

ESTELLE PEDNEAULT

**DISTRIBUTION, EXPRESSION ET DIVERSITÉ DE  
L'AMMONIUM MONOOXYGÉNASE (*AMO*A) DES  
ARCHAEA DANS LES EAUX DU NORD**

Mémoire présenté  
à la Faculté des études supérieures de l'Université Laval  
dans le cadre du programme de maîtrise en biologie  
pour l'obtention du grade de maître ès sciences (M.Sc.)

DÉPARTEMENT DE BIOLOGIE  
FACULTÉ DES SCIENCES ET GÉNIE  
UNIVERSITÉ LAVAL  
QUÉBEC

2009

## Résumé

Les micro-organismes constituent 90 % de la matière vivante océanique. Parmi eux, le *Crenarchaeota Candidatus Nitrosopumilus maritimus* a récemment été isolé, révélant que les membres du même clade sont responsables d'une grande part de la nitrification, dans le cycle de l'azote. Ce potentiel biochimique est détecté chez les *Crenarchaeota* avec une approche moléculaire, avec la présence et l'expression du gène de la sous-unité A de l'ammonium monooxygénase (*amoA*). L'environnement échantillonné, la région des Eaux du Nord (Nord de la Baie de Baffin, entre l'île Ellesmere et le Groenland) a une productivité biologique parmi les plus élevées en Arctique. Celle-ci résulte notamment d'apports externes de nitrate, le produit final de la nitrification. Les résultats obtenus démontrent que le gène *amoA* était abondamment présent et exprimé dans les Eaux du Nord, que sa distribution et son expression étaient influencées par certains paramètres environnementaux et que sa diversité phylogénétique diffèrerait entre les analyses basées sur l'ADN (présence) et l'ADNc (expression).

## Abstract

Microorganisms make up 90 % of organic matter in oceans. Among them, the crenarchaeon *Candidatus Nitrosopumilus maritimus* was recently isolated, revealing that members of its clade could be responsible for much of the nitrification, in the nitrogen cycle. This biochemical potential can be detected in Crenarchaeota using molecular tools, estimating the presence and the expression of the ammonium monooxygenase subunit A gene (*amoA*). The North Water area (Northern Baffin Bay, between Ellesmere island and Greenland) has a biological productivity among highest in the Arctic, which is thought to be linked to nitrate inputs, the final product of nitrification. Results show that the *amoA* gene was abundantly present and expressed, that its distribution and expression were influenced by some environmental variables and that *amoA* phylogenetic diversity differed between DNA (presence) and cDNA (expression) derived analyses.

## Avant-Propos

Ce mémoire de maîtrise a été réalisé sous la supervision de la Professeure Connie Lovejoy. Il se divise en trois chapitres. Le premier consiste en une introduction générale présentant les différents éléments nécessaires à la compréhension du projet. Le deuxième décrit les principaux résultats obtenus. Le troisième présente une conclusion générale du projet.

Le deuxième chapitre a été rédigé en anglais sous forme d'article scientifique, article qui sera prochainement soumis pour publication dans une revue scientifique internationale sous le titre «*Distribution, expression and diversity of Archaeal amoA in the North Water* ». J'ai accompli la majorité des analyses et je suis l'auteure principale de cet article, réalisé en étroite collaboration avec ma directrice de recherche, madame Connie Lovejoy. J'ai par ailleurs contribué à un article publié en 2009 dans la revue *Environmental Microbiology* sous le titre «*Archaeal diversity and a gene for ammonia oxidation are coupled to oceanic circulation* ». Également, certaines données recueillies en mer à bord du NGCC Amundsen et le résultat de leur analyse ont été présentés sous forme d'affiches scientifiques lors de l'*International Conference on Polar and Alpine Microbiology* à Banff en mai 2008 ainsi que lors de l'*International Arctic Change Conference* à Québec en décembre 2008. Enfin, des études ultérieures ont à leur disposition des données complémentaires ayant été l'objet d'analyses préliminaires.

Je tiens à exprimer ma gratitude envers les organismes, institutions et personnes qui ont contribué de près ou de loin à cet accomplissement. L'océanographie biologique est pour moi une passion depuis déjà de nombreuses années. Un rêve s'est concrétisé. J'ai acquis grâce à ce projet de maîtrise une expérience de grande valeur, tant scientifiquement que humainement.

Merci à madame Connie Lovejoy, ma directrice de recherche, dont la passion et la jovialité sont aussi vivaces que contagieux. Merci de m'avoir judicieusement conseillée et



encouragée. Merci à l'appui financier du *Conseil de recherche en sciences naturelles et en génie du Canada* (CRSNG), de *Québec-Océan*, du programme *ArcticNet* et du département de biologie de l'Université Laval. Merci aux membres de mon comité d'encadrement, messieurs Maurice Levasseur et Jean-Éric Tremblay, pour votre professionnalisme et vos conseils. Merci aux membres d'équipage et aux collaborateurs à bord du NGCC Amundsen en septembre 2006. Merci aux collègues membres des laboratoires de madame Connie Lovejoy et monsieur Warwick Vincent, pour leur indispensable coopération à chacune des étapes et toutes les couleurs qu'ils ont apporté dans ma vie. Merci aux membres de ma famille et mes ami(e)s, des gens si précieux. Enfin, je ne peux passer sous silence le fait que le déroulement de ce projet a été dérangé par un événement accidentel. Le plus grand défi de ma vie. Le soutien et le sourire de tous et toutes m'ont été d'une valeur inestimable. Merci infiniment.

*« La vie est comme une bicyclette, si on n'avance pas on perd l'équilibre. »*

- Albert Einstein

*À vous.*

# Table des matières

Résumé .....	i
Abstract .....	ii
Avant-Propos.....	iii
Table des matières .....	vii
Liste des tableaux .....	ix
Liste des figures .....	x
<b>Chapitre 1 - Introduction .....</b>	<b>1</b>
Les Archaea .....	1
Les méthodologies.....	3
Le cycle de l'azote .....	4
Les Eaux du Nord .....	6
Les hypothèses et objectifs.....	8
<b>Chapitre 2 - Distribution, expression and diversity of Archaeal amoA in the North Water .....</b>	<b>9</b>
Résumé .....	9
Abstract.....	10
Introduction.....	11
Materials and methods .....	13
Sampling.....	13
DNA and RNA extractions.....	15
Quantitative PCR.....	16
PCR amplification, cloning and sequencing.....	17
Statistical, phylogenetic and diversity analysis .....	18
Results .....	19
Physicochemical environment.....	19
Quantitative PCR.....	19
Correlations .....	20
Diversity and phylogeny.....	21
Discussion.....	22
Distribution of archaeal <i>amoA</i> and Crenarchaeota from DNA .....	22
Distribution of archaeal <i>amoA</i> from RNA.....	23
Diversity and phylogeny of archaeal <i>amoA</i> DNA and RNA (cDNA) .....	25
Conclusion .....	27
<b>Chapitre 3 - Conclusion .....</b>	<b>28</b>
<b>Tableaux .....</b>	<b>30</b>
<b>Figures .....</b>	<b>33</b>
<b>Bibliographie .....</b>	<b>43</b>
<b>Annexe 1 .....</b>	<b>50</b>



**Annexe 2 ..... 53**  
**Annexe 3 ..... 54**  
**Annexe 4 ..... 55**

# Liste des tableaux

## Chapitre 2

**Table 2.1.** Characteristics of the PCR primers and PCR conditions used in this study. QPCR Standard curves were derived from an *amoA* DNA clone generated with Arch-amoAF/Arch-amoAR for AOA-*amoA*, from 16S rDNA from clone M05a026.02 for marine Crenarchaeota group I (MCGI), and from 16S rDNA from clone M05a026.02 for Archaea (Galand *et al.*, 2009). AL : Amplicon length; bp : base pairs; AT : Annealing temperature; S : Sequencing; Q : QPCR.....30

**Table 2.2.** Spearman correlation coefficients between *amoA* gene (DNA and cDNA) and 20 other bio-physico-chemical parameters of the seawater sampled. Three stations are included. Negative values indicate inverse relationships. \* :  $P > 0.05$ ; \*\*:  $P > 0.01$ ; Number in parenthesis : Number of observations.....31

**Table 2.3.** Diversity of archaeal *amoA* gene (DNA and cDNA) assemblages. Enumeration of OTU and calculations of Chao1 non-parametric richness estimator and Shannon-Wiener diversity index were realized at a similarity level of  $\geq 97$  %.....32

# Liste des figures

## Chapitre 1

**Figure 1.1.** Arbre des trois domaines du vivant (Modifiée de Woese *et al.*, 1990).....33

**Figure 1.2.** Diversité du domaine des Archaea. L'hexagone rouge au centre indique l'enracinement de l'arbre dans le domaine bactérien. Les triangles de couleur pâle sont des branches ayant exclusivement des espèces non cultivées, les triangles de couleur foncée incluent au moins une espèce cultivée. La taille des triangles est proportionnelle au nombre de séquences analysées. Un total de 1344 séquences d'ARNr 16S y sont considérées. Les boîtes mettent en évidence des résultats d'études en génomique. Les lignes pointillées indiquent des positions phylogénétiques incertaines (Schleper *et al.*, 2005).....34

**Figure 1.3.** Cycle de l'azote en milieu aquatique oxygéné, mettant en évidence le segment de la nitrification (Adaptée d'après Karl *et al.*, 2002; Zehr *et al.*, 2003; Arrigo, 2005; Canfield *et al.*, 2005 et Ward, 2005)..... 35

**Figure 1.4.** Carte délimitée par le cercle Arctique et le pôle Nord, mettant en évidence le site d'étude. Le triangle noir indique la localisation de la région des Eaux du Nord, au nord de la Baie de Baffin et entre l'île d'Ellesmere et le Groenland (Modifiée de IBCAO, 2008)..... 36

## Chapitre 2

**Figure 2.1.** Map of the North Water area showing location of sampling stations 109 (76° 15.28 N, 74° 10.39 W), 123 (77° 20.48 N, 74° 38.37 W) and 129 (78° 19.75 N, 074° 0.87 W). Bathymetry is indicated in isobaths and colour shaded. Main currents are illustrated with arrows. AW : Arctic Water; WGC : West Greenland Current (Adapted from Hamilton *et al.*, 2008).....37

