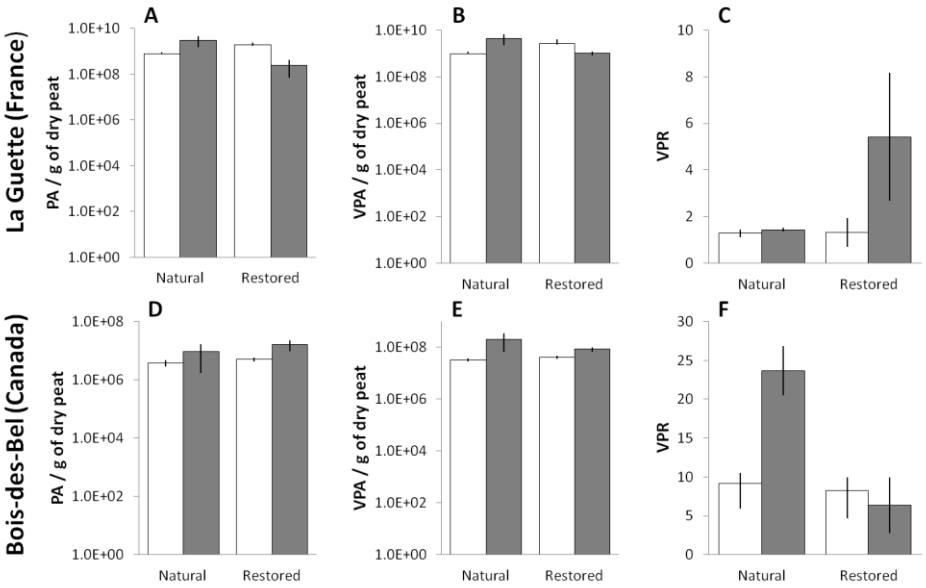


Figure 5 Prokaryotes abundance (A-D), viral particles abundance (B-E) and VPR (C-F) measured at the surface (white) and at depth (grey) of the peat matrix from the natural and restored area in the peatlands of La Guette (France) (A-B-C) and Bois-des-Bel (Canada) (D-E-F). Abundances are shown on a log-scale. For La Guette, N=3, for Bois-des-Bel, N=2. Bars represent standard deviation. Abbreviation : PA : prokaryotes abundance ; VPA : viral particles abundance ; VPR : virus to prokaryote ratio.

In la Guette, PA and VPA presented a different surface/depth pattern between the natural and the disturbed area (Kruskal-Wallis test, p= 0.02 and p=0.03, n= 24), with higher PA and VPA at depth in

the natural area, and the inverse pattern in the disturbed area (Figure 3A and B). In Canada, PA and VPA varied with depth (MW, p=0.01 and p=0.003, n=16), and with the combined effect of site and depth (KW, p=0.002 and p=0.0002, n=16). PA and VPA were higher at depth in both area (Figure 5D and E). In la Guette, VPR was influenced by depth (MW, p=0.04, n=24), while in Bois-des-Bel, VPR was influenced by the area (MW, p=0.02, n=16) and the combined effect of depth and area (KW, p=0.02, n=16).

In order to obtain a more detailed overview of the variation of the physico-chemical parameters between the natural and the restored area, pore-water samples were retrieved with the peat samples in La Guette, at the surface and depth in the natural area, and only at the surface in the disturbed area (Table 2). In the disturbed area, water samples could not be extracted as the water-content of the peat was too low.

Table 2 Nutrients and ions concentrations measured in the pore-water from La Guette. Mean (± standard deviation).

		DOC (mg.mL)	SON (mg.mL)	K+	Ca++	Cl-	NO3-	SO4—	PO4---
Natural	**Surface**	63 (25)	1.2 (0.9)	12.3 (6.8)	5.9 (4.1)	16.1 (7.9)	0.4 (0.2)	2.1 (0.4)	0.02 (0.01)
	Depth	48 (22)	1.4 (0.1)	7.5 (3.6)	4.7 (1.9)	8.4 (0.9)	0.1 (0.03)	1.3 (0.1)	0.27 (0.37)
Restored	**Surface**	43 (13)	1.4 (0.1)	3.4 (2.5)	2.6 (1.6)	5.9 (1.5)	0.5 (0.2)	0.7 (0.2)	0.01 (0.0)

In order to analyze the variation of the physico-chemical features and to look for potential covariation with VPA and PA, a principal component analysis (PCA) was performed on the data-set from La Guette.

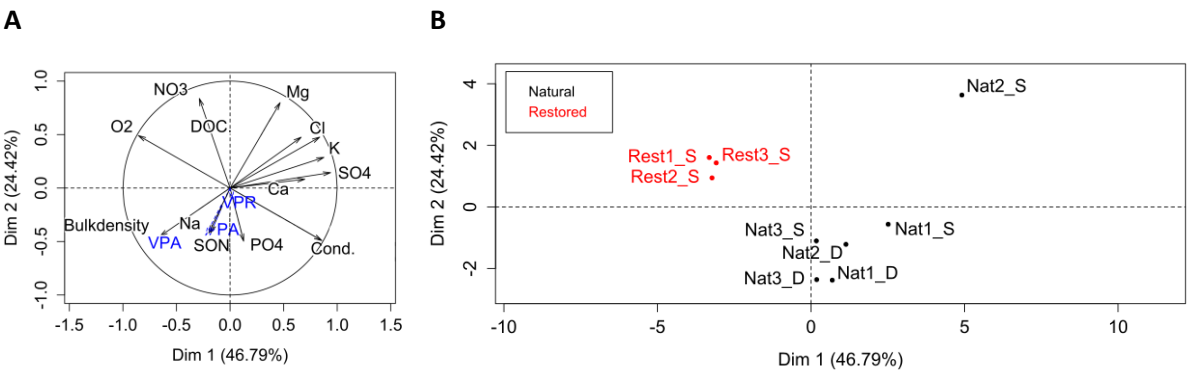


Figure 6 Principal component analysis on the physico-chemical variables measured in the natural and the restored area from La Guette (France): variables map (A) and individual mapping (B).

'Area' and depth were added as supplementary qualitative information, and VPA, PA and VPR as supplementary quantitative variables (Figure 6A). The first component carried 46.8 % of the variance,

and was associated to the difference between the natural and the restored area, with more sulfate, K+, Chloride, and Ca^{2+} in the natural area. The second component (24.42% of the variance) was associated to the difference between surface and depth in the natural area with more nitrate and Mg^{2+} at the surface. This PCA indicates that there was more variation between the natural and the restored area than between the surface and depth of the natural area (Figure 6B).

Viral diversity in the restored and natural areas from Bois-des-Bel

Fourteen pore-water (Table 3) and peat matrix metaviromes (Table 4) have been produced to study the viral diversity from the natural and restored areas in Bois-des-Bel. This represents a total of 10 105 252 sequences, of which 5 299 245 sequences represented the natural peat-matrix and pore-water viral communities (analyzed in Chapter 3), and 4 806 007 sequences represented the peat matrix and pore-water viral diversity from the restored area.

Table 3 Main properties of the pore-water metaviromes produced for the natural and restored area of in Bois-des-Bel (Canada)

Country	**Canada**					
Type	**Pore-water metaviromes**					
Area	**Natural**			**Restored**		
Peat column	**1**	**2**	**2**	**1**	**2**	**2**
Origin	**Depth**	**Surface**	**Depth**	**Surface**	**Surface**	**Depth**
Name	**vCan_Nat1_Dw**	**vCan_Nat2_Sw**	**vCan_Nat2_Dw**	**vCan_Rest1_Sw**	**vCan_Rest2_Sw**	**vCan_Rest2_Dw**
Sequencing type	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina
Metavir ID	6094	6095	6096	6097	6098	6099
N° of sequences	885 598	606 800	552 541	639 213	581 239	645 873
% assigned sequences	4.96	5.94	5.89	4.28	3.8	3.66

Table 4 Main properties of the peat matrix metaviromes produced for the natural and restored area in Bois-des-Bel (Canada).

Country	**Canada**							
Type	**Peat metaviromes**							
Area	**Natural**				**Restored**			
Peat column	**1**		**2**		**1**		**2**	
Depth	**Surface**	**Depth**	**Surface**	**Depth**	**Surface**	**Depth**	**Surface**	**Depth**
Name	**vCan_ Nat1_Sp**	**vCan_ Nat1_Dp**	**vCan_ Nat2_Sp**	**vCan_ Nat2_Dp**	**vCan_ Rest1_Sp**	**vCan_ Rest1_Dp**	**vCan_ Rest2_Sp**	**vCan_ Rest2_Dp**
Sequencing type	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina	Illumina
Metavir ID	6100	6101	6102	6103	6104	6105	6106	6107
N° of sequences	784 862	813 320	825 206	830 918	863 801	839 501	533 149	703 231
% assigned sequences	7.24	5.99	5.11	4.82	4.42	5.54	2.89	5.09

Comparison of the viral diversity based on the taxonomical assignations

In order to determine their taxonomical assignation, sequences from the pore-water and peat metaviromes have been uploaded on Metavir (Roux et al., 2011; Roux et al., 2014) and compared by tblastx against RefSeqVirus (bit score 50). On average, 5.53 ± 0.77 % and 4.24 ± 0.82 % of the sequences from the natural and restored area respectively matched to referenced viral genomes.

Taxonomical assignations of the sequences were retrieved using MEGAN (Huson et al., 2007). The viral communities from the restored area appeared dominated by the dsDNA viruses (no RNA stage), that represented 31.41 ± 5.34 % of the peat matrix and 43.36 ± 2.58 % of the pore-water metaviromes from this area (Figure 7).

In order to facilitate the analysis, the relative abundance of the different viral types (Figure 7) was used to build a contingency matrix, and the variance of this dataset was analyzed with a principal component analysis. The first component carried 42.35% of the variance and was correlated with the difference between the peat-matrix and pore-water metaviromes, with more Circoviridae, Satellites and Boiling-Spring viruses (chimere) in the peat matrix metaviromes, and more Mimiviridae, Caudovirales and other dsDNA viruses (no RNA stage) in the pore-water metaviromes. The second component carried 17.40% of the variance. However, differences associated with the area or depth were not significantly associated to this second component nor to the third one (Figure 8).

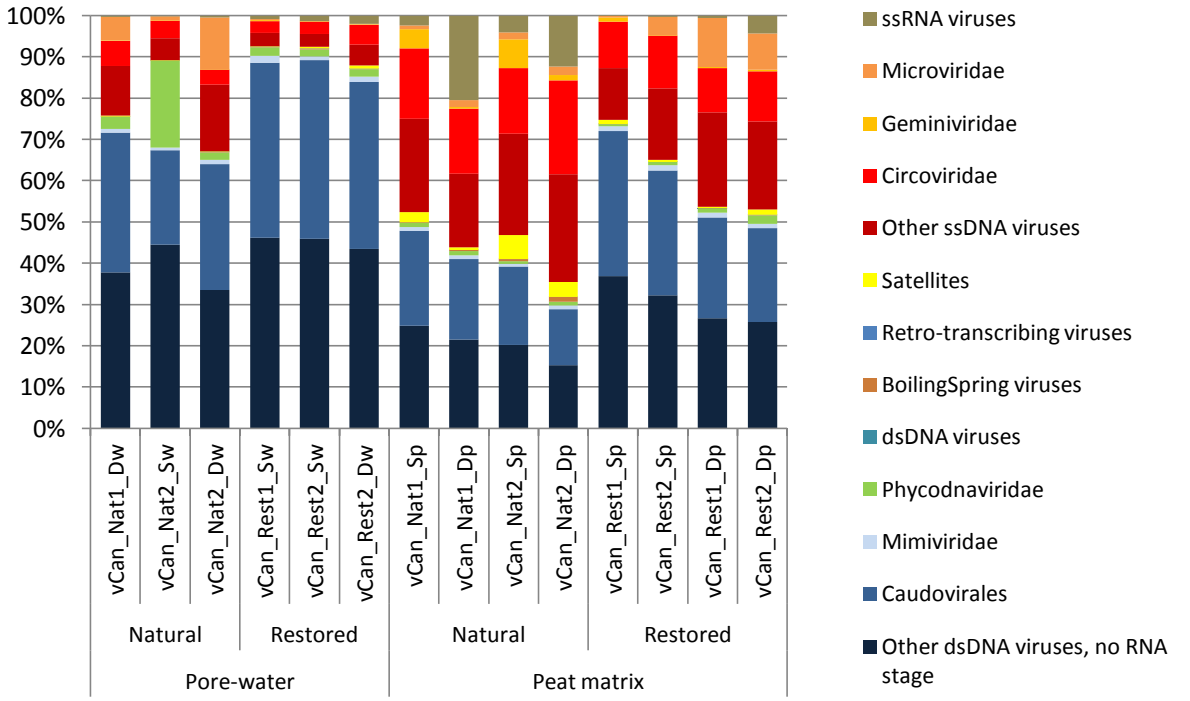


Figure 7 Relative abundances of the referenced viral types in the metaviromes produced from the surface and depth of the peat column, in the pore-water and peat matrix from the natural and restored area in the Bois-des-Bel peatland (Canada). Taxonomical assignations were obtained by tblastx against RefSeqVirus (bit score 50) on Metavir.

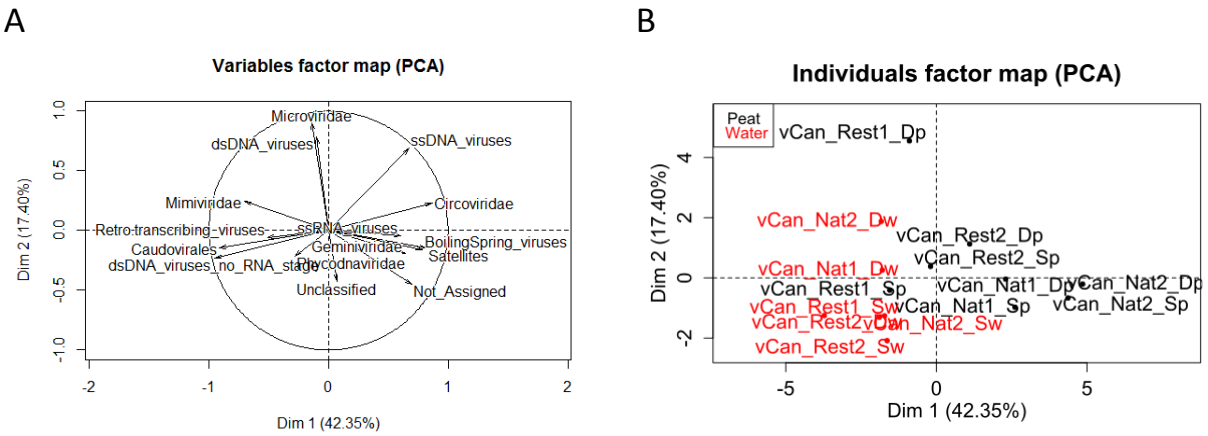


Figure 8 Variable correlation circle (A) and samples projection (B) of the principal component analysis performed depending on the relative composition of the different viral types assigned to the metaviromes from the natural and the restored area in Bois-des-Bel. In the name of the metaviromes, Nat/Rest: Natural or Restored, number= sampling point, S/D: surface or depth, p/w: pore-water or peat matrix origin.

Comparison of the viral communities independent on the taxonomical assignations

In order to detect differences between the viral communities from the natural and the restored area, the whole diversity carried in the metaviromes was compared using commet (k=33, t=4). Commet ouput is a distance matrix that gives a normalized percentage of similarity all the pairs of metaviromes. This matrix is used to represent the similarity between the samples through hierarchical clustering. The pattern of similarity was not very clear as the pore-water viral communities and the peat matrix viral communities from the two areas did not cluster distinctly (Figure 9A). The average normalized similarity between the metaviromes from the left-hand side group (Figure 9A) was almost equal to this from the right-hand side: respectively 39% and 40%, with a 19% of average similarity between metaviromes from different group. We thus split the analyses to obtain a better overview. It appears that the peat matrix and pore water metaviromes from the restored area respectively clustered into two distinct groups (Figure 9B). The viral communities from the natural and restored areas remained distinct when comparing only the metaviromes from the pore-water (Figure 9C) or from the peat matrix (Figure 9D). In addition, the similarity of the metaviromes was slightly higher between the metaviromes from the natural area (pore-water 44% of averaged normalized similarity; peat-matrix 48%) compared with the metaviromes from the restored area (pore-water 38%, peat-matrix 39%), but the inter-group similarity was high (pore-water 16%, peat-matrix 22%) (Figure 9C and D). Finally, the peat-matrix metaviromes from the natural area and the pore-water metaviromes from the natural and the restored area present a resembling pattern of similarity: the viral communities from a same sampling point (surface and depth of the peat column) shared more identical sequences. In contrast, the peat metaviromes from the restored area presented a surface/depth pattern.