**Figure 2.2.** Depth profiles for stations 129, 123 and 109. Column A) Salinity : black line; Dissolved oxygen concentration (DO<sub>2</sub>) : gray line; Column B) Nitrite concentration : black squares; Nitrate concentration : gray diamonds; Column C) Number of crenarchaeal 16S gene copy ml<sup>-1</sup> (16S Cren<sub>DNA</sub>) and Column E) Number of archaeal *amoA* gene copy ml<sup>-1</sup> (*amoA*<sub>DNA</sub>): black circles; Column D) Number of crenarchaeal 16S gene transcript copy ml<sup>-1</sup> (16S Cren<sub>cDNA</sub>) and Column F) Number of archaeal *amoA* gene transcript copy ml<sup>-1</sup> (*amoA*<sub>cDNA</sub>) : no-filled circles. All numbers of copy ml<sup>-1</sup> are on logarithmic scale. Arrows indicate samples for which clone libraries were constructed for diversity analysis in this study.....38

**Figure 2.3.** Rarefaction curves of archaeal *amoA* gene (DNA and cDNA) assemblages, calculated at a similarity level of  $\geq 97\%$ . DNA: black; cDNA: grey; Station 129: dots



(normal and bold); Station 123: dashes (normal and bold); Station 109: lines (normal and bold).....39

**Figure 2.4.** Community structure of archaeal *amoA* gene (DNA and cDNA) in clone libraries from seawater sampled at stations 129, 123 and 109. Sampling depths and number of clones per library are indicated. Caption shows the corresponding color of each OTU. S: station.....40

**Figure 2.5.** Phylogenetic relationships among archaeal *amoA* sequences from the North Water and previously reported sequences. A condensed tree shows relationships between groups (a), details of group A (b) and group B (c). Scale bar represents 10% sequence divergence. Sequences from this study are named following a code in which M stands for the cruise name (MMBOAS : Molecular Microbial Biodiversity of Arctic Seas), 06 stands for the sampling year, d or r stand for DNA or RNA (cDNA), am stands for *amoA*, the following number indicate the sample (26: RNA from station 123 at 200 m; 27: DNA from station 129 at 120 m; 29: DNA from station 129 at 35 m; 43: RNA from station 109 at 110 m; 46: DNA from station 123 at 280 m; 47: DNA from station 123 at 200 m; 74: DNA from station 109 at 400 m; 77: DNA from station 109 at 110 m) and the last number indicate the sequence number in the library. One representative sequence per library is shown and numbers in parenthesis indicate how many clones are represented by the single sequence. The outgroup was the soil fosmid clone 54d9 (AJ627422) and has been removed from the tree for clarity.....41-42



# Chapitre 1 - Introduction

Invisibles à l'œil nu, les micro-organismes dominent la biomasse des océans en raison de leur abondance inégalée. La matière vivante océanique est à 90 % de taille microscopique (Suttle, 2007). Sans compter que les océans couvrent 70 % de la surface de la planète et recueillent 96 % de l'eau sur Terre (Thurman et Burton, 2001). Le monde microbien, dissimulé sous la surface de l'eau, représente donc un acteur majeur dans la mobilisation des ressources vitales que sont le carbone et les nutriments. Cette dominance n'est pas fortuite. D'une part, l'histoire de l'évolution des microbes est considérée presque aussi longue que l'histoire de la vie sur Terre. D'autre part, l'évolution des micro-organismes ayant précédé de plusieurs milliards d'années celle des macro-organismes, les premiers jouent un rôle-clé dans l'écologie et la biogéochimie des océans. De nombreux processus fondamentaux nécessitent la participation des micro-organismes marins (Perry *et al.*, 2004). La récente application de techniques moléculaires a permis le développement d'un nouveau domaine de recherche, l'écologie microbienne moléculaire. Dans ce mémoire, l'écologie et la génétique de certains micro-organismes impliqués dans le cycle de l'azote, précisément les *Crenarchaeota* impliquées dans la nitrification, ont été étudiées à l'aide d'une approche moléculaire. Le présent chapitre introduit les notions sur lesquelles s'appuie la suite du mémoire, abordant successivement le domaine des *Archaea*, les méthodes utilisées, le cycle de l'azote ainsi que le site d'étude nommé les Eaux du Nord. Également, les hypothèses de recherche et les objectifs de l'étude y sont présentés.

## Les Archaea

Les *Archaea* ont été découvertes par Woese et Fox (1977) grâce aux techniques moléculaires et à l'étude des ARN ribosomiaux (ARNr) 16S. Auparavant, les micro-organismes étaient classés en deux domaines, les procaryotes et les eucaryotes. D'un point de vue cellulaire, chez les procaryotes (du grec signifiant « avant noyau ») la structure cellulaire est simple et le matériel nucléaire n'est pas délimité par une membrane, contrairement aux eucaryotes (du grec signifiant « vrai noyau »). Woese *et al.* (1990) ont alors proposé un nouveau système de classification, conservant le domaine des eucaryotes

(*Eucarya*) et scindant le domaine des procaryotes en deux domaines à part entière, les eubactéries (*Bacteria*) et les archaebactéries (*Archaea*) (figure 1.1). Ces derniers se distinguent d'un point de vue physiologique par le type de liaison des lipides membranaires et la composition de la paroi cellulaire (Perry *et al.*, 2004).

La phylogénie des *Archaea* indique une division en deux principaux embranchements, les *Crenarchaeota* et les *Euryarchaeota* (Woese *et al.*, 1990). D'autres embranchements sont évoqués dans la littérature, les *Korarchaeota* (Barns *et al.*, 1996), les *Nanoarchaeota* (Huber *et al.*, 2002) et les *Thaumarchaeota* (Brochier-Armanet *et al.*, 2008), mais sont à ce jour peu connus. Au moment de leur découverte, les *Archaea* étaient considérées comme des organismes extrémophiles. Les *Crenarchaeota* étaient décrites comme des hyperthermophiles alors que les *Euryarchaeota* étaient des méthanogènes anaérobies ou des halophiles. Mais, Fuhrman *et al.* (1992) ainsi que DeLong (1992), à l'aide de techniques d'hybridation, de clonage et de séquençage, ont détecté des *Archaea* dans la colonne d'eau d'un milieu marin aérobie. Depuis, une plus grande diversité d'*Archaea* a été découverte. En plus de leur présence en milieux extrêmes, les *Crenarchaeota* sont considérées omniprésentes et abondantes dans la zone aphotique de l'océan global, alors que les *Euryarchaeota* sont généralement en abondance supérieure dans la zone photique (Herfort *et al.*, 2007; Mincer *et al.*, 2007). Ces distinctions sont sujettes à des variations selon la région et la période de l'année, mais suggèrent néanmoins l'association à différentes niches écologiques (Church *et al.*, 2003).

Les *Crenarchaeota*, principalement abordées dans le cadre de ce mémoire, se détaillent en une arborescence de filières (figure 1.2). Le groupe I.1A regroupe la majorité des *Crenarchaeota* dans les océans, alors que son homologue, le groupe I.1B est surtout défini par des séquences originant des sols (Schleper *et al.*, 2005). La classification des micro-organismes à l'intérieur de ces groupes se détermine différemment de la classification classique des macro-organismes. Cette dernière, organisée de façon hiérarchique, se base généralement au niveau de l'espèce, soit un groupe de populations en mesure de se reproduire par voie sexuée. Or, le terme *espèce* est défini différemment chez les micro-organismes, soit un ensemble de souches qui ont de nombreuses propriétés stables en



commun et qui diffèrent significativement d'autres groupes de souches (Prescott *et al.* 1995). Deux autres systèmes de classification sont possibles, l'un phénotypique et l'autre phylogénétique. Les tests phénotypiques renseignent sur les caractéristiques exprimées des organismes, sur leurs capacités, et peuvent aider à comprendre leur rôle dans l'environnement où ils vivent. Ils sont simples à réaliser, mais présentent le désavantage de nécessiter une culture pure de l'espèce étudiée qui, elle, peut s'avérer très complexe à réaliser. La phylogénie réfère à l'évolution des organismes en se basant sur les séquences génétiques de certaines molécules ayant changé très lentement au cours de l'évolution. Avec une séquence d'ADN connue, il est possible d'adhérer à une définition moins arbitraire d'une espèce microbienne. La technique de l'hybridation ADN/ADN, difficile à réaliser, permet de regrouper dans une même espèce les individus ayant le même pourcentage de guanine plus cytosine et un taux d'au moins 70 % d'hybridation ADN/ADN (Perry *et al.*, 2004). Enfin, la définition adoptée dans cette étude implique la comparaison de séquences du gène codant pour l'ARNr 16S ou 18S. Les isolats sont ainsi rapidement identifiés en regroupant les individus dont la séquence a au moins 97 % de similarité (Wayne *et al.*, 1987).

## **Les méthodologies**

L'ensemble des géotypes présents sur Terre ou dans une région donnée correspond à la biodiversité. La présente étude concerne plutôt la diversité, c'est-à-dire les géotypes présents et « actifs » en un lieu et un moment donnés (Pedrós-Alió, 2006). La diversité d'un échantillon s'évalue en utilisant des méthodes moléculaires. Dans cette étude, les amorces sélectionnées ont été préalablement conçues pour cibler une portion de gène conservée chez les groupes microbiens choisis. D'une part, l'amplification du gène ciblé par la réaction de polymérisation en chaîne (PCR), suivi du clonage, du séquençage et de différentes analyses des séquences permet de décrire la diversité des individus faisant partie du groupe choisi. Dans un échantillon environnemental, la probabilité d'amplification est supérieure pour les gènes représentés en grande abondance que les gènes rares. Cette approche permet la détection de gènes connus ou inconnus faisant partie du groupe ciblé et non la découverte de nouvelles séquences (Pedrós-Alió, 2007). D'autre part, l'amplification d'un gène ciblé

par la PCR quantitative (qPCR), avec émission de fluorescence par le fluorochrome SYBR Green, permet d'estimer l'abondance relative de copies du gène présentes dans l'échantillon.

Les gènes ribosomiaux d'ARNr, relativement très conservés, ont permis la distinction du domaine des *Archaea* et le développement de la phylogénie du monde microbien. Cependant, rares sont les groupes de micro-organismes qui démontrent une corrélation entre phylogénie et physiologie. Certains gènes sont exprimés de façon constitutive (gènes domestiques, ou « housekeeping »), mais les gènes fonctionnels sont surtout inductibles, c'est-à-dire que leur expression est régulée par les besoins de la cellule et les conditions environnementales. Ainsi, l'étude de l'activité de ces gènes chez les micro-organismes non cultivés repose sur l'activité des enzymes, ces protéines qui catalysent les transformations biogéochimiques. L'ARN messager (ARNm), en tant que seul produit de la transcription de l'ADN qui est traduit en protéines, fait figure de matière première pour comprendre les facteurs régulant l'activité microbienne au sein des cycles biogéochimiques (Ward, 2005). Pour ce faire, une rétrotranscription du segment d'ARNm ciblé sert à synthétiser un segment d'ADN complémentaire (ADNc), représentant la partie codante. Ce dernier est utilisé pour étudier l'expression des gènes parce qu'il est plus stable que l'ARN et plus facile à utiliser.

## **Le cycle de l'azote**

À tous les paliers de la chaîne alimentaire, la synthèse de molécules organiques requiert l'incorporation de carbone (C), d'oxygène (O<sub>2</sub>), d'azote (N<sub>2</sub>) ainsi que de moindres quantités de soufre (S) et de phosphore (P). Dans le cycle de l'azote, l'écart de huit électrons entre le composé azoté le plus oxydé et le plus réduit est lié à l'implication de divers types de métabolisme microbien. La substantielle implication de certains procaryotes dans ce cycle sert notamment à subvenir à leurs propres besoins, l'azote étant un constituant essentiel des nucléotides, des acides aminés, des phospholipides et de la paroi cellulaire des micro-organismes. Par ailleurs, leur implication sert à fournir les nutriments azotés à d'autres groupes de l'écosystème, malgré que plusieurs facteurs physiques et



biologiques contrôlent la disponibilité de l'azote sous une forme ou l'autre (Karl *et al.*, 2002; Canfield *et al.*, 2005). Ainsi, les procaryotes en question participent à la régulation de la productivité biologique et, par conséquent, ont un impact sur le cycle du carbone (Yamamoto-Kawai *et al.*, 2006).

Dans les eaux marines de surface, la production biologique est limitée par la disponibilité de l'azote et du phosphore. Les besoins des producteurs primaires marins sont caractérisés par le rapport Redfield, C:N:P de 106:16:1 (Redfield *et al.*, 1963), et il s'avère que ces mêmes éléments chimiques s'y retrouvent dans des proportions moyennes de 1000:15:1. L'utilisation du carbone est de ce fait restreinte à 1/10 de la quantité disponible (Thurman & Burton, 2001). Les processus déterminants pour la réalisation du rapport Redfield sont de deux types, soient les modifications aux apports exogènes de nutriments et les métabolismes microbiens (Arrigo, 2005). Les apports exogènes de nutriments empruntent différentes voies qui sont spécifiques au site d'étude (voir ci-bas), alors que l'analyse de l'activité des *Crenarchaeota* dans une des réactions du cycle de l'azote, l'oxydation de l'ammonium dans le processus de la nitrification, fait l'objet de la présente étude.

La nitrification consiste en une série d'oxydations et par conséquent a lieu en milieu oxygéné. Ce processus nécessite la présence d'ammonium ( $\text{NH}_4^+$ ) en tant que substrat initial (figure 1.3). L'ammonium est produit via l'ammonification, où la majeure partie des composés azotés intégrés aux biomolécules de différents organismes est recyclée sous forme inorganique par la digestion animale ou par la dégradation microbienne. Il a été calculé en 1980 (Suess, cité par Canfield *et al.*, 2005) que 99 % de cette reminéralisation a également lieu dans la partie oxygénée de la colonne d'eau. L'ammonium disponible peut être réintégré dans des biomolécules ou oxydé, via la nitrification, en hydroxylamine ( $\text{NH}_2\text{OH}$ ), en nitrite ( $\text{NO}_2^-$ ) puis en nitrate ( $\text{NO}_3^-$ ) par des procaryotes chemolithoautotrophes. D'un point de vue microbien, la nitrification est divisée en deux étapes. D'abord, l'oxydation de l'ammonium est réalisée par des micro-organismes munis du gène codant pour l'ammonium monooxygénase (*amo*). Les mêmes organismes possèdent généralement le système enzymatique nécessaire à l'oxydation de l'hydroxylamine (Ward, 2005). Ensuite, l'oxydation du nitrite est assurée par d'autres

groupes de micro-organismes. Enfin, le nitrite et le nitrate peuvent être soit assimilés par des micro-organismes ou des algues, soit dénitrifiés pour produire du N<sub>2</sub> (Canfield *et al.*, 2005).

L'oxydation de l'ammonium en tant qu'activité métabolique a d'abord été détectée chez les  $\beta$ - et  $\gamma$ -protéobactéries. Ces bactéries, qui utilisent l'ammonium comme substrat, sont reconnues pour être souvent peu abondantes, c'est-à-dire 0.1 % de la communauté bactérienne (Ward, cité par Kirchmann, 2000). Un tel décalage entre leur abondance et l'importance de l'activité biogéochimique correspondante porte à croire qu'elles ne sont pas les seules à la réaliser. L'isolement du Crenarchaeota *Candidatus Nitrosopumilus maritimus* a amené la distribution phylogénétique des micro-organismes nitrifiants au-delà du domaine des bactéries. La culture (isolement de l'ADN et séquençage du génome) de ce représentant a permis la mise au point d'amorces et de sondes spécifiques et il est maintenant établi que les *Crenarchaeota*, un groupe aussi abondant qu'omniprésent, possèdent le gène codant pour la sous-unité A de l'ammonium monooxygénase (*amoA*) (Schleper *et al.*, 2005; Francis *et al.*, 2005; Wuchter *et al.*, 2006). Dans la mer du Nord, la variante du gène *amoA* des *Archaea* a été détectée de une à deux fois plus que la variante bactérienne (Wuchter *et al.*, 2006). Les récentes évaluations de la présence du gène *amoA* chez les *Crenarchaeota* dans des environnements marins (Francis *et al.*, 2005; Hallam *et al.*, 2006; Wuchter *et al.*, 2006; Mincer *et al.*, 2007; Galand *et al.*, 2009) confèrent un rôle important à ce groupe. Toutefois, l'évaluation de l'expression du gène *amoA* chez ce même groupe demeure à approfondir.

## Les Eaux du Nord

Le site d'étude, les Eaux du Nord, est une région de l'Arctique située au nord de la Baie de Baffin entre les latitudes 76° et 78,5° N et entre l'île d'Ellesmere et le Groenland (figure 1.4). Cette situation polaire implique des variations extrêmes de l'éclairement incident et de la longueur du jour d'une saison à l'autre (ACIA, 2005). Les masses d'eau y conservent une température légèrement supérieure au point de congélation (Ingram *et al.*, 2002). Celles-ci sont également sujettes à la stratification haline parfois forte, causée par les



interactions entre l'eau de mer, la glace et ses dérivés, les eaux de fonte et les eaux sursalées (Thurman et Burton, 2001).

La bathymétrie de cette région est complexe et peut atteindre plus de 600 m de profondeur (Ingram *et al.*, 2002). Un courant froid et particulièrement salé en provenance de l'océan Arctique rejoint cette zone par son extrémité nord, alors qu'un courant plus chaud et moins salé en provenance de l'Atlantique affecte le sud de la région. Ce qui donne lieu à un entrecroisement des masses d'eau (Melling *et al.*, 2001; Bâcle *et al.*, 2002). L'étude de certains groupes microbiens, dont les *Archaea*, (Lovejoy *et al.*, 2002; Hamilton *et al.*, 2008; Galand *et al.*, 2009) a d'ailleurs démontré l'influence de cet entrecroisement sur la distribution de ces groupes par un couplage physique-biotique. Non seulement les *Crenarchaeota* sont en abondance supérieure dans la zone aphotique et les *Euryarchaeota* dans la zone photique, tel que mentionné précédemment, mais en outre Galand *et al.* (2009) ont observé que 70 % des copies de gènes *amoA* détectées se retrouvaient dans l'une des trois masses d'eau de la région, identifiées par les profils de température et salinité. Cette distinction de la masse d'eau appelée *Arctic Basin halocline intrusions* (ABH) démontre l'importance des liens entre les facteurs biophysicochimiques et leur effet potentiel sur les cycles globaux comme celui de l'azote.

La production biologique de la région des Eaux du Nord est particulièrement élevée, comparée à des environnements aussi extrêmes (Tremblay *et al.*, 2002; Stirling, 1980). Deux facteurs sont en cause. D'une part la formation annuelle d'une polynie, une zone libre de glace au milieu de la banquise où le plancton bénéficie de la lumière qui pénètre dès la fin de l'hiver, et d'autre part, des apports exogènes de nutriments, surtout le nitrate. En ce qui concerne les nutriments, l'apport par ruissellement est faible, puisque la région terrestre entourant les Eaux du Nord est caractérisée par un continent gelé, une absence de grande rivière et une mosaïque de glaciers. L'apport par advection, c'est-à-dire les courants, est plus considérable. Cependant les masses d'eau provenant du Nord de l'océan Pacifique et de l'Arctique sont caractérisées par des teneurs faibles en nitrate et élevées en phosphate (Yamamoto-Kawai *et al.*, 2006; Karl *et al.*, 2002). Les besoins en azote des organismes dans la zone photique des Eaux du Nord sont plutôt considérés satisfaits par les nutriments

issus de la dégradation et la reminéralisation en profondeur qui sont transportés vers la surface par les résurgences. Les *Crenarchaeota* munies du gène *amoA* seraient impliquées dans ces apports exogènes de nitrate, mais les détails de ce processus demeurent à approfondir.

## **Les hypothèses et objectifs**

La première hypothèse de recherche est que les *Archaea* des Eaux du Nord expriment le gène *amoA*. L'objectif correspondant à cette hypothèse est de quantifier et évaluer la distribution du gène *amoA* (ADN) et de son expression (ARN). La deuxième hypothèse est que, chez les *Archaea* des Eaux du Nord, le gène *amoA* est diversifié. L'objectif de travail correspondant est d'analyser la phylogénie et la diversité du gène *amoA* (ADN) et de son expression (ARN).