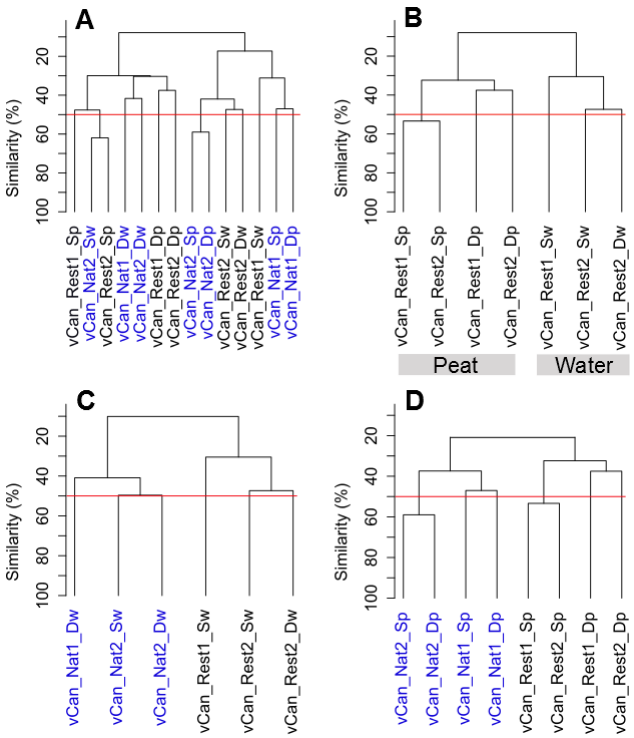


Figure 9 Dendrograms representing the normalized percentage of similarity between the viral communities in all the metaviromes from Bois-des-Bel (A), in the peat matrix and pore-water metaviromes from the restored area (B), in the pore-water metaviromes from the natural (blue) and restored (black) areas (C), and in the peat matrix metaviromes from the natural and restored areas (D). Similarity corresponds to the percentage of shared sequences between pairwise samples (commet, k=33, t=4). Red bars represent 50% of similarity.

Discussion

Effect of climatic changes

Whether peatlands functioning may turn from carbon sink to carbon sources in response to global changes is probably the main concern regarding peatlands ecosystems services (Dise, 2011; Kimmel and Mender, 2010). Rising global temperatures are expected to enhance microbial decomposing activity and carbon emission in the atmosphere, affecting peatlands carbon storage. However, the response of the microbial compartment is physiologically mediated and the C-CO_2 release can also decline, depending on how microbial communities adjust their carbon-use efficiency (Allison et al., 2010). These increased temperatures are also expected to affect the seasonal functioning through the modification of the water regime (precipitations, moisture changes and frozen peat soils thawing) (Dorrepaal et al.; 2009; Yergeau et al., 2010; Mackelprang et al., 2011; Tveit et al. 2013). As the microbial compartment is central to peatlands functioning, and as the viral compartment is dependent on the activity of the microbial compartment, and responds to the fluctuation of the water-table (Chapter 1.A.), it appeared important to document peatland viral ecology in the context of global changes. To address the effect of a moderate rising of air temperature, we compared viral and prokaryote abundances in the piezometer water of the fen and bog profiles of two experimental

stations equipped with Open Top Chambers that increase the mean annual temperature by around 1-2°C. The effect of water-table manipulations was explored in the fen profile of one of these stations (Mukhrino in Russia) where the experimental field warming treatment is coupled to an experimental manipulation of peat monoliths, i.e. + 10 et – 10 cm compared to the control peat surface.

Neither the warming treatment (OTC) nor the modification of the water-table level had significant effects on the viral particles and prokaryotes abundances (VPA and PA) measured in the piezometers from the experimental stations in Frasne (7 years of OTC treatment) and Mukhrino (1 year of OTC treatment). The highest VPA ($1.51x10^6$ VP per mL of piezometer water) measured in the fen profile from Mukhrino is close to the lowest VPA measured in pore-water expressed from the *Sphagnum* in chapter one ($1.7x10^6$ VP per mL of pore water) suggesting that the filters around the piezometers may adsorb particles, yielding lower counts in the water from the piezometers than in the pore-water expressed from the *Sphagnum*.

There are at least two points that plead for more monitoring of viral abundance in these experimental stations. The first point is the absence of difference in PA and VPA with the warming treatment, at least in Frasne, where OTC has been shown to affect microbial biomass and activity at the short term (Jassey, 2011; Gicquel, 2012). The second point concerns the combined effect of OTC and water-table manipulation in Mukhrino. In plots where water-table was lowered (+10 cm raise treatment), there tends to be a higher concentration of DOC (dissolved organic carbon) in the water of the piezometers. In Chapter 1.A. we observed that lower water-table was associated to higher DOC (nutrients more concentrated) and that higher DOC concentrations lead to increased prokaryotic abundance and subsequent higher viral abundance. In the case of the experimentally lowered water-table, the higher DOC is to the contrary associated to a lower VPA. To a further extend, any modification of the water-table (even just "cut") combined with the OTC treatment seems to be associated with both lower VPA and higher DOC (Figure 2D, Figure 4D). These results suggest that the combined effect of water-table manipulation and warming may influence more VPA than a single of these two disturbances. In order to assess modified relationships between DOC and VPA, and the potential loss of viral-particles in response to water-table disturbance and increased temperature it would be important to monitor the temporal dynamics of VPA and PA and DOC in the piezometers to observe whether these treatments affect the seasonality of the viral communities. Since under warming and altered precipitations experiments, the composition of the soil microbial community appeared unchanged while the activity of most phylotypes became affected (Sheik et al., 2011) it would be informative to follow the diversity of the viral communities.

Effect of peatland restoration

Peat exploitation is known to affect peatlands carbon balance. In regard to ecosystems services, the aim of ecological restoration is to return the disturbed ecosystem to a functional state (carbon-sink functioning). In the case of peatlands, restoration success is assessed by the recovery of the hydrological fluxes, the re-establishment of the nutrients cycle, the accumulation of peat (carbon imbalance) and the vegetation composition (*Sphagnum* type and associated vascular plants) (Rochefort et al., 2003). Ultimate success would be the recovery of a similar microbial diversity and activity (Andersen et al., 2006).

In La Guette where the restoration program started in 2013, mean VPA and PA were lower at depth of the newly restored area compared with the depth from the natural area. Drainage lead to vegetation changes, with *Betula spp.* and *Molinia caerulea* invading the peatland and promoting faster organic matter decomposition due to a higher degradability of the litters (Gogo et al., 2011). Lower VPA and PA at depth of the restored area, to the contrary of most undisturbed peatlands (Chapter 1.A., Chapter 3), may suggest that drainage of the peatland have lead to a degradation of the peat, vegetation and available organic matter and to a subsequent decrease in the prokaryotic and viral abundances, biomass and activity as shown in Bois-des-Bel peatland after 5 years of restoration (Andersen et al., 2006). This hypothesis is supported in La Guette with the variation of the physico-chemical features between the restored and the natural area, with higher abundances of the different ions in the natural area suggesting that more diverse sources of nutrients are available for bacterial consumption in the natural area as observed at Bois-des-Bel (Andersen et al., 2006). Nonetheless, in La Guette, higher VPA and PA at the surface of the restored area compared with the natural one suggest a colonization of the newly available *Sphagnum* matrix. In the restored area of Bois-des-Bel, higher microbial biomass at the surface compared with depth was also observed and was expected to be associated with the availability of substrates for bacterial growth thanks to the recovery of the vegetation (Andersen et al., 2010).

In Bois-des-Bel, the restoration program set-up 15 years ago lead to the recovery of the vegetation, especially the *Sphagnum* carpet and restored a carbon balance equivalent to this of the natural peatland (Strack and Zuback, 2013). Comparison of VPA and PA between the natural and the restored area at a single date suggest that the abundance of VPA and PA is higher at depth of the restored area, as observed in most of the undisturbed peatlands (Chapter 1.A., Chapter 3), indicating a possible recovery of the functioning of the microbial communities. This trend should be further confirmed by a monitoring of the temporal dynamics of viral abundance at both depths of the natural and restored area.

However, in Bois-des-Bel, the diversity of the viral communities indicates that the active microbial communities from the two areas remain different. Extraction of the viral particles from the peat-matrix has been shown to yield stronger and more coherent spatial patterns than the extraction from the pore-water (Chapter 3). The Bois-des-Bel metaviromes extracted from the peat matrix present a different spatial pattern in the two areas. In the natural area, the viral communities cluster depending on the sampling point, and the absence of surface/depth pattern seems associated to a high horizontal heterogeneity due to the low water-table (Chapter 3). This influence of the water-table in the homogeneization of the viral communities is emphazised through the fact that the pore-water metaviromes clustered in the same way (depending on the sampling point) but distinctively between the two areas. It is therefore likely that the same "water" fills the pores in the peat, but its content is very influenced locally (Vos et al., 2009) when the water-table is low and the peat is drier (Figure 9). In contrast to the viral communities from the pore water and the peat-matrix from the natural area, the viral communities from the peat matrix in the restored area clustered depending on the depth they originate. This suggests that there is still an influence of the degradation of the peat at depth on the diversity and/or activity of the microbial communities that host the virus from the restored area. Compared with the results from the water-table manipulation in Mukhrino, any modification of the peat column seems to affect the peat substrate and to impact microbial activity and subsequent viral abundance and diversity.

Références du Chapitre 4

Abergel C., Claverie, J.M. 2014. Pithovirus Sibericum: Awakening of a Giant Virus of More than 30,000 Years. *Médecine Sciences : M/S* 30 (3): 329–31. doi:10.1051/medsci/20143003022.

Allison, S. D., Wallenstein, M. D. and Bradford, M. A. 2010. Soil-Carbon Response to Warming Dependent on Microbial Physiology. *Nature Geoscience* 3 (5): 336–40. doi:10.1038/ngeo846.

Andersen R., Francez A.-J., Rochefort L. 2006. The physico-chemical and microbiological status of a restored bog in Québec: identifications of relevant criteria to monitor success. *Soil Biology and Biochemistry* 38: 1375-1387.

Andersen R., Grasset L., Thormann M.N., Rochefort L., Francez A.-J. 2010. Changes in microbial community structure and function following *Sphagnum* peatland restoration. *Soil Biology and Biochemistry* 42: 291-301.

Artz R.R.E., Chapman S.J., Siegenthaler A., Mitchell E.A.D., Buttler A., Bortoluzzi E., Gilbert D., Yli-Petays M., Vasander H., Francez A.-J. 2008. Functional microbial diversity in cutover peatlands responds to vegetation succession and is partly directed by labile carbon. *J. Appl. Ecol.* 45: 1799-1809.

Bullock, J. M., Aronson, J., Newton, A. C., Pywell, R. F. and Rey-Benayas, J. M. 2011. Restoration of Ecosystem Services and Biodiversity: Conflicts and Opportunities. *Trends in Ecology & Evolution* 26 (10): 541–49. doi:10.1016/j.tree.2011.06.011.

Delarue F., 2010. *Dynamique des matieres organiques labiles et recalcitrantes dans la tourbiere de Frasne (Jura): impact des conditions hydriques et d'un rechauffement simule in situ.* Geochemistry. Universite d'Orleans, 2010. French.<NNT:2010ORLE2054>.<tel-00574496v3>

Dise N.B., Phoenix G.K. 2011. Peatlands in a changing world. *New Phytol.* 191: 309-311.

Dorrepaal E., Toet S., van Loegtestijn R.S.P., Swart E., van de Weg M.J., Callaghan T.V., Aerts R. 2009. Carbon respiration from subsurface peat accelerated by climate warming in the subarctic. *Nature* 460: 616-619.

Francez A.-J. 2000. La dynamique du carbone dans les tourbières à Sphagnum, de la sphaigne à l'effet de serre. *L'Année biologique* 39: 205-270.

Francez, A.-J. and Vasander, H. (1995). Peat accumulation and peat decomposition after human disturbance in French and finnish mires. *Acta oecologica* 16, 599-608.

Frolking, S., Talbot, J., Jones, M. C., Treat, C. C., Kauffman, J. B., Tuittila, E-S. and Roulet, N. 2011. Peatlands in the Earth's 21st Century Climate System. *Environmental Reviews* 19 (NA): 371–96. doi:10.1139/a11-014.

Gicquel A. 2012. *Impact des changements globaux sur le fonctionnement des tourbières: couplage C-N-S et interactions biotiques.* Thèse de l'Université de Rennes 1 : 221 p.

Hooijer, A., S. Page, J. G. Canadell, M. Silvius, J. Kwadijk, H. Wösten, and J. Jauhiainen. 2010. Current and Future CO_2 Emissions from Drained Peatlands in Southeast Asia. *Biogeosciences* 7 (5): 1505–14. doi:10.5194/bg-7-1505-2010.

Jassey, V. E. J., Chiapusio, G., Binet, P., Buttler, A., Laggoun-Défarge, F., Delarue, F., Bernard, N. et al. 2013. Above- and Belowground Linkages in Sphagnum Peatland: Climate Warming Affects plant-microbial interactions. Glob Chang Biol 19:811–23. doi: 10.1111/gcb.12075Plant-Microbial Interactions. *Global Change Biology* 19 (3): 811–23. doi:10.1111/gcb.12075.

Jassey V. 2011. *Impact d'un réchauffement climatique sur le fonctionnement de la sphagnosphère : relations polyphenols-communautés microbiennes.* Thèse de l'Université de Franche-Comté, Besançon (France): 238 p.

Junk, W. J., An, S., Finlayson, C. M., Gopal, B., Květ, J., Mitchell, S. A., Mitsch, W. J. and Robarts, R. D. 2012. Current State of Knowledge Regarding the World's Wetlands and Their Future under Global Climate Change: A Synthesis. *Aquatic Sciences* 75 (1): 151–67. doi:10.1007/s00027-012-0278-z.

Kimmel, K., and Mander, U. 2010. Ecosystem Services of Peatlands: Implications for Restoration. *Progress in Physical Geography* 34 (4): 491–514. doi:10.1177/0309133310365595.

Legendre M., Lartigue A., Bertaux L., Jeudy S., Bartoli J., Lescot M., Alempic J.-M., et al. 2015. In-Depth Study of Mollivirus Sibericum , a New 30,000-Y-Old Giant Virus Infecting Acanthamoeba. *Proceedings of the National Academy of Sciences* 112 (38): 201510795. doi:10.1073/pnas.1510795112.

Mackelprang, R., Waldrop, M. P., DeAngelis, K. M., David, M. M., Chavarria, K. L., Blazewicz, S. J., Rubin, E. M. and Jansson, J. K. 2011. Metagenomic Analysis of a Permafrost Microbial Community Reveals a Rapid Response to Thaw. *Nature* 480 (7377): 368–71. doi:10.1038/nature10576.

Mitra, S., Wassmann, R., and Vlek, P. L. G. 2005. An Appraisal of Global Wetland Area and Its Organic Carbon Stock. *CURRENT SCIENCE* 88 (1): 25–35. Mitsch, WJ, and JG Gosselink. 2007. "Wetlands (4th Edition)." *New York: Wiley.*

Poulin, M., Andersen, R. and Rochefort, L. 2013. A New Approach for Tracking Vegetation Change after Restoration: A Case Study with Peatlands. *Restoration Ecology* 21 (3): 363–71. doi:10.1111/j.1526-100X.2012.00889.x.

Rochefort L. 2001. Restauration écologique in Serge Payette et Line Rochefort (dir.), *Écologie des tourbières du Québec-Labrador*, Les presses de l'Université Laval, Québec, pp 449-505.

Rochefort, L., Quinty, F., Campeau, S., Johnson, K. and Malterer, T. 2003, North American approach to the restoration of *Sphagnum* dominated peatlands, Wetlands Ecology and Management, 11: 3-20.