## **Chapitre 2 - Distribution, expression and diversity of Archaeal amoA in the North Water**

### **Résumé**

La productivité des Eaux du Nord (Nord de la Baie de Baffin, entre l'île Ellesmere et le Groenland) est parmi les plus élevées en Arctique et résulte notamment d'apports externes de nitrate. Les Crenarchaeota y sont omniprésentes, mais leur rôle était inconnu avant l'identification du gène de la sous-unité A de l'ammonium monooxygénase (*amoA*) chez *Candidatus Nitrosopumilus maritimus*. Les membres de ce clade sont considérés responsables d'une grande part de la nitrification océanique. Le gène *amoA* est un indicateur de la nitrification par différentes communautés d'Archaea. Les objectifs de cette étude étaient (1) d'estimer la distribution et l'expression du gène *amoA* dans les Eaux du Nord, par la PCR quantitative et (2) d'examiner la diversité phylogénétique de l'ADN et l'ADNc d'*amoA*. Les résultats démontrent une différenciation entre l'ADN et l'ADNc du gène *amoA*, notamment la corrélation entre leur distribution respective et des paramètres environnementaux distincts, ainsi qu'une diversité phylogénétique d'ADNc moindre.

## Abstract

The productivity of the North Water (Northern Baffin Bay between Ellesmere Island and Greenland) is among the highest in the Arctic and is thought to be linked to nitrate input. Crenarchaeota are ubiquitous in arctic waters but their function was unknown prior to the identification of the gene for ammonium monooxygenase subunit A (*amoA*) in the crenarchaeon *Candidatus Nitrosopumilus maritimus*. Members of this clade could be responsible for much of the nitrification in the ocean. *AmoA* can be used to track likely nitrification by archaeal communities. The goals of this study were (1) to estimate the distribution and expression of the *amoA* gene in the North Water using quantitative PCR and (2) to investigate the phylogenetic diversity of *amoA*, from DNA and cDNA. Results show differences between *amoA* gene from DNA and cDNA, with respective distribution correlated to different environmental parameters, and expression related to a phylogenetic subset.

## Introduction

Microbes involved in the nitrogen (N) cycle play a major role in the productivity of the ocean and the global carbon cycle. N is a key constituent of many important biomolecules. Biological productivity of most marine ecosystems is limited by N availability, which in turn is conditioned by microbial processes. In the pelagic ocean, the redox transformations of the N cycle are thought to be vertically separated. In the euphotic zone, phototrophs and heterotrophs take up their respective sources of N, and heterotrophs excrete excess N usually in the form of ammonia, with a fraction as urea or uric acid. Phytoplankton take up and recycle these reduced forms of N, a process referred to as regenerated production (Dugdale and Goring, 1967; Parker and Armbrust, 2005). The process of the remineralisation of organic N by Bacteria and Archaea occurs via the oxidation of ammonia ( $\text{NH}_3$ ) to nitrite ( $\text{NO}_2^-$ ) and nitrate ( $\text{NO}_3^-$ ) (Miller, 2004). Nitrification is thought to be inhibited by light and mostly occurs below the euphotic zone. Nitrate eventually re-enters the euphotic zone via molecular diffusion, upwelling and mixing events providing a source of new N to upper ocean phytoplankton, a process referred to as new production (Dugdale and Goring, 1967).

While historically nitrification was thought to be carried out by a limited number of Bacteria, the contemporary view of nitrification has been completely revised with the recent discovery of archaeal ammonia oxidizers and their abundance in marine systems (Prosser and Nicol, 2008). Confirmation that Marine Group 1 Crenarchaeota (MGI) could grow chemolithoautotrophically by aerobically oxidizing ammonia to nitrite followed the isolation and cultivation of *Candidatus Nitrosopumilus maritimus* (Könneke *et al.*, 2005). A key gene in this process, which codes for the ammonia monooxygenase  $\alpha$ -subunit (*amoA*), was discovered following environmental genomic studies of uncultivated, nonextremophilic Crenarchaeota (Hallam *et al.*, 2006; Wuchter *et al.*, 2006). Once sequences were known, PCR primers were specifically designed for the archaeal *amoA* gene and it was discovered that ammonia-oxidizing Archaea (AOA) are to be widespread in marine water columns and sediments (Francis *et al.*, 2005). In some environments,



including marine systems, AOA are much more abundant than ammonia-oxidizing bacteria (AOB) (Wuchter *et al.*, 2006; Coolen *et al.*, 2007).

Descriptions of the biogeography of the *amoA* gene over large regions and multiple depths is now feasible using quantitative PCR (qPCR). A study by Galand *et al.* (2009) reported a significant variation in AOA among three different water masses in the North Water, in Northern Baffin Bay (75-79°N, 66-80°W). Kalanetra *et al.* (2009) recently compared central Arctic Ocean and coastal Antarctic AOA. Agogué *et al.* (2008) reported vertical and latitudinal gradients of the Archaeal *amoA* gene throughout the North Atlantic, concluding that AOA were more abundant in the meso- and bathypelagic waters in the northern part of the Atlantic coinciding with the more recently formed deep waters. They also found a significant correlation between archaeal *amoA* gene abundance and dark CO<sub>2</sub> fixation. However, the relationship between biogeochemical characteristics of water masses and AOA distribution remains unclear. All earlier studies have used genomic DNA as a template, but the expression of the gene is circumstantial (Ward, 2005). The use of mRNA as a template or direct measurement of the production of different forms of N would provide more solid evidence of the importance of archaeal nitrification activity over different regions and depths.

The North Water, between Ellesmere Island and Greenland, is a region where several water masses form complex density defined interleaved layers (Båcle *et al.*, 2002). Phytoplankton production remains high for at least three months over the Arctic summer and this area has long been considered one of the most productive regions above the Arctic Circle (Stirling, 1980; Lewis *et al.*, 1996; Klein *et al.*, 2002). In other regions of the Arctic, the annual productivity is either limited by nutrient availability because of strong surface stratification due to river inflow and ice melt (Carmack *et al.*, 2007), or is restricted in time and space as it follows a leading ice edge (Wassmann *et al.*, 2006) where nutrients become quickly consumed. In contrast, the high total annual productivity of the North Water is thought to be a combination of a recurrent polynya, an area of open water surrounded by sea ice that allows light to reach phytoplankton in late winter or early spring, and inputs of additional nitrate either from deep waters or by advection. Thus there are persistent high chlorophyll

levels over the summer and into fall (Mei *et al.*, 2002; Tremblay *et al.*, 2002). However, details of how nitrate becomes readily available throughout the summer months are lacking. Galand *et al.* (2009) found that the Marine Crenarchaeote Group I (MGI) and the archaeal *amoA* gene were abundant in the North Water, particularly in Arctic Halocline influenced water masses. The question of whether these were from relic Archaea transported passively or whether they were active nitrifiers has not been addressed.

The first objective of the present study was to determine whether the archaeal *amoA* transcripts (from RNA) were present in the North Water, and their vertical distribution. To this end, three vertical profiles along the center of the North Water were investigated, and water masses with different biophysicochemical characteristics were targeted. Abundance of the archaeal *amoA* gene and the Crenarchaeota 16S rRNA gene was estimated from both genomic DNA and cDNA (from RNA) using qPCR. The second objective was to investigate the diversity of the *amoA* gene, from both DNA and cDNA, by constructing and sequencing clone libraries, enabling biogeographic comparisons with other regions.

## **Materials and methods**

### **Sampling**

Samples were collected from the North Water, Northern Baffin Bay between Ellesmere Island and Greenland, from 11-17 September 2006 on board the research icebreaker CCGS Amundsen. Three stations, 109, 123 and 129 (Figure 2.1), were sampled along a south-north transect. Seawater samples were collected with 12 L Niskin type polyvinyl chloride (PVC) bottles (Ocean Test Equipment) mounted on a SBE carousel rosette system equipped with a SBE-911 Conductivity-Temperature-Depth (CTD) profiler (Seabird). Other instruments on the rosette system recorded fluorescence (Seapoint), transmissivity (WetLabs C-Star Transmissometer), photosynthetically available radiation (PAR; Biospherical Instruments), pH (Seabird), and relative nitrate concentrations (Satlantic MBARI ISUS). Water sample depths were selected on the downward CTD cast, to ensure that they were from representative water masses (Galand *et al.*, 2009). In particular, for this



study we followed the nitrate profile and targeted peak and minimum abundance indicated from the ISUS output. Eight sampling depths were selected for DNA analysis, and six for RNA analysis. The bottles were closed on the upward cast. Seawater was collected from the sampling bottles into clean carboys, which had been rinsed with 10 % HCL acid, followed by three rinses with MilliQ water and three rinses with sample water. DNA and RNA samples were filtered separately, both using a peristaltic pump system. The entire filtration system used to filter RNA was treated with RNaseZap® prior to filtration. Six L of water were filtered at room temperature for DNA analysis, and the volume filtered for RNA varied from 1-7 L with a maximum filtration time of 3 hours at 4 °C. All filtration was begun within 30 min of collection with samples stored at 4 °C prior to filtration. The samples were prefiltered through a 53 µm nylon mesh to remove zooplankton, and filtered sequentially through a 47 mm diameter 3 µm pore size polycarbonate (PC) filter, followed by a 0.2 µm pore size filter. The 3-0.2 µm size fraction was used for this study. For DNA samples, the final filter was a 0.2 µm Sterivex filter unit (Millipore). When filtration finished, stabilizing buffer (1.8 ml of 40 mmol L<sup>-1</sup> EDTA; 50 mmol L<sup>-1</sup> Tris pH = 8.3; 0.75 mol L<sup>-1</sup> sucrose) was added to the units and samples were stored at -80 °C until further analysis. For RNA samples, the final filter was a 47 mm diameter 0.2 µm PC filter. These were placed in 1.8 ml cryovials and covered with RLT buffer (QIAGEN RNeasy) mixed with 1 % mercaptoethanol. The vials were quick-frozen in liquid nitrogen and stored at -80 °C until further analysis.

CTD data were calibrated at Institut national de la recherche scientifique, centre Eau, Terre et Environnement (INRS-ETE), Québec. Oxygen values from the CTD were calibrated using samples taken at discrete depths (Graneli and Graneli, 1991). Nutrients samples were taken from the same cast and depths. Concentrations of nitrite (NO<sub>2</sub><sup>-</sup>), nitrate (NO<sub>3</sub><sup>-</sup>), silicic acid (Si(OH)<sub>4</sub>), and phosphate (PO<sub>4</sub>) were determined on board with a Bran and Luebbe Autoanalyzer III with routine colorimetric methods (Grasshoff, 1976) and a detection limit of 0.05 µmole L<sup>-1</sup>. Total chlorophyll *a* (Chl *a*) samples were taken at each target depth by filtering 500 mL of seawater from the same cast onto Whatman GF/F filters. These were stored at -80 °C until pigments were extracted from the filters in 95 % ethanol at 70 °C for



5 min and concentrations were determined by spectrofluorometry (Cary Eclipse) (Nusch, 1980).