Serreze, M. C., J. E. Walsh, F. S. Chapin III, T. Osterkamp, M. Dyurgerov, V. Romanovsky, W. C. Oechel, J. Morison, T. Zhang, and R. G. Barry. 2000. Observational Evidence of Recent Change in the Northern High-Latitude Environment. *Climatic Change* 46 (1-2): 159–207. doi:10.1023/A:1005504031923.

Sheik, C. S., Beasley, W. H., Elshahed, M. S., Zhou, X., Luo, Y. and Krumholz, L. R. 2011. Effect of Warming and Drought on Grassland Microbial Communities. *The ISME Journal* 5 (10): 1692–1700. doi:10.1038/ismej.2011.32.

Strack M. 2008. *Peatlands and climate change.* International Peat Society, Jyväskyla, 223 p.

Strack, M., and Zuback, Y. C. A.. 2013. Annual Carbon Balance of a Peatland 10 Yr Following Restoration. *Biogeosciences* 10 (5): 2885–96. doi:10.5194/bg-10-2885-2013.

Tveit A., Schwacke R., Svenning M.M., and Urich T. 2013. Organic Carbon Transformations in High-Arctic Peat Soils: Key Functions and Microorganisms. *The ISME Journal*, September. doi:10.1038/ismej.2012.99.

Vile M.A., Wieder R.K., Zivkovic T., Scott K.D., Vitt D.H., Hartsock J.A., Iosue C.L., Quinn J.C., Petix M., Fillingim H.M., Popma J.M.A., Dynarski K.A., Jackman T.R., Albright C.M., Wykoff D.D. 2014. N2-fixation by methanotrophs sustains carbon and nitrogen accumulation in pristine peatlands. *Biogeochemistry* 121: 317-328.

Vos M., Birkett P.J., Birch E., Griffiths R.I., and Buckling A. 2009. Local Adaptation of Bacteriophages to Their Bacterial Hosts in Soil. *Science (New York, N.Y.)* 325 (5942): 833. doi:10.1126/science.1174173.

Ward S.E., Ostle N.J., Oaldey S., Quirk H., Henrys P.A., Bardgett R.D. 2013. Warming effects on greenhouse gas fluxes in peatlands are modulated by vegetation composition. Ecol. Letters

Waddington M., Rottenberg P.A., Warren F.J. 2001. Peat CO_2 production in a natural and cutover peatland: implications for restoration. *Biogeochemistry* 54: 115-130.

Yergeau, E., Hogues, H., Whyte, L. G. and Greer, C. W. 2010. The Functional Potential of High Arctic Permafrost Revealed by Metagenomic Sequencing, qPCR and Microarray Analyses. *The ISME Journal* 4 (9): 1206–14. doi:10.1038/ismej.2010.41.

Synthèse générale:

Avancées fondamentales et perspectives

Synthèse Générale

1. Introduction

L'accumulation de tourbe, et donc le stockage du carbone, dans les tourbières à Sphaignes est liée à la décomposition incomplète de la matière organique produite par photosynthèse. Dans les tourbières boréales et tempérées, de nombreux facteurs limitent en effet l'activité des procaryotes, comme les conditions physico-chimiques (faible température, saturation en eau, oligotrophie, acidité) et les interactions biotiques (production de polyphénols par les Sphaignes). Néanmoins, seulement 10 à 15% de la production primaire sont accumulés sous forme de tourbe.

Cependant, en dépit de l'importance de l'activité microbienne dans le fonctionnement des tourbières à Sphaignes, l'influence des virus n'a jamais été étudiée, alors qu'ils contrôlent et structurent les communautés microbiennes (contrôle top-down) et, de ce fait, le recyclage du carbone.

Le but de ce travail de thèse était donc de décrire la diversité, la structuration et le fonctionnement du compartiment viral des tourbières à Sphaignes, en lien avec la physico-chimie et le fonctionnement des communautés de procaryotes au niveau de l'acrotelme et de l'interface acrotelme/mesotelme, là où l'activité de décomposition est la plus intense.

Les recherches se sont appuyées sur l'analyse des abondances virale et procaryotique ainsi que sur le séquençage de l'ADN de communautés de petits virus (< 0,2 µm) libres (métaviromes), en fonction des saisons et par comparaison de plusieurs tourbières situées sous des latitudes différentes et plus ou moins éloignées géographiquement. L'étude des fluctuations d'abondance des particules virales et des procaryotes s'est accompagnée d'une caractérisation physico-chimique des différents biotopes. Le fonctionnement des communautés microbiennes et plus particulièrement des procaryotes a été analysé par une combinaison de métatranscritomique et de métagénomique à une échelle spatiale très fine de l'acrotelme et du mésotelme pour déterminer sa structuration taxonomique et fonctionnelle en fonction des stades dynamiques de la tourbière et du gradient de profondeur lié aux conditions d'aérobiose/anaérobiose, ainsi qu'à la concentration en carbone organique dissous.

Ces différentes approches ont permis de proposer un fonctionnement du compartiment viral, en lien avec l'activité procaryote et la physico-chimie et ce à différentes échelles, permettant de répondre aux questions suivantes :

i) La structuration de la communauté microbienne basée sur l'expression de gènes (métatranscriptome) en fonction du potentiel génétique (métagénomes) des communautés d'hôtes entre deux profondeurs proches (acrotelme et mésotelme) de la surface des deux principaux stades dynamiques (le fen et le bog) de la tourbe. Quels facteurs (espèce de Sphaigne, profondeur) sélectionnent l'activité et le potentiel génétique des procaryotes (diversité taxonomique et fonctionnelle)?

ii) Comment interpréter la dynamique spatio-temporelle de l'abondance et de la diversité des communautés virales de la même tourbière ? Retrouve-t-on une influence des mêmes facteurs ?

iii) Les modalités de fonctionnement des communautés virales et des interactions hôtes/virus de cette tourbière sont-elles généralisables à l'ensemble des tourbières ? A quelles échelles les communautés virales sont-elles structurées : celles du point d'échantillonnage, du site, de l'écosystème ?

iv) Les perturbations anthropiques modifient-elles les caractéristiques de ce fonctionnement ?

2. Avancées fondamentales

2.1 Fonctionnement des communautés d'hôtes

Les virus sont dépendants de l'activité de leurs hôtes procaryotes et eucaryotes. Au niveau de l'acrotelme et du mésotelme des deux principaux stades d'une tourbière à *Sphagnum*, la diversité active et fonctionnelle des microorganismes est sélectionnée par les conditions environnementales liées à la profondeur, avec des conditions plus aérobies dans les 5 à 10 premiers centimètres de la tourbe, où l'on retrouve plus de Protozoaires, plus de producteurs primaires, mais aussi des conditions plus stressantes pour les microorganismes, probablement liées à la dessiccation des Sphaignes. Entre 10 et 15 centimètres sous la surface des capitula de Sphaignes, c'est-à-dire dans le mésotelme, des conditions où alternent aérobiose et anaérobiose supportent une plus forte abondance de procaryotes, une plus forte activité hétérotrophe, et l'activité d'archées méthanogènes. La sélection à la fois taxonomique et fonctionnelle de l'activité des microorganismes semble se faire de manière identique dans les deux principaux stades (Figure 1).

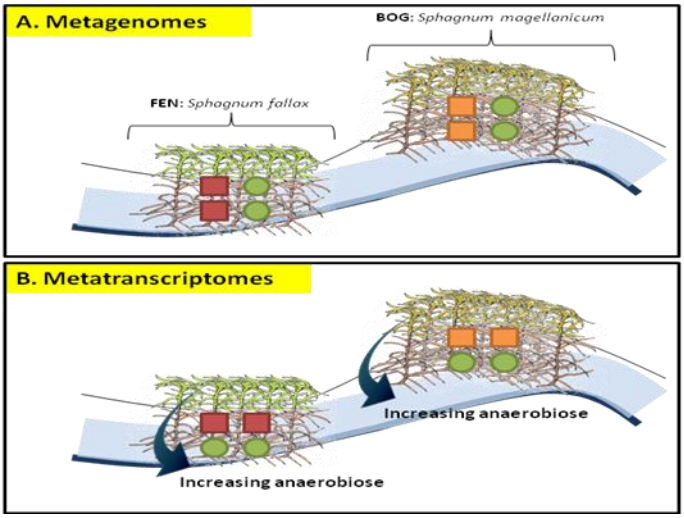


Figure 1 Patrons de diversité taxonomique (couleur) et fonctionnelle (forme) des communautés de procaryotes en fonction du stade, de la profondeur.

Cependant, en lien avec les conditions environnementales comme le pH et la disponibilité en nutriments, on observe une diversité taxonomique procaryote spécifique à chaque espèce de Sphaigne. Cette différence ne semble pas s'appliquer à la diversité fonctionnelle, puisque les communautés des deux stades présentent de fortes similarités fonctionnelles. Ceci suggère l'existence d'un « core-microbiome » fonctionnel des communautés procaryotes des tourbières à Sphaigne, c'est-à-dire un set de gènes/fonctions permettant leur maintien et leur évolution dans la Sphagnosphère. Ces fonctions seraient portées par des organismes différents en fonctions des espèces de Sphaignes, mais leur expression est sélectionnée de la même manière entre les deux stades en fonction des conditions environnementales liées à la profondeur et à la variation de la nappe d'eau.

2.2 L'abondance et la diversité des particules virales varient de manière temporelle en fonction des fluctuations de la nappe d'eau et de la disponibilité en nutriments

La fluctuation de la nappe n'affecte pas que l'activité microbienne, c'est aussi l'un des principaux facteurs structurant l'abondance et la diversité virale présente dans l'eau des Sphaignes. Le suivi saisonnier de l'abondance virale dans l'acrotelme et le mésotelme du fen et du bog, et l'analyse des fluctuations physico-chimiques montrent en effet que la variation temporelle de la nappe d'eau, inondant plus ou moins les deux stades dynamiques, affecte la concentration des nutriments, tels que le carbone organique dissous. En été et au début de l'automne, la nappe d'eau étant plus basse, et la production végétale moins élevée, les nutriments disponibles pour les communautés d'hôtes permettent une plus forte activité microbienne, qui se traduit par des abondances plus élevées de procaryotes et de virus (Figure 2). Ces variations de nappe, de nutriments et d'abondance des virus

et des procaryotes sont moins marquées en profondeur. La profondeur apparaît donc comme un milieu plus « tamponné » qui supporte généralement de plus fortes abondances de procaryotes et donc de virus[2].

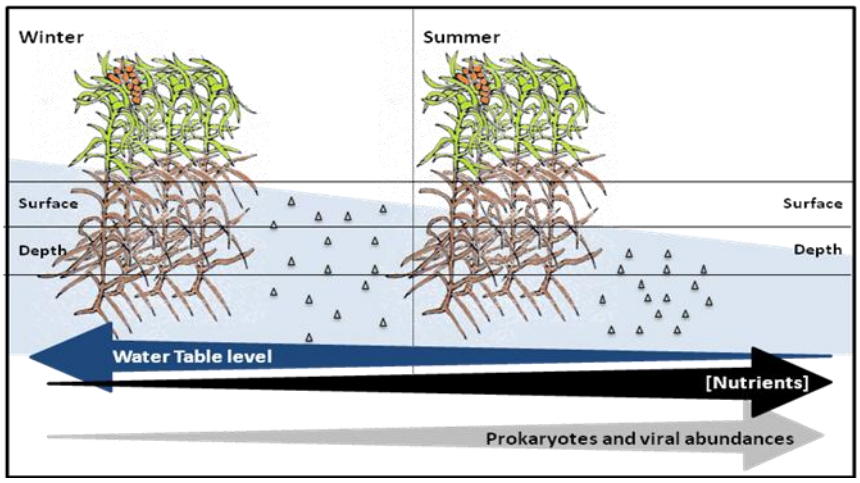


Figure 2 Effet de la variation saisonnière de la nappe d'eau sur la concentration des nutriments et l'abondance des procaryotes et des particules virales dans les échantillons d'eau extraite des Sphaignes.

La plus forte abondance de procaryote et de virus en été par rapport à l'hiver est associée à une différence profonde de la diversité virale qui présente deux types de communautés génétiquement distinctes, celles d'été/automne (août/octobre) et celles de l'hiver/printemps (juin 2011, mars 2012). Cette saisonnalité de la diversité virale s'accompagne donc potentiellement d'une récurrence annuelle.

Cette succession saisonnière pourrait s'expliquer par une transition de stratégie lysogénique vers une stratégie lytique quand l'hôte devient actif en été, et par un retour au cycle lysogénique à la fin de l'automne ou au début du printemps (Figure 3).

Ces résultats ont donc permis de d'intégrer la variation temporelle des paramètres environnementaux, et de les relier au fonctionnement du compartiment microbien en intégrant celui du compartiment viral. Par contraste avec cette saisonnalité, la diversité des communautés virales d'hiver présentes dans l'eau des Sphaignes ne présentait pas de différences importantes entre le fen et le bog, malgré la distinction taxonomique des communautés d'hôtes, ce qui renforce le rôle de la nappe et de ses fluctuations en fonction des conditions climatiques comme un des éléments structurant l'activité microbienne et la diversité virale.

[2] Cette observation d'une plus forte abondance en profondeur est également liée au fait que nous travaillons dans la couche très supérieure de la tourbe (acrotelme/mesotelme), où la matière organique est encore facilement assimilable par les micro-organismes. Plusieurs travaux ont en effet démontré qu'à des profondeurs plus élevées où la nappe ne fluctue plus (catotelme) et où la matière organique est plus dégradée, l'abondance des micro-organismes est moins élevée qu'en surface.

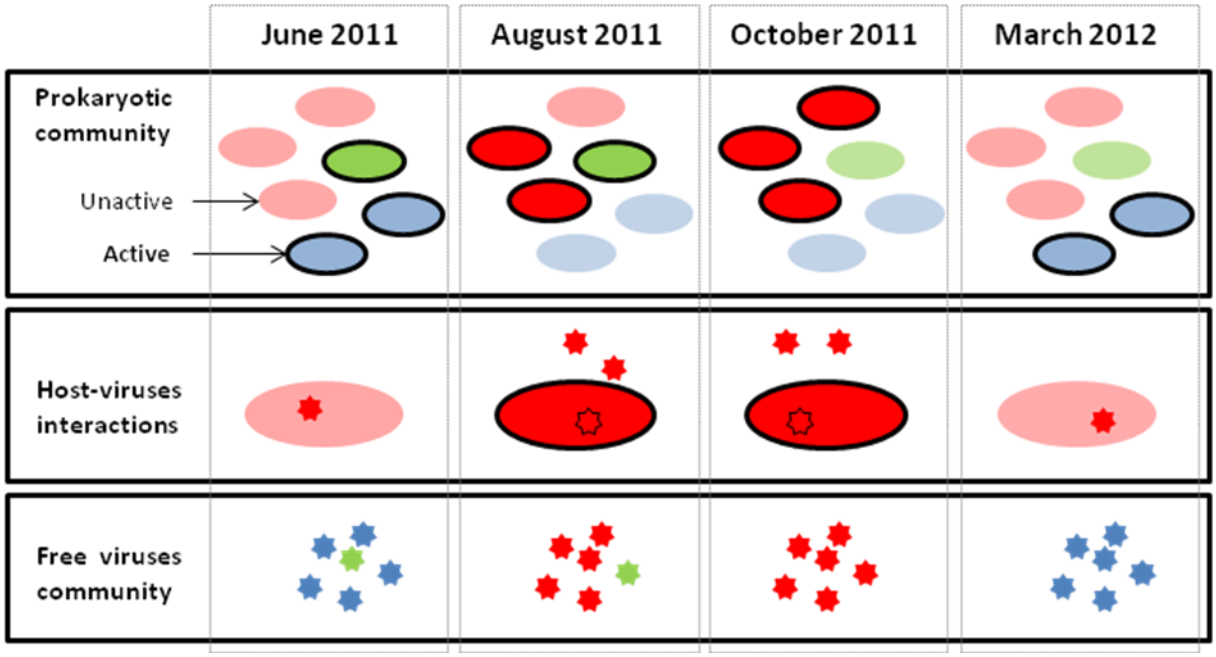


Figure 3 Effet de la variation saisonnière sur la diversité des hôtes procaryotes actifs, sur les stratégies d'infection virales pour l'un des hôtes de la communauté, et sur la diversité des communautés virales libres. Les couleurs des communautés d'hôtes représentent leur diversité, le contour noir autour de l'hôte indique qu'il est actif.

2.3 L'abondance et la diversité virale varient en fonction des tourbières et présentent une structure spatiale très fine

A fine échelle spatiale, la structuration des communautés microbiennes a été étudiée à partir d'échantillons de la matrice de tourbe et d'échantillons d'eau extraite des Sphaignes. Nous avons adapté un protocole d'extraction de virus du sol pour tenter d'obtenir une meilleure résolution spatiale de la diversité et de l'abondance virale en travaillant sur les particules virales extraites de la matrice de tourbe.

L'abondance des particules virales varie significativement entre les 5 tourbières échantillonnées au Canada, en Finlande, en France et sur l'île Subantarctique d'Amsterdam. Cependant, certaines caractéristiques communes ont été observées dans la plupart des sites, comme la plus forte abondance des particules virales et des procaryotes en profondeur, ainsi que de plus fortes abondances de ces entités biologiques associées à de plus fortes concentrations en carbone organique dissous et en sulfate. Ces observations suggèrent le fait que des conditions environnementales un peu plus stables, ou qu'une plus faible prédation par les Protozoaires dans le mésotelme s'associent à une plus forte abondance de procaryotes et donc de virus. Elles permettent aussi de soutenir l'hypothèse qu'une plus forte activité procaryote génère une plus forte production virale, et que cette augmentation d'activité varie positivement avec le carbone organique dissous. De plus, dans le fen et le bog de Lakkasuo (FIN) et des Pradeaux (F), aux végétations dominantes

comparables, l'abondance des particules virales et des procaryotes présentaient peu de différences, renforçant le fait qu'à l'échelle intra-site, la fluctuation temporelle des abondances des entités biologiques est plus marquée qu'entre les deux stades dynamiques.

Les tourbières du Canada et de Finlande présentent par ailleurs une diversité virale très distincte.

En Finlande, la diversité des communautés virales issues de la matrice de tourbe est structurée par le stade ou l'espèce de Sphaigne, et au sein d'un stade, la diversité est structurée en fonction de la profondeur. Ceci confirme donc les résultats obtenus sur la diversité microbienne, avec des communautés taxonomiquement différentes en fonction de chaque espèce de Sphaignes, et une sélection similaire dans les deux stades de l'activité des hôtes en fonction des conditions environnementales liées à la profondeur. En comparaison, la diversité virale de l'eau de Sphaigne est très proche de celle des communautés virales de la matrice de tourbe du fen (stade plus inondé), renforçant l'hypothèse de l'importance de la fluctuation de nappe sur la diversité virale des communautés d'eau de Sphaigne.

Dans les sites plus secs (Canada), lorsque la nappe d'eau est plus basse, la dessiccation génère une très forte hétérogénéité spatiale horizontale et gomme le gradient vertical d'oxie/anoxie, ce qui se traduit par une plus forte ressemblance entre les différents niveaux d'une même colonne de tourbe que pour les mêmes niveaux de différents points d'échantillonnage. Cette hétérogénéité spatiale génère une forte différentiation entre les communautés virales attachées à la matrice de tourbe et celles de l'eau extraite des Sphaignes, ce qui confirme la spécificité des Sphaignes en tant que micro-habitats.

2.4 Les communautés virales des tourbières se distinguent de celles des autres écosystèmes

Dans le contexte de 'sphagnosphère' définie comme la sphère d'influence de la sphaigne, et caractérisée par de nombreuses interactions entre les microorganismes et la Sphaigne, le fait que le fonctionnement des communautés microbiennes et virales soient comparables entre des sites très éloignés, et que la diversité taxonomique des communautés microbiennes soit spécifique d'une espèce de Sphaigne amène à faire l'hypothèse d'une similarité de communautés virales d'une même espèce de sphaigne dans des tourbières géographiquement éloignées. Certaines de nos observations semblent conforter cette hypothèse et amènent à se poser la question de la vitesse à laquelle évoluent les communautés virales.

Dans l'état actuel de nos connaissances, l'analyse des 31 communautés virales de tourbières à Sphaignes nous permet de souligner la spécificité des tourbières qui se distinguent très fortement des autres biotopes ou écosystèmes dans lesquels la diversité virale a été analysée. Cette spécificité renforce l'idée d'un réseau d'interactions et d'un fonctionnement propres à la Sphagnosphère.

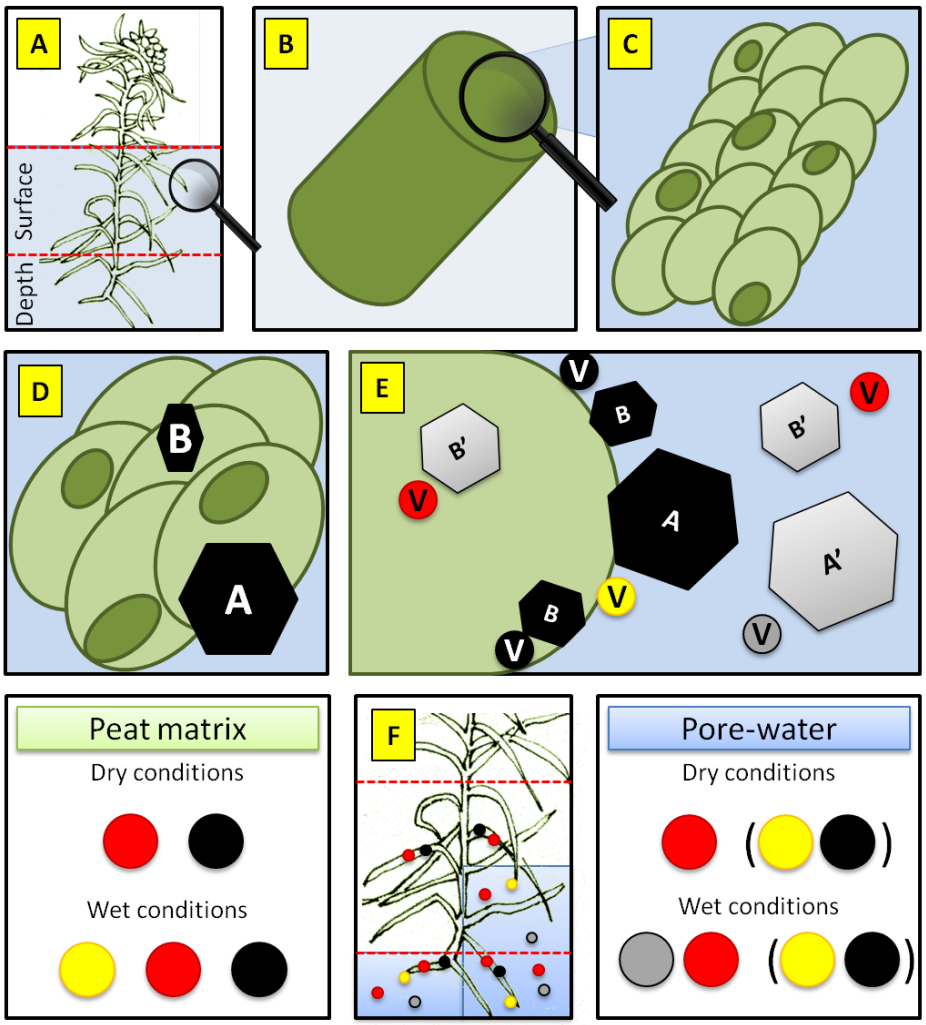


Figure 4 Diversité virale observable dans les métaviromes de la matrice de tourbe et d'eau de Sphaigne. L'eau extraite des Sphaignes provient de différentes sources : il s'agit de l'eau de la nappe (A), qui par capillarité imbibe le tapis de Sphaignes (B). Cette eau peut pénétrer jusqu'à l'intérieur des hyalocystes, grâce aux pores que présentent ces cellules (C). Ainsi, en fonction du niveau de la nappe et de la teneur en eau, l'eau des hyalocystes est plus ou moins homogénéisée avec le milieu extérieur. La taille des pores conditionne l'entrée ou non des virus et des microorganismes (procaryotes, protozoaires, petits métazoaires) dans les hyalocystes (D). Les microorganismes vivent donc à l'intérieur ou à l'extérieur des cellules et des « tiges » et « feuilles » des Sphaignes (E). L'extraction de l'eau de Sphaigne libère donc les communautés virales qui vivent de manière libres dans la Sphaigne, l'extraction depuis la matrice de tourbe donne en plus accès aux particules virales adsorbeées non récupérables par simple extraction de l'eau. Ainsi, quand la tourbe est très humide et que la nappe est haute, on récupère les virus présents à la fois dans les hyalocystes, l'eau d'imbibition et l'eau de la nappe, ainsi que quelques virus attachés aux surfaces des cellules de Sphaignes (F). La diversité des communautés de virus de la matrice de tourbe est elle aussi influencée par le niveau de la nappe et l'eau d'imbibition qui permettent de diffuser les particules virales. Dans le cas de conditions très sèches, elle sera essentiellement représentative des virus attachés à l'intérieur des hyalocytes. Il parait ici important de noter que si on récupère une très grande proportion des protozoaires présents dans les hyalocystes en pressant simplement les sphaignes, il semble que les particules virales s'attachent plus fortement à la surface des cellules, peut-être en raison de la nature de la capside. En effet, les protéines qui les constituent sont souvent des molécules plus chargées que les phospholipides des membranes lipidiques. Dans cette hypothèse, cet attachement serait d'autant plus fort que les particules sont petites.

2.5 Les perturbations anthropiques semblent affecter les communautés virales de tourbières et par conséquent les communautés microbiennes

Dans un contexte de réchauffement climatique simulé et d'une modification du niveau de la nappe d'eau, il semble que l'abondance virale soit diminuée, et qu'elle varie négativement avec le carbone organique dissous. Cette inversion de la relation entre l'abondance virale et le carbone organique dissous peut avoir de nombreuses origines que des études complémentaires permettront de préciser.

Par ailleurs, la comparaison des communautés virales d'une tourbière naturelle dont une partie a été exploitée, puis restaurée depuis une quinzaine d'année montre que si le ré-haussement de la nappe permet de ré-humecter les couches supérieures de la tourbe, la remise en eau n'a pas permis de reconstituer une diversité microbienne et virale comparable à celle des tourbières naturelles. De plus, une forte différence surface/profondeur des communautés virales de la matrice de tourbe exploitée indique que les communautés microbiennes présentent dans les Sphaignes qui repoussent sont différentes de celles de la tourbe sous-jacente qui avait été dégradée lors de l'exploitation du site.

2.6 Décrire de nouveaux génomes viraux

La description de nouveaux génomes viraux n'était pas l'objectif principal de ce travail. Elle illustre pourtant les limites qui existent encore à une compréhension globale de la diversité contenue dans les métaviromes, et la richesse de cette information. Cette description de nouveaux génomes de virus constitue donc un élément important puisque l'analyse des métaviromes est actuellement très limitée par le manque de génomes de références qui permettraient une comparaison plus aisée, et surtout plus fonctionnelle des communautés virales. C'est ce manque de références qui a demandé un travail important de réflexion sur la façon dont on pouvait extraire un maximum d'information de la diversité totale portée dans les métaviromes utilisés pour la dynamique saisonnière. Dans le cas des *Microviridae*, ce référencement est donc d'autant plus intéressant qu'ils infectent un groupe d'hôtes identifié (les bactéries), via des cycles lytiques et tempérés (Krupovic & Forterre, 2011), et que ces virus constituent une proportion importante des communautés virales des tourbières.

Perspectives

Il reste bien des choses à faire suite à ce travail de thèse. Tout d'abord, comme nous venons de le voir, parce que les métaviromes ne servent pas qu'à l'analyse spatiale de la diversité virale, ils sont aussi des outils complexes qui contiennent une telle quantité d'information qu'elle nécessite en elle-même qu'une thèse leur soit consacrée. Deuxièmement parce que si nous avons plus ou moins réussi à intégrer le fonctionnement du compartiment viral à celui des procaryotes, il nous reste à comprendre comment ce compartiment s'insère dans le fonctionnement de la boucle microbienne et dans le cycle du carbone. Enfin, parce que nous n'avons qu'à peine effleuré la réponse du compartiment viral aux perturbations d'origine anthropique alors qu'il semble représenter un outil intéressant pour appréhender à une échelle de temps court l'effet d'un réchauffement climatique sur le fonctionnement d'un écosystème, et qu'il permet d'évaluer l'impact d'une mesure de restauration.

3.1 Diversité virale

Nous avons terminé la synthèse de nos résultats en traitant de la description et de l'analyse de la diversité taxonomique et fonctionnelle contenue dans les métaviromes, mais elle représente à court-terme la première voie à explorer. Tout d'abord, il sera important d'intégrer les métaviromes encore non séquencés à nos analyses, pour confirmer le fonctionnement du compartiment viral sur la totalité du gradient géographique (île d'Amsterdam, et deux tourbières de France, dont l'une présente la même végétation que celle de Finlande, et celle ayant été étudiée pour la dynamique spatio-temporelle des virus). Par ailleurs, en raison de la faible proportion de séquences assignées à des génomes viraux référencés, il reste compliqué de donner un sens écologique à la présence de telle ou telle séquence dans les métaviromes. Il est donc important de continuer à essayer de décrire de nouveaux génomes à partir de ce matériel.

La forte similarité de certains métaviromes des tourbières de Finlande et de France laisse espérer que l'on retrouve des séquences proches dans les jeux de données, et que certaines, voire une grande partie de ces séquences similaires soient associées à des gènes de protéines de capside ou de réplication. Il serait donc envisageable d'assembler non pas des génomes mais des gènes à partir des séquences Illumina (moins de biais), ce qui permettrait de constituer une base de données contre laquelle nous pourrions comparer et identifier les séquences des autres métaviromes de tourbière.

Etant donnée la spécificité des communautés virales de tourbières, l'analyse de la diversité de ces séquences identifiées pourrait permettre de travailler sur les différences d'évolution des séquences

de ces gènes, comme suggéré par Krupovic et al., (2011) dans des écosystèmes qui semblent caractérisés par une très forte co-évolution des interactions procaryotes/plantes (Chapitre 2). Les protéines de réplication semblent en effet évoluer moins rapidement que les protéines de capsides car elles ne sont pas soumises aux mêmes contraintes de sélection (ici la reconnaissance hôte/virus). Dans le cadre de la dynamique spatio-temporelle, il serait également intéressant d'analyser les zones hypervariables trouvées dans les grandes protéines de capsides et potentiellement impliquées dans la reconnaissance de l'hôte, car une étude tend à montrer que les procaryotes échantillonnés à une date sont plus sensibles à l'infection par la communauté virale du mois suivant qu'à celle du mois précédent l'échantillonnage (Koskella, 2013). Ceci permettrait de mieux comprendre le cycle annuel et la spécificité des communautés virales de tourbières, et probablement d'expliquer les différences à très fine échelle (surface/profondeur d'un même site).

Toutes ces perspectives concernant la diversité virale s'inscrivent dans les limites actuelles des connaissances des interactions virus/hôtes dans les communautés environnementales, mais aussi dans les limites des concepts théoriques de l'écologie virale qui ont motivé l'étude des virus des tourbières, à savoir le rôle des virus dans le cycle du carbone de ces écosystèmes, qui incluent les concepts du shunt-viral et de kill-the-winner. Ces limites étant actuellement liées au manque de connaissances et à la difficulté d'estimer les aspects de l'infectivité et de la disparition des particules virales (decay), de la spécificité d'hôte et de l'adsorption des particules virales sur les hôtes (Winter et al., 2010 ; Koskella and Meaden, 2013).

3.3 Lien entre la saisonnalité et la lysogénie

Il serait important de quantifier la variation temporelle de la proportion de bactéries lysogènes dans les communautés procaryotes de l'acrotelme et du mésotelme afin de confirmer l'hypothèse d'un changement d'équilibre entre ces stratégies en été en relation avec la productivité bactérienne et le carbone organique dissous.