Nanoplankton (ca. 2-20  $\mu\text{m}$ ) concentrations were estimated using epifluorescent microscopy after staining with 4'6-diamino-2-phenylindole (DAPI) (Invitrogen, Burlington, ON, CAN) (Porter and Feig, 1980). Stained preparations were examined using an Olympus 1X71 (Olympus, Markham, ON, CAN) microscope at 1000X under both blue and UV excitation to visualize chlorophyll fluorescence of chloroplasts and DAPI stained nuclear DNA. Small flagellates were classified by size (1-3  $\mu\text{m}$ , 3-6  $\mu\text{m}$ , 6-10  $\mu\text{m}$ ) and diatoms, dinoflagellates and ciliates were enumerated separately. Cell sizes were estimated using a calibrated grid and ocular micrometer. Biovolume values were determined for each category (Hillebrand *et al.*, 1999). Bacterial grazers were considered to belong to the smaller size categories < 6 microns irrespective of chlorophyll content, whereas protist grazers were dinoflagellates, ciliates and other protists > 6 microns. Bacteria and virus concentrations were estimated after staining with SYBR Gold (Molecular Probes) (Noble and Fuhrman, 1998). All microscopic counts were carried out using an Olympus 1X71 microscope at 1000X under Blue excitation.

### **DNA and RNA extractions**

DNA was extracted within 6 months of collection by a salt extraction method adapted from Aljanabi and Martinez (1997). Lysozyme was added to the Sterivex filter unit (final concentration of 1 mg ml<sup>-1</sup>) and incubated at 37 °C for 45 min. Proteinase K (final concentration of 0.2 mg ml<sup>-1</sup>) and sodium dodecyl sulfate (SDS) (10 %) were then added, and the filter was incubated at 55 °C for 60 min. The lysate was recovered from the unit. The filter was rinsed with 1 ml of lysis buffer, incubated at 55 °C for 15 min, and the lysate was recovered. DNA in the pooled lysates was then extracted by addition of NaCl 2.3 M and centrifugation at 1500 g for 15 min. Following ethanol precipitation with 5 ml of cold 70 % ethanol on ice for 10 min, and centrifugation at 18 000 g for 10 min. The pellet was rinsed with 200  $\mu\text{l}$  of cold 70 % ethanol, and centrifuged at 18 000 g for 5 min. Pellets were

then dried before being solubilized in 200  $\mu$ l of TE buffer. Nucleic acid extracts were stored at -80 °C until further processing.

RNA was also extracted within 6 months of collection, using a method adapted from Church *et al.* (2005). The samples were thawed on ice, vortexed and their contents transferred to new microtubes containing glass beads (VWR International). These were placed on a bead beater and agitated for 2 minutes at 7 °C. The tubes were then centrifuged at 8,500 g for 30 s. The supernatants were transferred to clean 1.5 ml microtubes with an equal volume of 70 % ethanol. Samples were purified using QIAGEN RNeasy minicolumns (QIAGEN), following the manufacturer's specifications. Carry-over DNA was eliminated with DNase I by using the QIAGEN on-column DNase I RNA extraction protocol. RNA was eluted off the RNeasy minicolumn with 30  $\mu$ l RNase-free water. RNA extracts were reverse transcribed and PCR amplified (RT-PCR) by using a High capacity cDNA archive kit (Applied Biosystems) as specified by the manufacturer.

DNA and cDNA concentrations were quantified, using PicoGreen (Invitrogen) with all measurements in duplicate. A mix of 50  $\mu$ l of DNA in TE buffer and 50  $\mu$ l of a 200-fold dilution of PicoGreen was incubated for 5 min at 22 °C in the dark. Fluorescence intensity was detected using a TBS-380 fluorometer (Turner Biosystems) concentrations were determined from a standard curve generated using known amounts of Lambda DNA (Invitrogen).

### **Quantitative PCR**

Real time PCR was performed on a Dyad Disciple thermal cycler with Chromo 4 Real-Time Detector (Bio-Rad, Hercules, CA) in 96 well opaque qPCR plates with adhesive seals (Bio-Rad). Primers used in this study are listed in Table 2.1. Reactions were run in triplicate in a 20  $\mu$ l reaction mixtures that consisted of 1  $\mu$ l of template (between 0.2 and 2.2 ng DNA), 0.2  $\mu$ mol L<sup>-1</sup> of primers, MilliQ, Rnase free, autoclaved water, and 10  $\mu$ l of iQ SYBR Green Supermix (Bio-Rad, 100 mmol L<sup>-1</sup> KCl, 40 mmol L<sup>-1</sup> Tris-HCl,



0.4 mmol L<sup>-1</sup> of each dNTP, iTaq DNA polymerase 50 units/mL, 6 mmol L<sup>-1</sup> MgCl<sub>2</sub>, SYBR Green I, 20 nmol L<sup>-1</sup> fluorescein).

The initial denaturing step consisted of 95 °C for 5 min. For quantification of *amoA*, Crenarchaea and Archaea, the qPCR consisted of 45 cycles beginning with 94 °C for 30 s, then 58.5, 61 °C or 63 °C depending on the primers (Table 2.1) for 40 s, and 72 °C for 40 s. The fluorescent signal was read in each cycle at 78 °C for 25 s. Standard curves were generated from DNA obtained by PCR amplification of *amoA* clones (see below). Primer specificity and amplicons size had been previously verified by Galand *et al.* (2009). Quantification standards consisting of 10-fold dilutions ranging from 10<sup>2</sup> to 10<sup>7</sup> copies of target DNA amplicons that had been purified with Qiaquick PCR Purification Kit (Qiagen). The standards were subjected to qPCR along with the samples. *R*<sup>2</sup> values were always > 0.9762. Negative control reactions were run in triplicate in a 20 µl reaction without DNA. All sample reactions were run in triplicate as well.

Copy number per ml was calculated for each value of the triplicate as described in Zhu *et al.* (2005). For each sample, when one result of the triplicate had a threshold cycle (*C*<sub>t</sub>) value very different from the two other results (difference > 1), melting curves were examined and in most cases the outlier values were suspect and removed from further analysis. On several occasions the three results of the triplicates were quite different but with no clear outlier and no obvious problems with the melting curve characteristics, in these cases all three values were kept.

### **PCR amplification, cloning and sequencing**

A portion (target of 635 bp) of the archaeal *amoA* gene was amplified with the Arch-amoAF and Arch-amoAR primers (Francis *et al.*, 2005; Table 2.1). Three separate PCR reactions were run in triplicate in a 25 µl reaction mixtures that consisted of 2, 3 or 4 µl of template, 0.4 µmol L<sup>-1</sup> of primers, 200 µM dNTPs, 0.5 U *Taq* DNA polymerase (New-England Biolab), PCR reaction buffer and MilliQ, RNAase free, autoclaved water. Thermal cycling reaction conditions began with 95 °C for 60 s, then 30 cycles of 95 °C for 30 s,



53 °C for 30 s, and 68 °C for 60 s, followed by 68 °C for 10 min. The size of the resulting amplicons was verified by gel electrophoresis. The three PCR products were pooled, and then purified with Qiaquick PCR Purification Kit (Qiagen), and cloned with TA cloning kit (Invitrogen) according to the manufacturer's instructions. Colonies were transferred to 96 well plates containing Luria-Bertani medium with 7 % glycerol. Inserts were amplified with the vectors' M13 primers. In this case, thermal cycling reaction conditions began with 94 °C for 5 min, then 25 cycles of 94 °C for 1 min, 60 °C for 1 min, and 72 °C for 2 min, followed by 72 °C for 6 min. Resulting amplicons were analysed by restriction fragment length polymorphism (RFLP). PCR products were digested with RsaI (New England Biolab) for 5 h at 37 °C and digested products were run on 3 % low-melting point agarose gel for 3 h at 80 V. One or two clones per pattern, depending on pattern frequency, were sequenced to cover all RFLP patterns. Sequencing (single read) used the vector's T7p universal primer and was performed at Centre hospitalier de l'Université Laval (CHUL), Québec, with an ABI 3730xl system (Applied Biosystems).

The sequences were compared to the NCBI GenBank nr/nt database using NCBI BLAST (Altschul *et al.*, 1990), closest matches were all to uncultured environmental archaeal *amoA* genes. Nucleotide and translated amino acid sequences obtained in this study will be available soon from GenBank database ([www.ncbi.nlm.nih.gov/Genbank/index.html](http://www.ncbi.nlm.nih.gov/Genbank/index.html)).

### **Statistical, phylogenetic and diversity analysis**

One-Sample Kolmogorov-Smirnov non-parametric test with Lilliefors distribution was used to examine data normality. Log transformation was applied when distributions were non-normal. Pearson correlation was applied to parameters that passed the normality test. Spearman correlation to all quantitative parameters (Systat v.11; Systat Software Inc.). Both tests yielded similar results.

The 202 Archaeal *amoA* gene sequences from this study and 148 sequences selected from the GenBank database to represent reference, including sequences of clades derived from Francis *et al.* (2005), were aligned using Clustal X v.1.83 (Thompson *et al.*, 1997), then

manually checked and trimmed using GeneDoc, v. 2.6.0.2 (Nicholas *et al.*, 1997). After trimming and incorporation of additional sequences, 519 nt segments were included for phylogenetic tree construction using the DNAML program from PHYLIP (available at <http://evolution.genetics.washington.edu/phylip.html>). The topology of dendrograms using maximum likelihood and neighbour-joining were similar.

Operational taxonomic units (OTU) defined at a similarity level of  $\geq 97\%$ , Chao1 non-parametric richness estimators, Shannon-Wiener diversity indexes, and rarefaction curves were calculated for each clone library. The Shannon-Wiener diversity index was also calculated with 19 randomly selected clones per library, except for station 129 at 35 m where only 6 clones were available. Analyses were done using DOTUR (Schloss and Handelsman, 2005) through a Jukes-Cantor corrected distance matrix using the DNADIST program from PHYLIP (available at <http://evolution.genetics.washington.edu/phylip.html>).

## Results

### Physicochemical environment

Salinity varied between 31.4 and 34.4, and dissolved oxygen concentrations varied between 5.9 and 8.2 mL L<sup>-1</sup> (Figure 2.2 column A). At station 123, there was a low oxygen region between 200 and 300 m, coinciding with slight increases in salinity (Figure 2.2 column A) and temperature (Annexe 1). Nitrite concentrations ranged from 0.049 and 0.171  $\mu\text{mole L}^{-1}$  (detection limit of 0.05  $\mu\text{mole L}^{-1}$ ), whereas nitrate concentrations ranged from below the detection to 15.34  $\mu\text{mole L}^{-1}$ . The nitrite maxima were ca. 100-200 m, and were generally 100 to 200 m above the nitrate maxima (Figure 2.2 column B).