3.4 Lien entre la boucle virale et la boucle microbienne

Il apparait ensuite important de relier le fonctionnement du compartiment viral et du compartiment procaryote aux boucles virale et microbienne, en quantifiant les transferts de carbone entre ces compartiments afin de mieux connaître le rôle des virus dans le recyclage du carbone des tourbières. Deux points intéressants sont à relever dans les résultats obtenus au cours de cette thèse : premièrement l'abondance des Protozoaires est plus forte dans l'acrotelme, où l'abondance bactérienne et virale est généralement plus faible, alors qu'en profondeur on observe plus de virus et

de procaryotes et moins de Protozoaires, ce qui suggère un contrôle top-down assuré en surface (acrotelme) par les protozoaires au sein de la boucle microbienne et, plus en profondeur (mésotelme), par les virus au sein de la boucle virale.

La prise en compte du contrôle top-down des virus et des Protozoaires sur la biomasse et l'activité bactériennes, et l'effet des virus sur la diversité des communautés d'hôtes pourrait se faire par quantification des flux de carbone en conditions contrôlées en utilisant des molécules marquées. L'usage de la nanoSIM pourrait utilement compléter cette approche en permettant la visualisation des organismes et les réseaux trophiques.

3.2 Diversité microbienne fonctionnelle

De la même manière que les métaviromes, les 12 métagénomes et 12 métatranscriptomes produits méritent que l'on approfondisse leur étude. La difficulté à appréhender ces jeux de données est liée au fait qu'ils contiennent potentiellement « tout » mais pas forcément en grande quantité, ce qui fait que par rapport à des études qui ciblent un mécanisme, leur analyse, notamment statistique, est plus complexe. D'après les résultats obtenus jusqu'à ce jour, la poursuite des travaux sur ces jeux de données devrait s'orienter sur des gènes et des ARNm impliqués dans la dégradation de composés carbonés plus ou moins complexes, en fonction de la profondeur, ainsi que sur la détermination de leur diversité. L'analyse phylogénétique de cette diversité pourrait permettre de confirmer ou non l'hypothèse selon laquelle les mêmes fonctions sont portées dans les microbiomes de toutes les Sphaignes mais par des organismes différents.

Dans cette même perspective, la caractérisation du potentiel « core-microbiome » fonctionnel pourrait s'effectuer en déterminant pour quels grands types de métabolismes les principaux génomes ont été recrutés, afin de mieux connaître le cœur fonctionnel de ce microbiome.

La réponse expérimentale de la communauté microbienne à l'ajout d'un substrat et l'utitilisation de proxis (ici la diversité des gènes d'ARNr) permettrait d'identifier la façon dont la communauté est modifée, et de déterminer quels sont les organismes potentiellement impliqués dans l'utilisation de ce substrat.

3.5 Utiliser les communautés virales et microbiennes pour évaluer la restauration des tourbières anciennement exploitées

Comme nous l'avons vu dans le chapitre 3, l'étude de la diversité virale qui reflète la diversité et l'activité des hôtes est un bon moyen de valider des hypothèses sur les facteurs qui structurent les

communautés microbiennes. Dans le cas des tourbières exploitées puis restaurées, la recolonisation par la végétation ne valide le recouvrement d'un fonctionnement « normal » qu'au niveau de la surface de la colonne de tourbe, alors que ce fonctionnement peut rester perturbé plus en profondeur. La comparaison de métaviromes de différentes profondeurs (mésotelme et au delà catoleme) de sites naturels, exploités et restaurés est un bon outil pour évaluer le retour à un fonctionnement d'avant perturbation. Cependant, tant que l'identification des hôtes reste compliquée avec cet outil, il pourrait être envisageable de s'appuyer sur d'autres méthodes moins couteuses pour comparer le fonctionnement de ces communautés, tout en utilisant les virus.

Ces méthodes pourraient s'appuyer sur les travaux effectués dans le cadre des analyses de réseaux d'infections, en comparant par exemple la susceptibilité des bactéries des zones et/ou couches perturbées et naturelles à des pools de virus récupérés dans les différentes zones ou couches comparées (Vos et al., 2009 ; Poisot et al., 2011 ; Koskella et al., 2013).

Références

Koskella, Britt. 2013. "Phage-Mediated Selection on Microbiota of a Long-Lived Host." *Current Biology : CB* 23 (13): 1256–60. doi:10.1016/j.cub.2013.05.038.

Koskella, Britt, and Sean Meaden. 2013. "Understanding Bacteriophage Specificity in Natural Microbial Communities." *Viruses* 5 (3): 806–23. doi:10.3390/v5030806.

Krupovic, Mart, and Patrick Forterre. 2011. "Microviridae Goes Temperate: Microvirus-Related Proviruses Reside in the Genomes of Bacteroidetes." *PloS One* 6 (5): e19893. doi:10.1371/journal.pone.0019893.

Krupovic, Mart, David Prangishvili, Roger W Hendrix, and Dennis H Bamford. 2011. "Genomics of Bacterial and Archaeal Viruses: Dynamics within the Prokaryotic Virosphere." *Microbiology and Molecular Biology Reviews : MMBR* 75 (4): 610–35. doi:10.1128/MMBR.00011-11.

Poisot, Timothée, Gildas Lepennetier, Esteban Martinez, Johan Ramsayer, and Michael E Hochberg. 2011. "Resource Availability Affects the Structure of a Natural Bacteria-Bacteriophage Community." *Biology Letters* 7 (2): 201–4. doi:10.1098/rsbl.2010.0774.

Vos, Michiel, Philip J Birkett, Elizabeth Birch, Robert I Griffiths, and Angus Buckling. 2009. "Local Adaptation of Bacteriophages to Their Bacterial Hosts in Soil." *Science (New York, N.Y.)* 325 (5942): 833. doi:10.1126/science.1174173.

Winter, Christian, Thierry Bouvier, Markus G Weinbauer, and T Frede Thingstad. 2010. "Trade-Offs between Competition and Defense Specialists among Unicellular Planktonic Organisms: The 'Killing the Winner' Hypothesis Revisited." *Microbiology and Molecular Biology Reviews : MMBR* 74 (1): 42–57. http://mmbr.asm.org/content/74/1/42.full.

Annexes

Annexe 1

Supplementary Tables

Dynamics of viral abundance and diversity in a *Sphagnum*-dominated peatland: temporal fluctuations prevail over habitat

Flore Ballaud[1], Alexis Dufresne[1], André-Jean Francez[1], Jonathan Colombet[2], Télesphore Sime-Ngando[2] and Achim Quaiser[1]

[1]Université de Rennes 1, UMR CNRS 6553 ECOBIO, Campus de Beaulieu, 35042 Rennes, France

CNRS, UMR 6553 ECOBIO, Campus de Beaulieu, 35042 Rennes, France

[2] Université Clermont Auvergne, Université Blaise Pascal, BP 10448, F-63000 Clermont-Ferrand, France. CNRS, UMR 6023, Laboratoire Microorganismes: Génome et Environnement (LMGE), 24 avenue des Landais, F-63171 Aubière, France

Correspondence: Achim Quaiser, ECOBIO, Université de Rennes 1, Campus de Beaulieu, 263 avenue du Général Leclerc, 35042 cedex Rennes, France. Building 14A, office 222. Tel.: +33 2 23 23 53 51 37. E-mail: Achim.Quaiser@univ-rennes1.fr

Table S1. Detailed information about peat samples.

		Abundance (VPA, PA)	Physico-chemistry	Metaviromes	Metagenomes
Year	**Month**	N° of samples (fen/bog; surface/depth)	N° of samples (fen/bog)		
2010	May	4			
2011	June	4		vFen_June11, vBog_June11	pFen_S (1, 2, 3), pFen_D (1, 2, 3), pBog_S (1, 2, 3), pBog_D (1, 2, 3)
	Aug_A	4		vFen_Aug11, vBog_Aug11	
	Aug_B	4			
	October	4		vFen_Oct11, vBog_Oct11	
2012	March		2	vFen_Mar12 (A, B, C) vBog_Mar12 (A, B, C)	
	May		2		
	June	4	2		
	July	4	2		
	August		2		
	September	4	2		
	November		2		
	Total samples	32	14	8	4
	Triplicates	96	42	12 (2 samples in triplicate)	12

Surface (S): 5 to 10cm below the active *Sphagnum* layer; Depth (D): 10 to 15cm below the active *Sphagnum* layer.

Table S2. Main characteristics of the 6 fen and 6 bog metagenomes.

	pFen_S1	pFen_S2	pFen_S3	pFen_D1	pFen_D2	pFen_D3	pBog_S1	pBog_S2	pBog_S3	pBog_D1	pBog_D2	pBog_D3
Depth*		**Surface**			**Depth**			**Surface**			**Depth**	
N° of sequences	93 863	204 336	9 963	176 881	11 987	333 485	2 091	19 844	25 841	9 543	3 851	1 542
Sequence average size (bp)	454	455	452	455	452	455	452	452	455	456	454	450
Total nucleotides (Mbp)	42.67	93.04	0.45	80.64	5.43	152.05	0.95	8.99	11.77	4.36	1.75	0.7
% matches in database	54.00%	52.00%	52.00%	52.00%	53.00%	53.00%	60.35%	57.14%	56.13%	53.86%	54.09%	57.20%
****Matches to prokaryotes**	98.49%	99.39%	96.83%	99.01%	94.14%	99.42%	92.27%	94.19%	96.76%	93.53%	96.71%	95.81%
****Matches to viruses**	0.02%	0.01%	0.02%	0.01%	0.01%	0.01%	0.05%	0.01%	0.00%	0.00%	0.00%	0.00%
****Matches to eukaryotes**	1.37%	0.52%	2.99%	0.87%	5.76%	0.46%	7.56%	5.73%	3.19%	6.40%	3.29%	4.19%
N° of shared virome sequences***	67	73	17	78	7	78	1	14	10	3	1	1
% of shared virome sequences	0.07%	0.04%	0.17%	0.04%	0.06%	0.02%	0.05%	0.07%	0.04%	0.03%	0.03%	0.06%

*Surface: 5 to 10 cm below the *Sphagnum* capitula layer; Depth: 10 to 15 cm below the *Sphagnum* capitula layer;
**Relative abundances are given as a proportion of the sequences that matched against NCBI “nr” protein database;
***Shared: sequences that are common between at least a metagenome and a metavirome (Compareads1.2.2, t=4, k=33).

Supplementary Figures

Dynamics of viral abundance and diversity in a *Sphagnum*-dominated peatland: temporal fluctuations prevail over habitat

Flore Ballaud[1], Alexis Dufresne[1], André-Jean Francez[1], Jonathan Colombet[2], Télesphore Sime-Ngando[2] and Achim Quaiser[1]

[1]Université de Rennes 1, UMR CNRS 6553 ECOBIO, Campus de Beaulieu, 35042 Rennes, France

CNRS, UMR 6553 ECOBIO, Campus de Beaulieu, 35042 Rennes, France

[2] Université Clermont Auvergne, Université Blaise Pascal, BP 10448, F-63000 Clermont-Ferrand, France. CNRS, UMR 6023, Laboratoire Microorganismes: Génome et Environnement (LMGE), 24 avenue des Landais, F-63171 Aubière, France

Correspondence: Achim Quaiser, ECOBIO, Université de Rennes 1, Campus de Beaulieu, 263 avenue du Général Leclerc, 35042 cedex Rennes, France. Building 14A, office 222. Tel.: +33 2 23 23 53 51 37. E-mail: Achim.Quaiser@univ-rennes1.fr

Figure S1

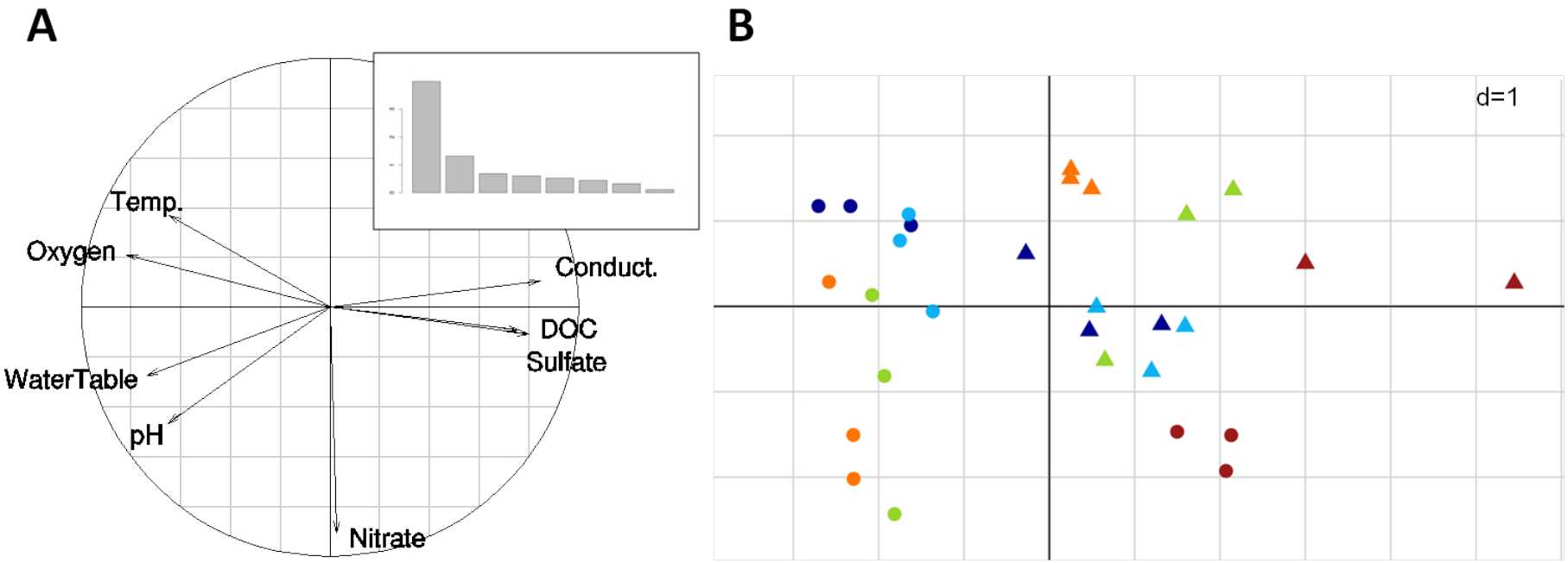


Figure S1. Principal Component Analysis on the physico-chemical dataset. **(A)** Variance explained per component, and correlation circle. The 1st component carries 49.8% of the dataset's variation, and the 2nd component 16.4%. **(B)** Projection of the samples. Dot shape represents the habitat (circle: fen sample, triangle: bog sample). Color represents the sampling date (dark blue: March 2012, light blue: May 2012, light green: June 2012, orange: July 2012, red: September 2012).

Figure S2

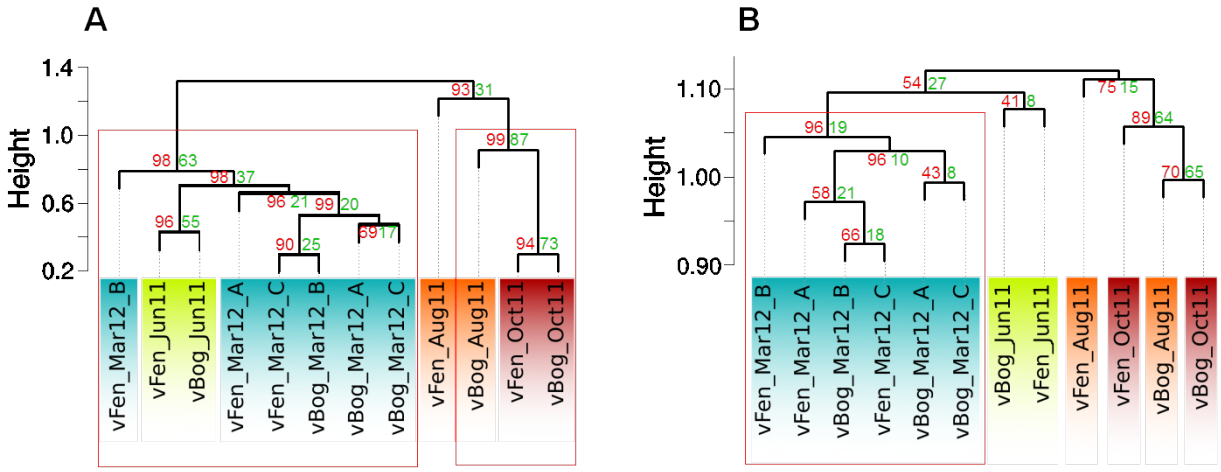


Figure S2. Metavirome similarity based on Compareads analysis **(A)** and Sørensen index dissimilarity **(B)**. Compareads analysis is based on the amount of identical sequences between pairwise samples (k=33, t=4). Sørensen index is based on the presence/absence of sequences from each metaviromes in the different clusters. Red rectangles represent the most robust groups obtained following hierarchical clustering (pvclust, au=0.05).