### Quantitative PCR

QPCR was used to estimate the relative abundance of the Crenarchaeota 16S rRNA gene from DNA (16S DNA Cren; Figure 2.2 column C), and cDNA, indicating active growth (16S cDNA Cren; Figure 2.2 column D). QPCR was also used to estimate the abundance of



the archaeal *amoA* gene from DNA (*amoA* DNA; Figure 2.2 column E), and cDNA, indicating gene expression (*amoA* cDNA; Figure 2.2 column F). The Crenarchaeota 16S DNA ranged from  $10^3$  to  $10^6$  gene copies  $\text{mL}^{-1}$ , while its cDNA ranged from ca.  $10^4$  to  $10^{10}$  copies  $\text{mL}^{-1}$ . The archaeal *amoA* DNA ranged from ca.  $10^2$  to  $10^5$  gene copies  $\text{mL}^{-1}$ , while its cDNA ranged from ca.  $10^3$  to  $10^6$  copies  $\text{mL}^{-1}$ .

The qPCR results from station 129 indicated that *amoA* DNA values closely tracked the Crenarchaeota 16S gene DNA profile, and both peaked at ca. 120 m at the nitrite maxima. The profile for Crenarchaeota 16S cDNA was similar to the nitrite profile, except for a slight decrease at 160 m below the maximum. In contrast the maximum *amoA* cDNA values were higher in the water column at 35 m. At station 123, the maximum of Crenarchaeota 16S DNA and cDNA and archaeal *amoA* DNA coincided with the nitrate maxima at 280 m, at the base of the low oxygen intrusion. The maximum *amoA* cDNA was above this region, but below the nitrite maxima. At station 109, all four qPCR profiles showed peaks near the nitrate maxima at 195 m (Figure 2.2).

## Correlations

Spearman's correlation was used to evaluate the relationships among archaeal *amoA* DNA and cDNA copy number and 19 different biological and physicochemical characteristics of the water sampled. Values of the variables are listed in Annexe 1. The *amoA* DNA gene copy number was positively correlated with depth, density, salinity, and concentrations of nitrate, phosphorus, Archaea 16S gene copy number from DNA, and Crenarchaea 16S DNA and cDNA copy number ( $P > 0.01$ ), and also with concentration of silicic acid ( $P > 0.05$ ). While it was negatively correlated with pH, dissolved oxygen concentration, and bacterial concentration ( $P > 0.01$ ), and also with chlorophyll *a* concentration ( $P > 0.05$ ). *AmoA* DNA gene copy number was not significantly correlated with PAR, temperature, nitrite, grazers biomass, virus concentration nor Crenarchaeota 16S cDNA (Table 2.2). Archaeal *amoA* cDNA copy number was positively correlated with Crenarchaeota 16S cDNA ( $P > 0.01$ ), nitrite, silicate and phosphorus concentrations ( $P > 0.05$ ). While it was negatively correlated with virus concentration ( $P > 0.01$ ), and PAR ( $P > 0.05$ ). Other



biophysicochemical parameters analysed were not significantly correlated with *amoA* cDNA, including *amoA* DNA gene copy number.

### **Diversity and phylogeny**

Rarefaction curves indicated that the two libraries from Stn 109-110 m, DNA and cDNA, and the cDNA library from Stn 123-200 m reached an asymptote indicating good sampling coverage. Whereas four other DNA *amoA* libraries, from Stn 129-120 m, Stn 123-200 and 280 m and Stn 109-400m, reached near asymptote indicating diversity was mostly sampled. The DNA *amoA* library from Stn 129-35 m was undersampled, with rarefaction curves not reaching an asymptote (Figure 2.3). Over the eight libraries, Chao richness estimator ranged from 1 to 14.5, and both the Shannon-Wiener diversity index with all clones and with subsample of 19 clones ranged from 0 to 1.7. The highest richness was within three of the DNA libraries, from Stn 129-35 and 120 m, and Stn 109-400 m (Table 2.3).

A total of 423 clones yielded 19 OTU defined at  $\geq 97\%$  similarity level (Figure 2.4). There were one to six *amoA* OTUs per library (Table 2.3). Two of the OTUs, OTU 3 and OTU 12, were recovered from all of our DNA *amoA* clone libraries. The most common OTU, OTU 19, was recovered from both cDNA *amoA* libraries, and from all DNA *amoA* libraries except from Stn 123-280 m. Other OTUs were rarer, OTU 7 and OTU 16 were both only from Stn 129-120 m, OTU 13 only from Stn 123-280 m, and OTU 15 from Stn 109-400 m. Both abundant and rare OTUs were retrieved from Stn 129-120 m, Stn 123-200 and 280 m, and Stn 109-400 m. Stn 109 at 110 m and Stn 123 at 200 m had 5 OTUs each. At the extremes, all 6 clones from Stn 129-35 m were different OTUs, and only one OTU was recovered among the 46 clones analysed from the Stn 109-110 m cDNA library (Figure 2.4).

Neighbor joining (not shown) and Maximum likelihood phylogenetic dendrograms revealed that the archaeal *amoA* sequences from North Water fell into the previously defined groups A and B but none into group C (Francis *et al.*, 2005). The remaining clones fell into previously unidentified or novel clades. While group A and unidentified clades

include sequences derived from both DNA and cDNA, none of our cDNA *amoA* clones clustered within group B (Figure 2.5). Alignment of amino acid sequences show position of differences between DNA and cDNA sequences (Annexe 3).

## Discussion

Our Quantitative PCR tracking Archaea, Crenarchaeota and the archaeal *amoA* gene in the North Water confirmed earlier results, indicating the importance of Crenarchaeota in this region of the Arctic (Galand *et al.*, 2009). Overall we found that as in most temperate oceans, Crenarchaeota were rare in well illuminated surface waters (DeLong *et al.*, 2006), and generally increased in abundance with depth and there was a positive correlation with Crenarchaeota estimated from 16S rRNA gene abundance from genomic DNA using qPCR and most of the physical and biological variables measured. This ensemble of physical and biological variables reflects the watermass history in the North Water (Hamilton *et al.*, 2008) and our results are consistent with Crenarchaeota abundance being closely linked to parent water masses (Galand *et al.*, 2009).

### Distribution of archaeal *amoA* and Crenarchaeota from DNA

A number of recent studies report that *amoA* gene distribution and abundance are correlated with the distribution and abundance of Crenarchaeota MG I (Wuchter *et al.*, 2006; Mincer *et al.*, 2007; Francis *et al.*, 2005; Herfort *et al.*, 2007; Hallam *et al.*, 2006; Galand *et al.*, 2009; Pouliot *et al.*, 2009). The use of qPCR for quantification of gene abundance in environmental samples is a relatively new technique and although results may not represent absolute abundance, qPCR yields are comparable and generally consistent to other independent techniques (Prosser and Nicol, 2008).

In this study,  $10^2$  to  $10^5$  copies  $\text{mL}^{-1}$  of the archaeal *amoA* gene were found from DNA samples in different water masses of the North Water in September 2006. Galand *et al.* (2009) reported values near our upper range,  $7.5 \times 10^4$  to  $9 \times 10^5$  copies  $\text{mL}^{-1}$ , from samples collected in 2005 in broadly the same Arctic region below the euphotic zone. The lower



values from our 2006 results reflect on average the increased sampling effort in the upper water column, where copy numbers were lower. Estimates from this study were equivalent or higher than those reported elsewhere, for example Mincer *et al.* (2007) retrieved a maximum of  $2 \times 10^4$  copies mL<sup>-1</sup> in the Monterey Bay and Agogue *et al.* (2008) reported  $2.5 \times 10^3$  to  $2.5 \times 10^4$  copies mL<sup>-1</sup> in the subsurface 100-150 m Atlantic Ocean with much lower estimates in the deeper waters.

The average ratio of North Water archaeal *amoA* gene to crenarchaeotal 16S rRNA gene copy number was one-to-two, which was similar to values reported by Agogue *et al.* (2008) in the upper Atlantic and by Wuchter *et al.* (2006) in the North Sea and the Atlantic and Beman *et al.* (2008) in the Gulf of California, while in Monterey Bay the ratio was near one (Mincer *et al.*, 2007). The two fully sequenced marine Crenarchaeota, *Candidatus Nitrosopumilus maritimus* (Könneke *et al.*, 2005) and *Cenarchaeum symbiosum* (Preston *et al.*, 1996; Hallam *et al.*, 2006) contain a single copy of the functional *amoA* gene. If other marine Crenarchaeota also have a single copy of the *amoA* gene then either ca. half of the Crenarchaeota in this Arctic Region may not have an *amoA* gene or some Crenarchaeota may have multiple 16S gene copy numbers (Acinas *et al.*, 2004).

AOA may be associated with accumulation of Nitrite and Nitrate, and *amoA* and 16S Crenarchaeota gene copy numbers estimated from DNA tended to follow the abundance of Nitrate down the water column and was positively correlated with nitrate concentrations. Similarly in Monterey Bay the crenarchaeal *amoA* gene copy numbers followed nitrate concentrations (Mincer *et al.*, 2007). In a North-South transect of the Atlantic Ocean AOA concentrations were greatest at the base of the euphotic zone, and decreased with depth (Agogue *et al.*, 2008), although the base of the photic zone typically coincides with increasing nitrate concentration the nitrate concentrations were not reported.

### **Distribution of archaeal *amoA* from RNA**

The presence of a functional gene does not mean that the gene is active and expression may only occur under specific and possibly rare combinations of environmental conditions



(Prosser and Nicol, 2008; Ward, 2005). The advantage of using RNA (as cDNA) is that only transcription product, mRNA, is quantified. This provides an estimate of microbial activity, in this case ammonia oxidation (Ward, 2005). Ammonia ( $\text{NH}_4^+$ ) acts as an N source, which is taken by phytoplankton or other microorganisms or oxidized into nitrite ( $\text{NO}_2^-$ ). Nitrification provides energy to chemolithoautotroph Archaea and Bacteria (Könneke *et al.*, 2005). Nitrite is the initial measurable product of nitrification, and an indication of ammonia monooxygenase activity. However since nitrite may also be a substrate for phytoplankton and other micro-organisms, low concentrations do not preclude nitrite production or lack of ammonia monooxygenase activity.