Figure S3

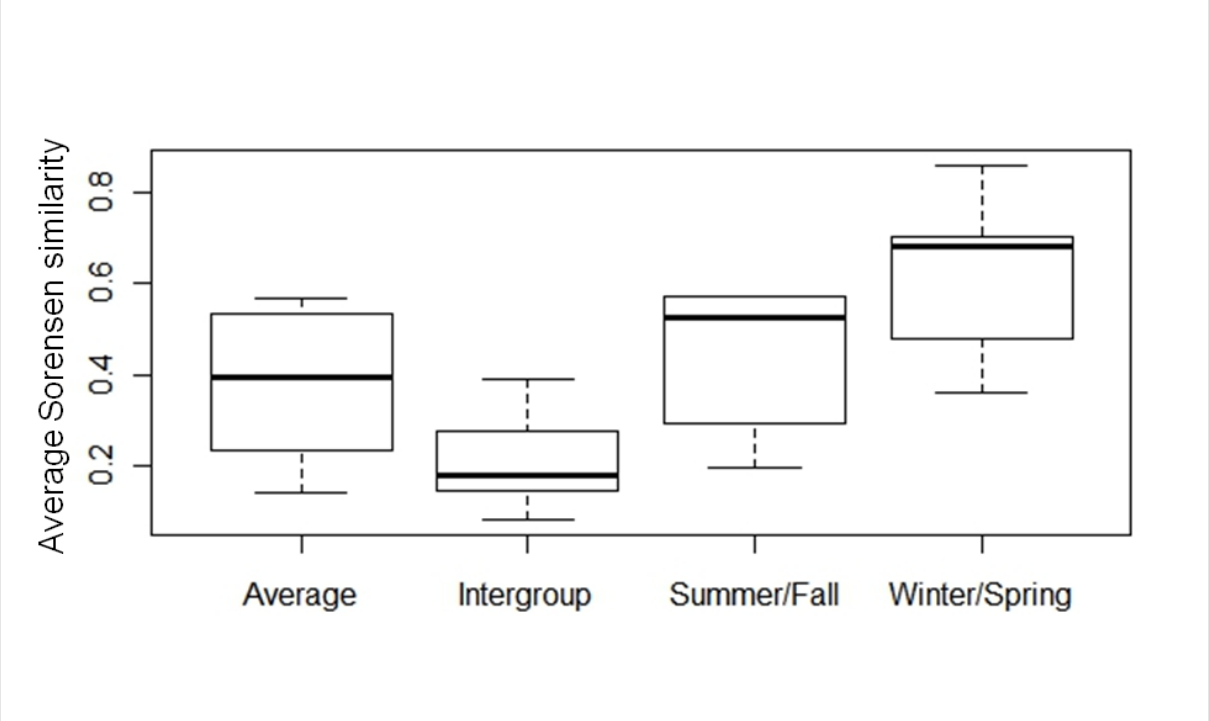


Figure S3. Average Sørensen similarity for the different summer/autumn and winter/spring groups of metaviromes. The groups were determined using the most robust clusters from Figure 5. For each size category, except the last one ($4 \leq CS \leq 3$, only one robust group), average Sørensen similarity (1- Sørensen dissimilarity) was calculated among the metaviromes from each group, and between metaviromes that originated from different groups (Intergroup). Values for each group from each size category were averaged corresponding to the average similarity between all the metaviromes, independently of their assignation to a group.

Figure S4

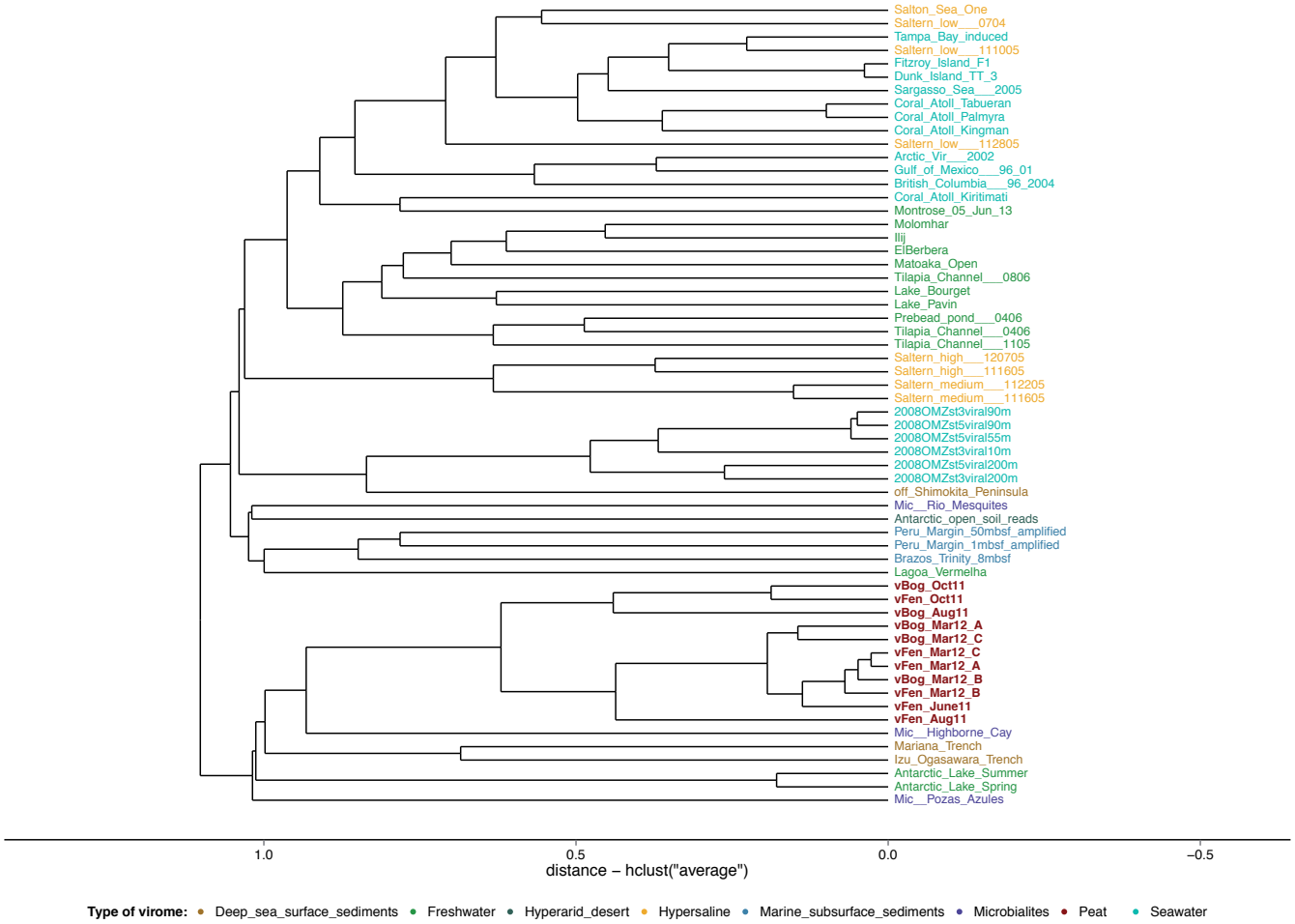


Figure S4. Hierarchical clustering of the peatland metaviromes and metaviromes available for different types of ecosystems (Metavir blast-based comparison). Peatland metaviromes (dark red) form one group. Lake Bourget and lake Pavin are two freshwater lakes (green) geographically close to the peatland. The viral communities from these two lakes strongly differ from the peatland communities. The difference between the communities from the two lakes is equivalent to the difference between the peatland summer/autumn and winter/spring communities.

Figure S5

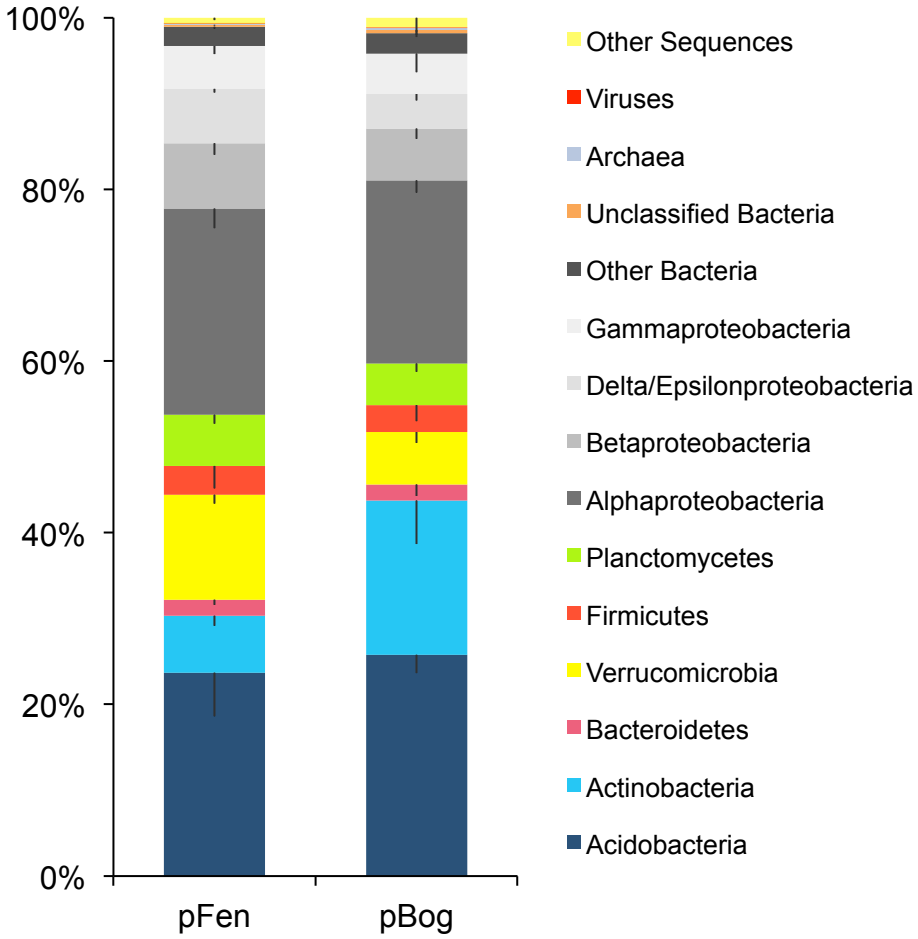


Figure S5. Relative abundance of the prokaryotic phyla in fen and bog microbial metagenomes based on tBLASTx assignations against NCBI non-redundant protein database (relative proportions of assigned sequences). Number of sequences fen: 830515; number of sequences bog: 62712. Error bars represent the negative value of the standard deviation (N=6).

Figure S6

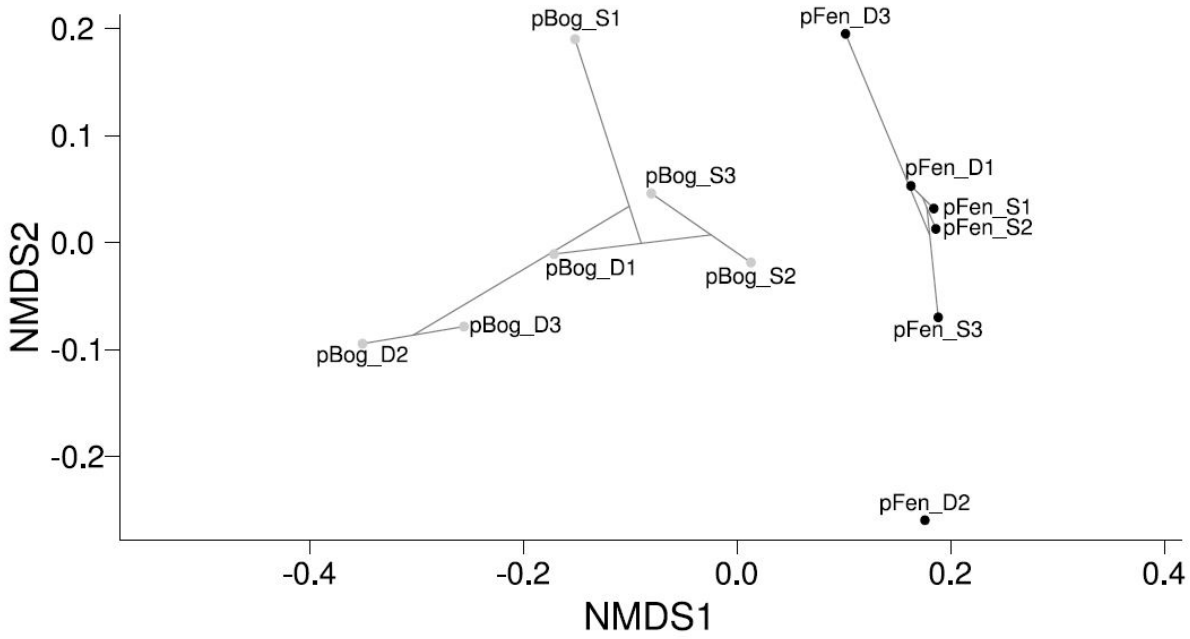


Figure S6. NMDS representing fen and bog microbial communities depending on the euclidean distance between each sample at the phyla level. Stress= 0.03.

Figure S7

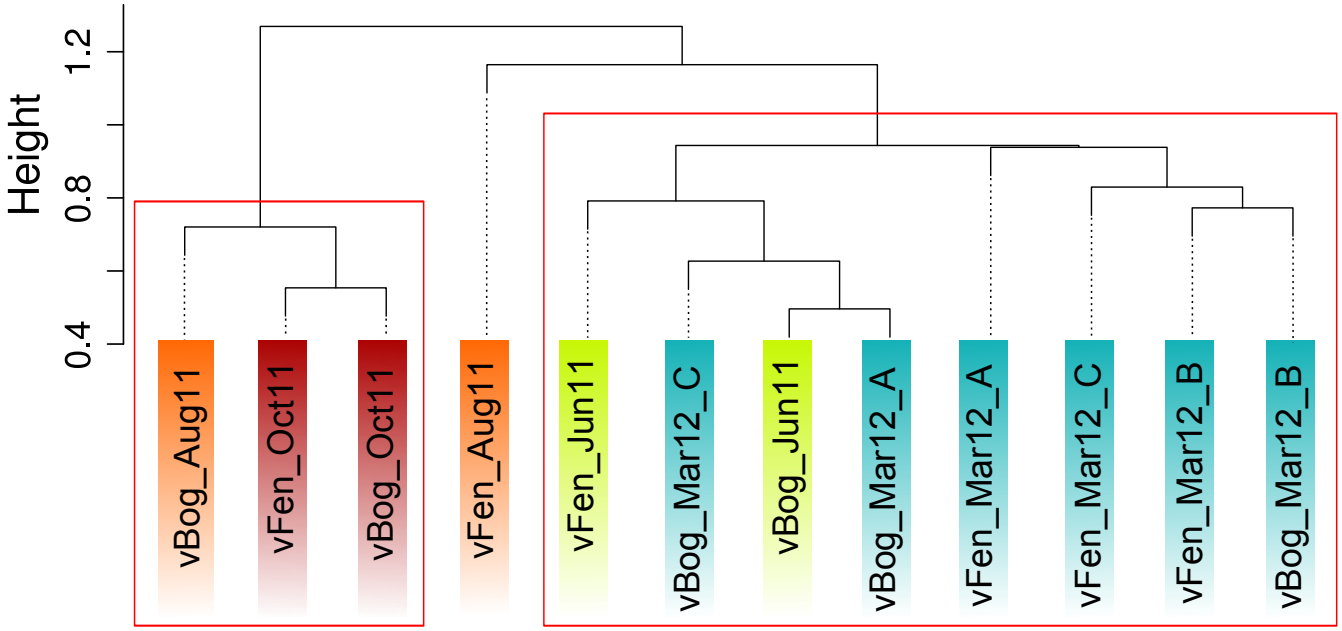


Figure S7. Hierarchical clustering of the metaviromes sequences shared with the metagenomes based on Sørensen dissimilarity (18 326 sequences retrieved with Compareads1.2.2, k=33. t=4). Sequences have been clustered using CD-HIT-EST and Sørensen dissimilarity index was calculated between pairwise metaviromes. Red rectangles represent the most robust clusters (pvclust, au=0.05).