Although low levels of *amoA* expression is normally linked to low substrate levels and lack of activity, Ward *et al.* (2005), pointed out that mRNA levels cannot be used to determine the rates of enzymatic transformations in the environment. Since this relationship varies with the gene-enzyme system analysed, direct quantification of the corresponding enzymatic reaction is advisable. As a general rule, high levels of mRNA signal are only measured during transition phases and relatively low mRNA levels are associated with the highest rates of corresponding enzyme activity. That is because after initial induction of gene expression, once enzyme levels are adequate, transcription ceases and mRNA levels subsequently fall. Despite this caveat, quantification of expression may be used to assess the biotic and abiotic factors that may regulate the gene activity.

This is the first study of crenarchaeal *amoA* gene expression in Arctic Seas. Among the few recent studies of archaeal *amoA* expression, only one was carried out in marine waters (Lam *et al.*, 2007), where from  $0-70 \times 10^4$  crenarchaeal *amoA* mRNA copies  $\text{mL}^{-1}$  were found down the water column of the Black Sea. In that study a peak in *amoA* gene abundance within the nitrification zone coincided with the lower oxic and suboxic zones. In the present study peaks of crenarchaeal *amoA* cDNA copies  $\text{mL}^{-1}$  were recorded within a likely nitrification zone, below the surface oxygen maximum zones. Arctic *amoA* gene transcripts ranged from ca.  $10^4$  to  $10^{10}$  copies  $\text{mL}^{-1}$  and depth of cDNA peaks differed from DNA peaks in gene abundance. Maximum *amoA* expression was generally higher in the water column than gene abundance estimated from DNA, above nitrate peaks and in most

cases just below the nitrite peaks, consistent with this being a region of higher substrate concentrations (ammonia) derived from sinking decaying phytoplankton. This result suggests that at least in the North Water the primary nitrite maximum is a result of nitrification and not phytoplankton excretion as indicated by Lomas *et al.* (2006).

However the significant correlations of *amoA* expression and environmental variables suggest that higher nutrients levels, particularly nitrite, silice and phosphorus, are indicative of nitrification and that irradiance (PAR) and high viral abundance indicate lower likelihood of nitrification activity. Additional data using this approach could facilitate the search for optimal conditions for *amoA* expression or inhibitors. *AmoA* transcript abundance was also positively correlated with Crenarchaeota 16S cDNA abundance, but not with *amoA* from DNA or Crenarchaeota from DNA. Indicating very different factors were controlling the abundance of the biomass, estimated from DNA, and the activity, gleaned from cDNA.

### **Diversity and phylogeny of archaeal *amoA* DNA and RNA (cDNA)**

Diversity of archaeal *amoA* in the North Water area was heterogeneous and moderate compared to other studies of Archaea and *amoA* diversity (Francis *et al.*, 2005; Galand *et al.*, 2009). No correlation was found between *amoA* diversity and abundance of gene copies estimated from DNA or cDNA (data not shown), indicating that diversity was independent of biomass. However, rarefaction curves indicated that five of the eight *amoA* libraries were undersampled (Figure 2.3), with a total of 202 sequences analysed.

Sequences from the North Water were phylogenetically analysed along with their closest matches and sequences selected from the major clusters defined by Francis *et al.* (2005). Those included soils and marine derived sequences from Monterey Bay, the Black Sea, the Eastern tropical North Pacific oxygen minimum zone (ETNP OMZ), the Arctic Ocean, the Japan Deep Sea, Antarctic seawater, Elkhorn Slough, San Francisco Bay and sequences from other environments; Arctic meromictic lakes, wastewater and sediments from hydrothermal spring. Francis *et al.* (2005) suggested that there were three water column



clades (A, B and C), which rarely contained sequences from sediments. Sequences from the North Water were phylogenetically close to sequences from a variety of marine environments, and no sequence fell into the sediment/soil clusters. Our *amoA* gene sequences from DNA fell into the previously identified groups A and B and into an unidentified or novel cluster we have designated group D, but not into group C. Whereas those from cDNA fell into groups A and D, but none in groups B and C. Sequences from a same library were dispersed among clusters, indicating that no phylogenetic cluster was associated to a specific library, except for the Stn 109-110 m cDNA library where only one OTU was recovered (Figure 2.5).

Among the total of 19 OTU, four, OTU 3, OTU 11, OTU 12 and OTU 19, were recovered from all but one or two clone libraries, from DNA and cDNA (Figure 2.4). These common OTU are identifiable as part of phylogenetic groups. OTU 3 clustered within group B (Figure 2.5 (c)), close to sequences from Monterey Bay, whereas OTU 11 and OTU 12 within group A (Figure 2.5 (b)), both close to sequences from Arctic Ocean. OTU 19 clustered within a previously unidentified or novel group (Figure 2.5 (a)). In this group, two low oxygen clusters were identified. These include sequences from sediments, aquarium substratum, marine derived Arctic meromictic lakes with low oxygen haloclines, the low oxygen zone of the Black Sea and wastewater. OTU 19 clustered apart from these, and was most similar to sequences from Antarctic seawater and Japan Deep Sea, suggesting a subcluster, D1.

Further sequencing of the archaeal *amoA* gene from diverse environments would be needed to establish whether the environmental clusters above are robust. Environmental or physiological factors that determine archaeal *amoA* sequences expression in the North Water in group A warrants further investigation.



## Conclusion

Analyses of archaeal *amoA* gene profile distribution, expression, correlation with environmental variables, diversity and phylogeny shows that the North Water contains an active population of MGI Crenarchaota at depths that would favour nitrification and that ammonium oxidation by Archaea is important for the Nitrogen cycle of this productive ecosystem. Not only was the *amoA* gene abundant but it was expressed and diverse. *AmoA* expression is positively correlated with nutrients such as nitrite, silice and phosphorus and Crenarchaeal growth (Crenarchaea 16S cDNA), and negatively correlated with irradiance (PAR) and virus abundance. New sequences enriched the phylogenetic analysis of the *amoA* gene. This first study using cDNA suggested that *amoA* genes clustering in group B were not expressed since no sequence from cDNA fell into that group. Understanding the link between diversity and function in the environment is still a major remaining challenge for predicting factors affecting biological production in the global ocean.

## Chapitre 3 – Conclusion

Le résultat de l'analyse du gène *amoA* des communautés d'Archaea, particulièrement des segments d'ADN et d'ADNc des Crenarchaeota, échantillonnées sur trois profils verticaux dans la région des Eaux du Nord en septembre 2006 démontre que la nitrification qui y a lieu est un processus important, contribuant à la relativement forte production biologique de cet environnement marin arctique. En effet, la variante archéenne du gène *amoA* a été abondamment détectée dans les profils échantillonnés. L'estimation de l'expression du gène, via la quantification de l'ADNc, confirme que le gène *amoA* y est actif et que l'oxydation de l'ammonium est catalysée par l'ammonium monooxygénase. Cette expression est par ailleurs positivement corrélée avec certains nutriments, soit la concentration en nitrites, en silice et en phosphore, et la croissance active des Crenarchaeota (nombre de copies d'ADNc du gène 16S des Crenarchaeota). Les résultats de corrélation sont cohérents avec la conception que l'expression du gène *amoA* est plus inhibée avec l'augmentation du rayonnement photosynthétiquement actif (RPA ou PAR, de l'anglais « photosynthetically active radiation ») ainsi qu'avec l'augmentation de l'abondance de virus. L'analyse phylogénétique du gène *amoA*, dont les groupes A, B et C étaient prédéfinis, a été enrichie des séquences provenant des Eaux du Nord. Des séquences d'ADN se sont jointes aux groupes A et B ainsi qu'à un groupe indéfini ou nouveau nommé groupe D, alors que les séquences d'ADNc ne se sont jointes qu'aux groupes B et D. Cette observation suggère une différenciation phylogénétique entre les gènes *amoA* exprimés et ceux non-exprimés.

Une meilleure compréhension des différents processus du cycle de l'azote et des micro-organismes impliqués s'avère nécessaire afin d'obtenir un portrait détaillé de ces mécanismes et de la façon dont ils affectent la production biologique océanique. En ce sens, les analyses subséquentes à ce sujet bénéficieront des données biophysicochimiques analysées dans ce mémoire et des ratios mettant en relation certaines d'entre elles, présentés à l'annexe 1.

La relation entre la diversité génétique du gène *amoA* et sa fonction au sein d'un écosystème marin demeure également à approfondir. Dans cette optique, l'annexe 2 met à disposition la liste des séquences que regroupe chaque OTU présenté dans ce mémoire (figure 2.4). De même, il est possible que le lien entre la diversité génétique et la fonction écosystémique apparaisse dans l'analyse des séquences d'acides aminés du gène *amoA*. Cette dernière hypothèse est suggérée par les résultats présentés à l'annexe 3, qui détaille la position des acides aminés qui varient entre les séquences protéiques d'AmoA, selon leur groupe et selon leur type (ADN et ADNc).

L'effet de l'éclairement sur l'expression du gène *amoA* est le dernier sujet d'étude à approfondir ici suggéré. Les résultats de corrélations du présent mémoire sont cohérents avec l'hypothèse que l'éclairement inhibe l'expression du gène *amoA*. Néanmoins, les résultats de l'expérience complémentaire présentée à l'annexe 4 viennent nuancer la conception actuelle de ce sujet. Ainsi, l'expression du gène *amoA* serait favorisée sous des conditions d'obscurité, mais pas significativement inhibée sous des conditions d'éclairement. Il est donc suggéré d'analyser l'effet de l'éclairement sur l'expression du gène *amoA* d'un plus grand nombre d'échantillons et accompagner cette expérience d'une analyse étroite des nutriments utilisés et produits par la réaction enzymatique.



## Tableaux

**Table 2.1.** Characteristics of the PCR primers and PCR conditions used in this study. QPCR Standard curves were derived from an *amoA* DNA clone generated with Arch-*amoA*F/Arch-*amoA*R for AOA-*amoA*, from 16S rDNA from clone M05a026.02 for marine Crenarchaeota group I (MCGI), and from 16S rDNA from clone M05a026.02 for Archaea (Galand *et al.*, 2009). AL : Amplicon length; bp : base pairs; AT : Annealing temperature; S : Sequencing; Q : QPCR.