Annexe 2

Supplementary Material

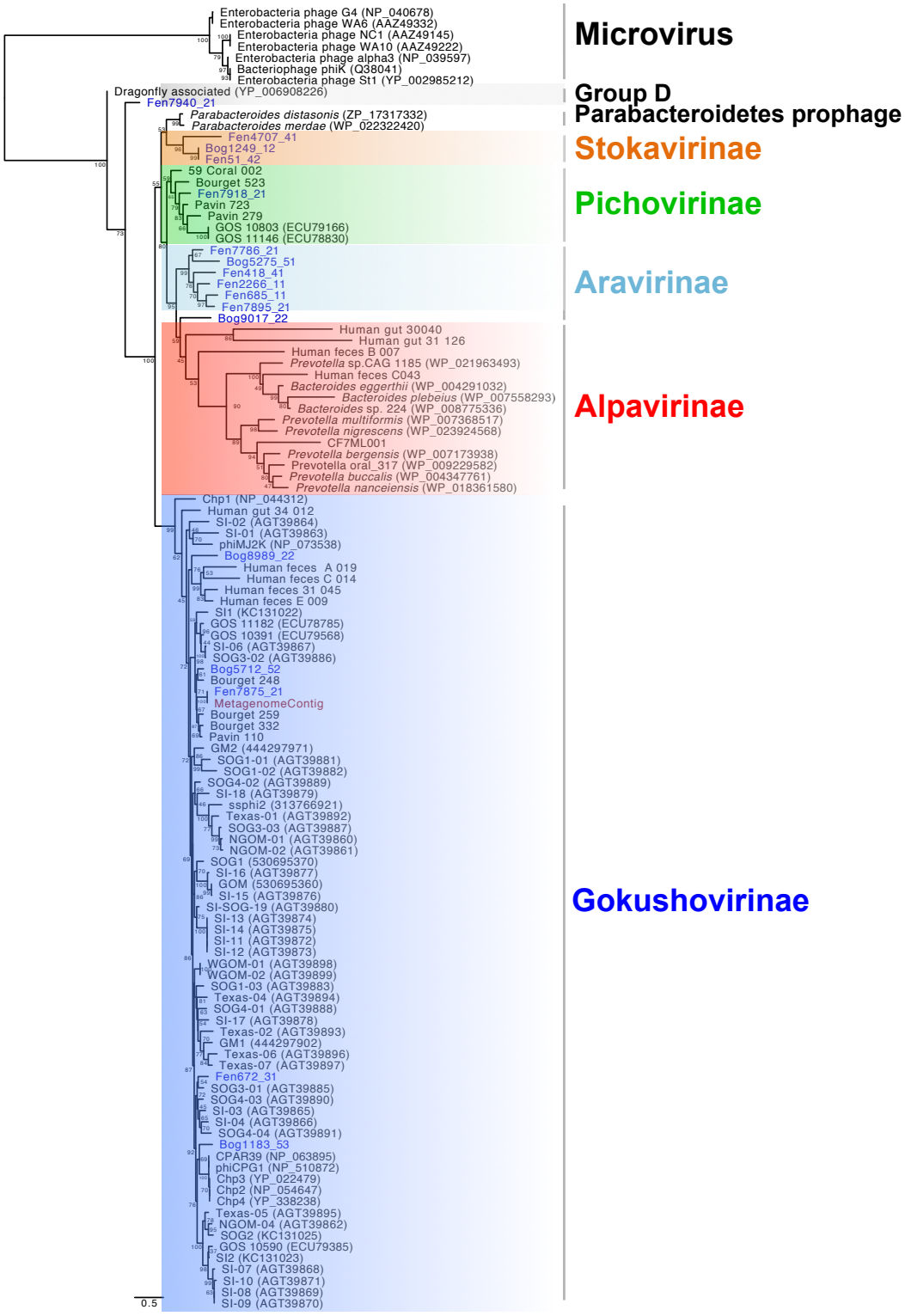
Diversity, distribution and comparative genomics of *Microviridae* in *Sphagnum*-peat soils

Achim Quaiser[1]*, Alexis Dufresne[1], Flore Ballaud[1], Simon Roux[2], Yvan Zivanovic[3], Jonathan Colombet[4], Télesphore Sime-Ngando[4] and André-Jean Francez[1]

[1]Université de Rennes 1, CNRS UMR6553 EcoBio, Rennes, France; [2]Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ, USA; [3]Université Paris Sud, CNRS UMR8621, 91405 Orsay, France. [4]Université Blaise Pascal, UMR CNRS 6023 LMGE, Clermont-Ferrand, France

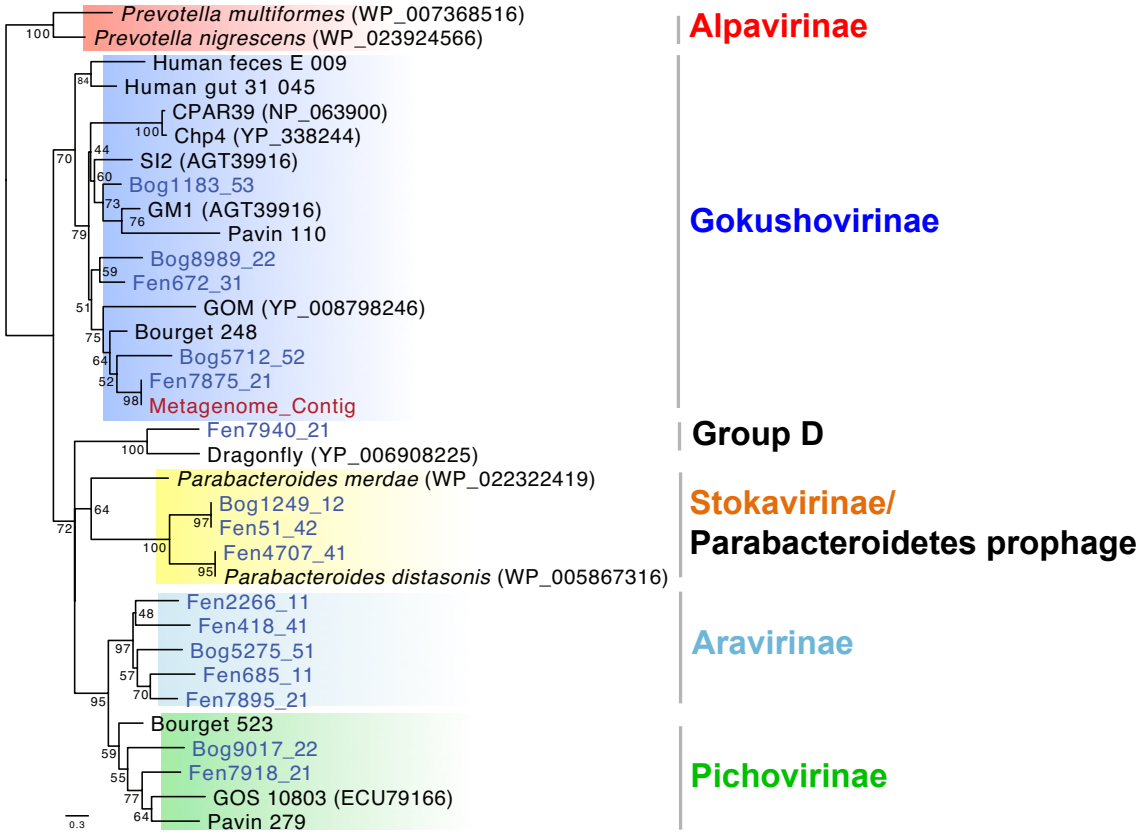
***Correspondence:** Achim Quaiser, Université de Rennes 1, CNRS UMR6553 EcoBio, Rennes, France; achim.quaiser@univ-rennes1.fr

Supplementary Figure S1



Supplementary Figure S1 Maximum likelihood phylogenetic analysis of major capsid protein sequences from *Microviridae* including recently amplified sequences. A total of 157 unambiguously aligned positions from 115 sequences were used in the analysis. Bootstrap values above are indicated at the nodes. The scale bar indicates the number of substitutions per position for a unit branch length.

Supplementary Figure S2



Supplementary Figure S2 Maximum likelihood phylogenetic analysis of replication proteins from *Microviridae*. A total of 144 unambiguously aligned positions from 34 sequences were used in the analysis. The replication protein from Fen7786_21 was too short to align and not included. Bootstrap values above are indicated at the nodes. The scale bar indicates the number of substitutions per position for a unit branch length.

Supplementary Table S1 Characteristics of *Microviridae* genomes assembly

Phage Genome	Original name	Clade	Sampling date	Contig length (bp)	N° of reads_assembled (Metassembler)	Average % identity of assembled reads	Coverage (x times)	Total N° of matches from the same sample (>90% id)	N° of matches from other samples (>90% id)
Fen685_11	11_Contig685	*Aravirinae*	June 2012	4653	54	99.33	4.82	48	20
Fen2266_11	11_Contig2266	*Aravirinae*	June 2012	4366	136	99.53	22.43	98	2
Bog1249_12	12_Contig1249	*Stokavirinae*	June 2012	4453	164	99.4	15.39	150	282
Fen7786_21	21_Contig7786	*Aravirinae*	Aug. 2012	4684	129	99.4	11.46	69	196
Fen7875_21	21_Contig7875	*Gokushovirinae*	Aug. 2012	4630	411	99.31	36.93	337	0
Fen7895_21	21_Contig7895	*Aravirinae*	Aug. 2012	4424	263	99.55	24.73	238	19
Fen7918_21	21_Contig7918	*Pichovirinae*	Aug. 2012	4199	89	99.52	8.82	81	0
Fen7940_21	21_Contig7940	GroupD	Aug. 2012	4453	89	99.61	8.31	89	23
Bog8989_22	22_Contig8989	*Gokushovirinae*	Aug. 2012	4656	117	99.33	10.43	90	35
Bog9017_22	22_Contig9017	*Aravirinae*	Aug. 2012	4671	458	99.44	40.69	396	0
Fen672_31	31_Contig672	*Gokushovirinae*	Oct. 2012	4622	200	99.43	17.70	80	27
Fen418_41	41_Contig418	*Aravirinae*	Mars 2013	4949	147	99.44	12.33	103	1932
Fen4707_41	41_Contig4707	*Stokavirinae*	Mars 2013	4552	215	99.4	19.60	192	242
Fen51_42	42_Contig51	*Stokavirinae*	Mars 2013	4437	300	98.94	28.13	260	170
Bog5275_51	51_Contig5275	*Aravirinae*	Mars 2013	4815	470	99.52	40.61	442	25
Bog5712_52	52_Contig5712	*Gokushovirinae*	Mars 2013	4499	671	99.44	61.75	179	13
Bog1183_53	53_Contig1183	*Gokushovirinae*	Mars 2013	4609	1051	99.4	94.63	357	125
	Metagenome								335
average						99.41	26.99		

Supplementary Table S2 Distance matrix of whole genome phylogeny (Figure 4) based on 2052 unambiguously aligned positions.

	Chp1	Chp3	Chp2	Chp4	CPAR39	phiCPG	SpV4	Bog8989_22	Bog1183_53	Fen672_31	MPSH06248_GM2	SI1	Bog5712_52	Fen7875_21	Metagenome	MPOG05594_GM1	SOG1	phiMH2K	SARssphi2	GOM	SI2	SOG2
Chp1 (Chlamydia, 9629143)	100	59.7	59.6	60.2	59.6	59.8	49	56.1	60.4	55.5	52.6	54.8	54.6	52.9	52.9	50	53.2	48.6	48.4	51.5	55.6	54.3
Chp3 (Chlamydia, 47566140)	59.7	100	97	95.6	95	94.7	50.6	58.5	62.4	59	55.9	57.3	58.6	57.7	57.7	55.9	57.7	53.2	55	55.2	61	59.2
Chp2 (Chlamydia, 9634948)	59.6	97	100	95.6	94.6	94.8	50.5	58.3	62.4	59.3	55.8	57.6	58.7	58	58	55.9	57.7	53.9	55.1	55.4	61.1	59.6
Chp4 (Chlamydia, 77020114)	60.2	95.6	95.6	100	96.1	96.3	50.5	57.9	62.1	59	55.8	57.3	58.3	57.5	57.5	55.8	57.7	53.5	54.3	55.3	60.6	59.8
CPAR39 (Chlamydia, 9791176)	59.6	95	94.6	96.1	100	97.9	50.5	57.7	62.3	59	55.3	57.5	58.9	57.5	57.5	56	57.8	53.4	54.5	55.4	61.4	59.3
phiCPG (Chlamydia, 9632287)	59.8	94.7	94.8	96.3	97.9	100	50.5	57.7	62.6	58.9	55.8	57.5	58.7	57.7	57.7	55.8	57.8	53.2	54.3	55.3	61.4	59.1
SpV4 (Spiroplasma, 19387568)	49	50.6	50.5	50.5	50.5	50.5	100	50.5	53.7	48.2	45.9	48.2	46.8	46.3	46.3	45.7	46.1	44.3	44.1	49.8	49.2	51
Bog8989_22	56.1	58.5	58.3	57.9	57.7	57.7	50.5	100	62.9	59.1	55.9	56.5	57.9	57.7	57.7	51.4	56	50.5	51.7	53.5	56.3	55.2
Bog1183_53	60.4	62.4	62.4	62.1	62.3	62.6	53.7	62.9	100	63.6	58.4	59.6	60.1	59.3	59.3	56.1	57.4	48.2	50.3	53.6	59.5	58.1
Fen672_31	55.5	59	59.3	59	59	58.9	48.2	59.1	63.6	100	58.5	57.8	61.1	62.2	62.2	53.2	59.5	49.8	50.1	52.7	60.1	57
MPSH06248_GM2 (444297971)	52.6	55.9	55.8	55.8	55.3	55.8	45.9	55.9	58.4	58.5	100	57.4	58.6	57.8	57.8	49.7	56.5	48.1	51.9	51.8	55.6	54.6
SI1 (Gokushovirus, 530695349)	54.8	57.3	57.6	57.3	57.5	57.5	48.2	56.5	59.6	57.8	57.4	100	60.4	60.2	60.2	51.1	56.5	50.6	52.5	53.3	58.7	57.4
Bog5712_52	54.6	58.6	58.7	58.3	58.9	58.7	46.8	57.9	60.1	61.1	58.6	60.4	100	64.7	64.7	51.7	57.8	50	51.4	51.6	58.1	55.6
Fen7875_21	52.9	57.7	58	57.5	57.5	57.7	46.3	57.7	59.3	62.2	57.8	60.2	64.7	100	100	51.4	58.7	50	52	52	56.7	54.2
Metagenome	52.9	57.7	58	57.5	57.5	57.7	46.3	57.7	59.3	62.2	57.8	60.2	64.7	100	100	51.4	58.7	50	52	52	56.7	54.2
MPOG05594_GM1 (444297902)	50	55.9	55.9	55.8	56	55.8	45.7	51.4	56.1	53.2	49.7	51.1	51.7	51.4	51.4	100	52.7	49	48.8	50.4	53.6	52.3
SOG1 (Gokushovirus, 530695370)	53.2	57.7	57.7	57.7	57.8	57.8	46.1	56	57.4	59.5	56.5	56.5	57.8	58.7	58.7	52.7	100	52.6	53.6	55.3	58.2	57.3
phiMH2K (Bdellovibrio, 12085135)	48.6	53.2	53.9	53.5	53.4	53.2	44.3	50.5	48.2	49.8	48.1	50.6	50	50	50	49	52.6	100	53.6	51.9	53.2	53.8
SARssphi2 (313766921)	48.4	55	55.1	54.3	54.5	54.3	44.1	51.7	50.3	50.1	51.9	52.5	51.4	52	52	48.8	53.6	53.6	100	53.8	55.6	55.1
GOM Gokushovirus, 530695342)	51.5	55.2	55.4	55.3	55.4	55.3	49.8	53.5	53.6	52.7	51.8	53.3	51.6	52	52	50.4	55.3	51.9	53.8	100	58.3	60
SI2 (Gokushovirus, 530695360)	55.6	61	61.1	60.6	61.4	61.4	49.2	56.3	59.5	60.1	55.6	58.7	58.1	56.7	56.7	53.6	58.2	53.2	55.6	58.3	100	66.5
SOG2 (Gokushovirus, 530695382)	54.3	59.2	59.6	59.8	59.3	59.1	51	55.2	58.1	57	54.6	57.4	55.6	54.2	54.2	52.3	57.3	53.8	55.1	60	66.5	100

Grey shadings indicate different levels of percentage identities. White: < 50%, light grey >50% - 60%; grey > 60%; dark grey > 90%; black = 100%

Supplementary Table S3 Major capsid protein matching sequences (VP1) and general characteristics of viromes analyzed

	Sample name	Availability	Sample origin	Sample type	N° of seqs	Average size (bp)	N° of VP1 matches (Microviridae)	Reference
	vFen June11	Metavir id 1368	Sphagnum dominated peatland	Fen soil	90 933	415	1597	This study
	vBog June11	Metavir id 1369	Sphagnum dominated peatland	Bog soil	39 125	418	936	This study
	vFen Aug11	Metavir id 1373	Sphagnum dominated peatland	Fen soil	106 973	416	1851	This study
	vBog Aug11	Metavir id 1374	Sphagnum dominated peatland	Bog soil	171 231	415	931	This study
	vFen Oct 1	Metavir id 1375	Sphagnum dominated peatland	Fen soil	67 298	409	828	This study
	vBog Oct11	Metavir id 1376	Sphagnum dominated peatland	Bog soil	105 029	413	116	This study
	vFen Mar12_A	Metavir id 1377	Sphagnum dominated peatland	Fen soil	75 147	415	1107	This study
	vFen Mar12_B	Metavir id 1378	Sphagnum dominated peatland	Fen soil	62 351	416	470	This study
	vFen Mar12_C	Metavir id 1379	Sphagnum dominated peatland	Fen soil	78 081	416	1229	This study
	vBog Mar12_A	Metavir id 1380	Sphagnum dominated peatland	Bog soil	114 473	416	1429	This study
	vBog Mar12_B	Metavir id 1382	Sphagnum dominated peatland	Bog soil	122 143	414	5301	This study
	vBog Mar12_C	Metavir id 1370	Sphagnum dominated peatland	Bog soil	65 424	415	1192	This study
1	Lake Pavin	Metavir (id: 6)	Lake Pavin - France	Freshwater	649 290	412	1601	(Roux et al., 2012)
2	Lake Bourget	Metavir (id: 7)	Lake Bourget - France	Freshwater	593 084	433	36879	(Roux et al., 2012)
3	Airborn Forest	Metavir id 1050	Central Region Korea	Air	23 316	544	85	(Whon et al., 2012)
4	Airborn Industrial Komplex	Metavir id 1051	Korea	Air	20 578	539	403	(Whon et al., 2012)
5	Airborn Rainwater	Metavir id 1052	Seoul city - Korea	Air	18 829	542	24	(Whon et al., 2012)
6	Airborn Residential	Metavir id 1053	Seoul city - Korea	Air	24 721	542	71	(Whon et al., 2012)
7	Sediment-Ogasawara	Metavir id 164	Ogasawara Trench - Japan	Pelagic sediment	46 458	336	3058	(Yoshida et al., 2013)
8	Sediment-Mariana	Metavir id 165	Mariana Trench - Japan	Pelagic sediment	49 584	279	173	(Yoshida et al., 2013)
9	Sediment-Shimokita	Metavir id 166	Shimokita Peninsula- Japan	Pelagic sediment	76 498	365	5598	(Yoshida et al., 2013)
10	OMZ st3 viral 10m	Metavir id 897	Chile	Marine OMZ	128 441	246	380	(Cassman et al., 2012)
11	OMZ st3 viral 200m	Metavir id 898	Chile	Marine OMZ	96 706	164	3	(Cassman et al., 2012)
12	OMZ st3 viral 90m	Metavir id 992	Chile	Marine OMZ	361 488	253	34	(Cassman et al., 2012)
13	OMZ st5 viral 200m	Metavir id 993	Chile	Marine OMZ	163 531	242	4	(Cassman et al., 2012)
14	OMZ st5 viral 55m	Metavir id 899	Chile	Marine OMZ	226 628	246	6	(Cassman et al., 2012)
15	OMZ st5 viral 90m	Metavir id 901	Chile	Marine OMZ	300 835	252	9	(Cassman et al., 2012)
16	Human feces A	Metavir id 196	Healthy human	Human feces	113 054	431	22 918	(Kim et al., 2011)
17	Human feces B	Metavir id 293	Healthy human	Human feces	109 569	435	15 619	(Kim et al., 2011)
18	Human feces C	Metavir id 298	Healthy human	Human feces	68 391	437	5 275	(Kim et al., 2011)
19	Human feces D	Metavir id 299	Healthy human	Human feces	115 121	433	20 706	(Kim et al., 2011)
20	Human feces E	Metavir id 300	Healthy human	Human feces	98 511	417	22 458	(Kim et al., 2011)
21	Indian ocean GS108	Metavir id 1478	Coccos Keeling - Australia	Marine plancton (1.8m)	320 104	391	1	(Williamson et al., 2012)
22	Indian ocean GS112	Metavir id 1477	Indian ocean - Australia	Marine plancton (8m)	494 832	361	0	(Williamson et al., 2012)
23	Indian ocean GS117	Metavir id 1479	Saint-Anne Island -Seychelles	Marine plancton (1.8m)	480 375	404	0	(Williamson et al., 2012)
24	Indian ocean GS122	Metavir id 1480	Madagaskar-South Africa	Marine plancton (1.9m)	341 386	350	0	(Williamson et al., 2012)
25	Human salivary Sub1 D1	Metavir id 711	USA	Human salivary	63476	443	8	(Pride et al., 2012)
26	Human salivary Sub1 D30	Metavir id 712	USA	Human salivary	86362	439	0	(Pride et al., 2012)
27	Human salivary Sub2 D30	Metavir id 1102	USA	Human salivary	119 621	406	1 671	(Pride et al., 2012)
28	Human salivary Sub3 D30	Metavir id 1147	USA	Human salivary	103744	287	10	(Pride et al., 2012)
29	Human salivary Sub5 D1	Metavir id 1148	USA	Human salivary	604 957	306	110 355	(Pride et al., 2012)
30	Saanich Inlet (10-200m)	CAMERA	Saanich Inlet - USA	Seawater	96 950	512	269	(Labonté and Suttle, 2013)
31	Gulf of Mexico GOM	CAMERA	Gulf of Mexico - USA	Seawater	87 274	483	149	(Labonté and Suttle, 2013)
32	Coastal water SOG	CAMERA	British Columbia Strait - USA	Seawater	95 402	541	1 184	(Labonté and Suttle, 2013)
33	V1 Typhoon freshwater	SRR371573	Taiwan	Freshwater reservoir	110 339	245	935	(Tseng et al., 2013)
34	V2 Typhoon freshwater	SRR371574	Taiwan	Freshwater reservoir	93 171	247	520	(Tseng et al., 2013)
35	V3 Typhoon freshwater	SRR648311	Taiwan	Freshwater reservoir	75 645	250	37	(Tseng et al., 2013)
36	V4 Typhoon freshwater	SRR648312	Taiwan	Freshwater reservoir	67 123	254	591	(Tseng et al., 2013)
37	V5 Typhoon freshwater	SRR648313	Taiwan	Freshwater reservoir	61 773	254	653	(Tseng et al., 2013)
38	V6 Typhoon freshwater	SRR648314	Taiwan	Freshwater reservoir	59 007	263	11	(Tseng et al., 2013)
39	Reclaimed water Effluent DNA	Metavir id 1351	USA ((FL)	Freshwater	262 097	245	54	(Rosario et al., 2009)
40	Reclaimed water Effluent RNA	Metavir id 1352	USA ((FL)	Freshwater	247 589	227	0	(Rosario et al., 2009)
41	Reclaimed water NurseryDNA	Metavir id 1353	USA ((FL)	Freshwater	283 753	243	3197	(Rosario et al., 2009)
42	Reclaimed water Nursery RNA	Metavir id 1354	USA ((FL)	Freshwater	295 823	236	0	(Rosario et al., 2009)
43	Reclaimed water Park DNA	Metavir id 1355	USA ((FL)	Freshwater	202 436	105	1	(Rosario et al., 2009)
44	Potable Water	Metavir id 1356	USA ((FL)	Freshwater	240 259	226	59	(Rosario et al., 2009)
45	Antarctic Lake Spring	Metavir id 10	Livingston Island, Antarctica	Freshwater	41 322	237	277	(López-Bueno et al., 2009)
46	Antarctic Lake Summer	Metavir id 11	Livingston Island, Antarctica	Freshwater	38 475	221	34	(López-Bueno et al., 2009)
47	Coral_A1	Metavir id 850	US Virgin Islands (Brewers_Bay)	Coral	5 677	385	2	(Soffer et al., 2014)
48	Coral A10	Metavir id 879	US Virgin Islands (Brewers_Bay)	Coral	15 939	404	1352	(Soffer et al., 2014)
49	Coral A2	Metavir id 851	US Virgin Islands (Brewers_Bay)	Coral	3 033	386	303	(Soffer et al., 2014)
50	Coral A3	Metavir id 852	US Virgin Islands (Brewers_Bay)	Coral	4 068	399	128	(Soffer et al., 2014)
51	Coral A5	Metavir id 853	US Virgin Islands (Brewers_Bay)	Coral	24 953	405	5485	(Soffer et al., 2014)
52	Coral A6	Metavir id 854	US Virgin Islands (Brewers_Bay)	Coral	11 924	381	157	(Soffer et al., 2014)
53	Coral A9	Metavir id 855	US Virgin Islands (Brewers_Bay)	Coral	12 942	380	0	(Soffer et al., 2014)
54	Coral D1	Metavir id 880	US Virgin Islands (Brewers_Bay)	Coral	60 629	398	5607	(Soffer et al., 2014)
55	Coral D10	Metavir id 886	US Virgin Islands (Brewers_Bay)	Coral	12 313	395	4073	(Soffer et al., 2014)
56	Coral D2	Metavir id 881	US Virgin Islands (Brewers_Bay)	Coral	14 905	393	214	(Soffer et al., 2014)
57	Coral D3	Metavir id 882	US Virgin Islands (Brewers_Bay)	Coral	25 484	393	3849	(Soffer et al., 2014)
58	Coral D5	Metavir id 883	US Virgin Islands (Brewers_Bay)	Coral	20 186	392	31	(Soffer et al., 2014)
59	Coral D6	Metavir id 884	US Virgin Islands (Brewers_Bay)	Coral	27065	381	1597	(Soffer et al., 2014)
60	Coral D9	Metavir id 885	US Virgin Islands (Brewers_Bay)	Coral	27065	380	1597	(Soffer et al., 2014)
61	Coral H10	Metavir id 892	US Virgin Islands (Brewers_Bay)	Coral	698	400	140	(Soffer et al., 2014)
62	Coral H2	Metavir id 887	US Virgin Islands (Brewers_Bay)	Coral	17786	392	15	(Soffer et al., 2014)
63	Coral H3	Metavir id 888	US Virgin Islands (Brewers_Bay)	Coral	2336	407	0	(Soffer et al., 2014)
64	Coral H5	Metavir id 889	US Virgin Islands (Brewers_Bay)	Coral	9409	392	0	(Soffer et al., 2014)
65	Coral H6	Metavir id 890	US Virgin Islands (Brewers_Bay)	Coral	25037	381	5	(Soffer et al., 2014)
66	Coral_sur. Water d	Metavir id 895	US Virgin Islands (Brewers_Bay)	Coral	22685	407	1330	(Soffer et al., 2014)
67	Coral sur. Water H	Metavir id 896	US Virgin Islands (Brewers_Bay)	Coral	17405	392	1405	(Soffer et al., 2014)
68	Coral Ub1	Metavir id 893	US Virgin Islands (Brewers_Bay)	Coral	1608	381	0	(Soffer et al., 2014)
69	Coral Ub2	Metavir id 894	US Virgin Islands (Brewers_Bay)	Coral	9993	407	1495	(Soffer et al., 2014)

References to Supplementary Table S3

Cassman, N., Prieto-Davó, A., Walsh, K., Silva, G. G. Z., Angly, F., Akhter, S., Barott, K., Busch, J., McDole, T., Haggerty, J. M., et al. (2012). Oxygen minimum zones harbour novel viral communities with low diversity. *Environ Microbiol* 14, 3043–3065. doi:10.1111/j.1462-2920.2012.02891.x.

Kim, M.-S., Park, E.-J., Roh, S. W., and Bae, J.-W. (2011). Diversity and abundance of single-stranded DNA viruses in human feces. *Applied and Environmental Microbiology* 77, 8062–8070. doi:10.1128/AEM.06331-11.

Labonté, J. M., and Suttle, C. A. (2013). Previously unknown and highly divergent ssDNA viruses populate the oceans. *ISME J* 7, 2169–2177. doi:10.1038/ismej.2013.110.

López-Bueno, A., Tamames, J., Velázquez, D., Moya, A., Quesada, A., and Alcamí, A. (2009). High diversity of the viral community from an Antarctic lake. *Science* 326, 858–861. doi:10.1126/science.1179287.

Pride, D. T., Salzman, J., Haynes, M., Rohwer, F., Davis-Long, C., White, R. A., Loomer, P., Armitage, G. C., and Relman, D. A. (2012). Evidence of a robust resident bacteriophage population revealed through analysis of the human salivary virome. *ISME J* 6, 915–926. doi:10.1038/ismej.2011.169.

Rosario, K., Nilsson, C., Lim, Y. W., Ruan, Y., and Breitbart, M. (2009). Metagenomic analysis of viruses in reclaimed water. *Environ Microbiol* 11, 2806–2820. doi:10.1111/j.1462-2920.2009.01964.x.

Roux, S., Enault, F., Robin, A., Ravet, V., Personnic, S., Theil, S., Colombet, J., Sime-Ngando, T., and Debroas, D. (2012). Assessing the diversity and specificity of two freshwater viral communities through metagenomics. *PLoS ONE* 7, e33641. doi:10.1371/journal.pone.0033641.

Soffer, N., Brandt, M. E., Correa, A. M. S., Smith, T. B., and Thurber, R. V. (2014). Potential role of viruses in white plague coral disease. *ISME J* 8, 271–283. doi:10.1038/ismej.2013.137.

Tseng, C.-H., Chiang, P.-W., Shiah, F.-K., Chen, Y.-L., Liou, J.-R., Hsu, T.-C., Maheswararajah, S., Saeed, I., Halgamuge, S., and Tang, S.-L. (2013). Microbial and viral metagenomes of a subtropical freshwater reservoir subject to climatic disturbances. *ISME J* 7, 2374–2386. doi:10.1038/ismej.2013.118.

Whon, T. W., Kim, M.-S., Roh, S. W., Shin, N.-R., Lee, H.-W., and Bae, J.-W. (2012). Metagenomic characterization of airborne viral DNA diversity in the near-surface atmosphere. *J. Virol.* 86, 8221–8231. doi:10.1128/JVI.00293-12.

Williamson, S. J., Allen, L. Z., Lorenzi, H. A., Fadrosh, D. W., Brami, D., Thiagarajan, M., McCrow, J. P., Tovchigrechko, A., Yooseph, S., and Venter, J. C. (2012). Metagenomic exploration of viruses throughout the Indian Ocean. *PLoS ONE* 7, e42047. doi:10.1371/journal.pone.0042047.

Yoshida, M., Takaki, Y., Eitoku, M., Nunoura, T., and Takai, K. (2013). Metagenomic analysis of viral communities in (hado)pelagic sediments. *PLoS ONE* 8, e57271. doi:10.1371/journal.pone.0057271.