Target gene	Use	Primer pair	Primer sequence (5' to 3')	AL (bp)	Reference	AT (°C)
AOA	S	Arch- <i>amoA</i> F	STAATGGTCTGGCTTAGACG	635	Francis <i>et al.</i> (2005)	53
<i>amoA</i>	S	Arch- <i>amoA</i> R	GCGGCCATCCATCTGTATGT			
AOA	Q	AOA- <i>amoA</i> F	CTG AYT GGG CYT GGA CAT C	256	Coolen <i>et al.</i> (2007)	58.5
<i>amoA</i>	Q	AOA- <i>amoA</i> R	TTC TTC TTT GTT GCC CAG TA			
MCGI	Q	MCGI-391f	AAG GTT ART CCG AGT GRT TTC	122	Takai <i>et al.</i> (2004)	61
16S	Q	MCGI-554r	TGA CCA CTT GAG GTG CTG			
Archaea	Q	Parch-519f	CAG CMG CCG CGG TAA	420	Ovreas <i>et al.</i> (1997) Stahl and Amann (1991)	63
16S	Q	Arch-915r	GTG CTC CCC CGC CAA TTC CT			

**Table 2.2.** Spearman correlation coefficients between *amoA* gene (DNA and cDNA) and 20 other biophysico-chemical parameters of the seawater sampled. Three stations are included. Negative values indicate inverse relationships. \* :  $P > 0.05$ ; \*\* :  $P > 0.01$ ; Number in parenthesis : Number of observations.

	<i>amoA</i> DNA		<i>amoA</i> cDNA	
<b>Physic :</b>				
Depth (m)	0.585	** (24)	0.243	(18)
PAR	-0.323	(24)	-0.579	* (18)
Density	0.576	** (24)	0.214	(18)
Potential temperature (°C)	0.277	(24)	0.255	(18)
<b>Chemistry :</b>				
Salinity	0.582	** (24)	0.243	(18)
pH	-0.616	** (24)	-0.199	(18)
Dissolved oxygen (mL L <sup>-1</sup> )	-0.683	** (24)	-0.234	(18)
Nitrite (µM)	0.246	(24)	0.568	* (18)
Nitrate (µM)	0.724	** (24)	0.207	(18)
Silice (µM)	0.568	* (16)	0.685	* (12)
Phosphorus (µM)	0.679	** (16)	0.692	* (12)
<b>Biology :</b>				
Large grazers (pgC L <sup>-1</sup> )	0.237	(24)	0.164	(18)
Bacteria grazers (pgC L <sup>-1</sup> )	-0.376	(24)	-0.162	(18)
Virus (10 <sup>9</sup> VLP mL <sup>-1</sup> )	-0.230	(24)	-0.671	** (18)
Chlorophyll <i>a</i> (µg L <sup>-1</sup> )	-0.441	* (24)	0.007	(18)
Bacteria (cells L <sup>-1</sup> )	-0.570	** (24)	-0.121	(18)
Archaea 16S (copies mL <sup>-1</sup> )	0.944	** (24)	0.156	(18)
Crenarchaea 16S DNA (copies mL <sup>-1</sup> )	0.848	** (24)	0.108	(18)
Crenarchaea 16S cDNA (copies mL <sup>-1</sup> )	0.695	** (24)	0.538	* (18)
<b><i>amoA</i> related :</b>				
<i>amoA</i> DNA (copies mL <sup>-1</sup> )	X		0.154	(18)
<i>amoA</i> cDNA (copies mL <sup>-1</sup> )			X	

**Table 2.3.** Diversity of archaeal *amoA* gene (DNA and cDNA) assemblages. Enumeration of OTU and calculations of Chao1 non-parametric richness estimator and Shannon-Wiener diversity indexes were realized at a similarity level of  $\geq 97\%$ . na : non applicable.

Station	Depth (m)	Template	No. of clones	No. of OTU	Chao1	Shannon-Wiener	Shannon-Wiener (subsample of 19 clones)
129	35	DNA	6	6	14.5	1.7	na
129	120	DNA	91	12	13.0	2.0	1.5
123	200	DNA	51	8	8.3	1.6	1.4
123	200	cDNA	32	5	5.0	1.5	1.3
123	280	DNA	86	11	11.3	2.0	0.9
109	110	DNA	19	4	4.0	1.3	1.2
109	110	cDNA	46	1	1.0	0	0.0
109	400	DNA	92	13	13.3	2.2	1.7



# Figures

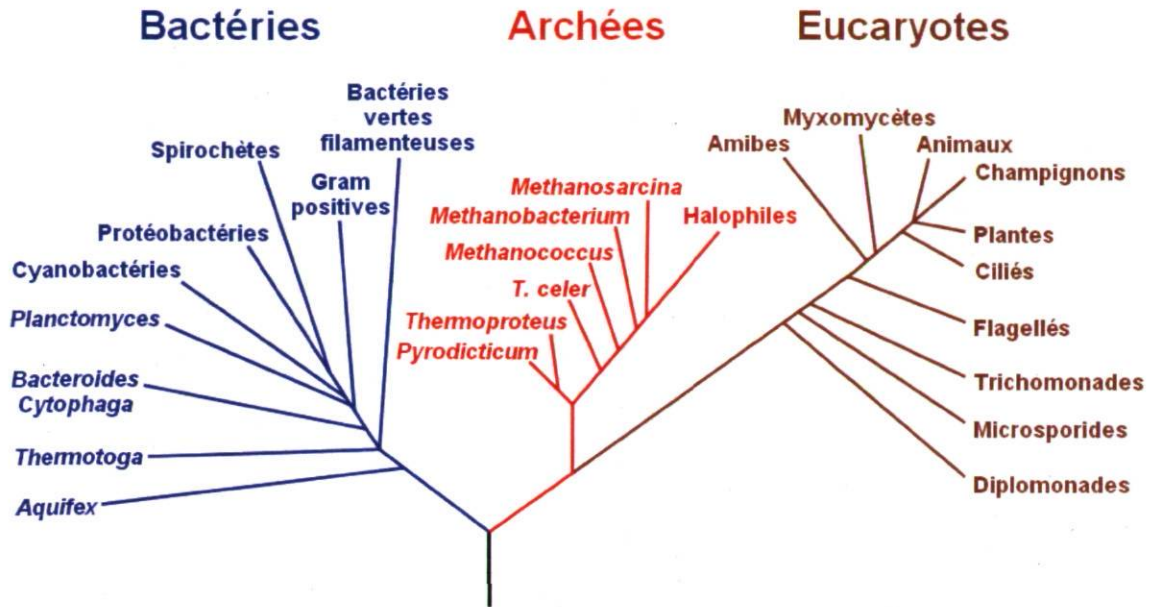
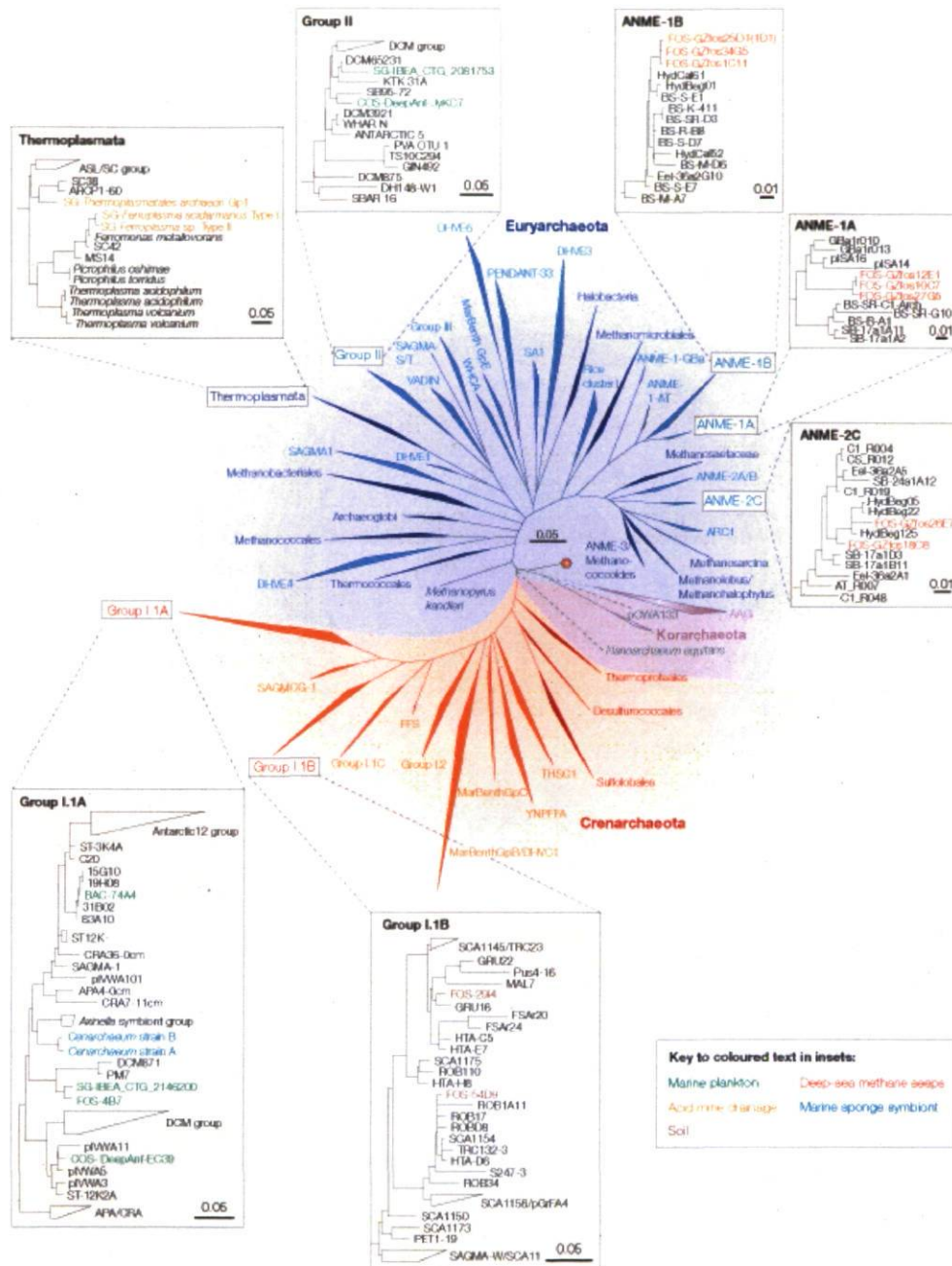
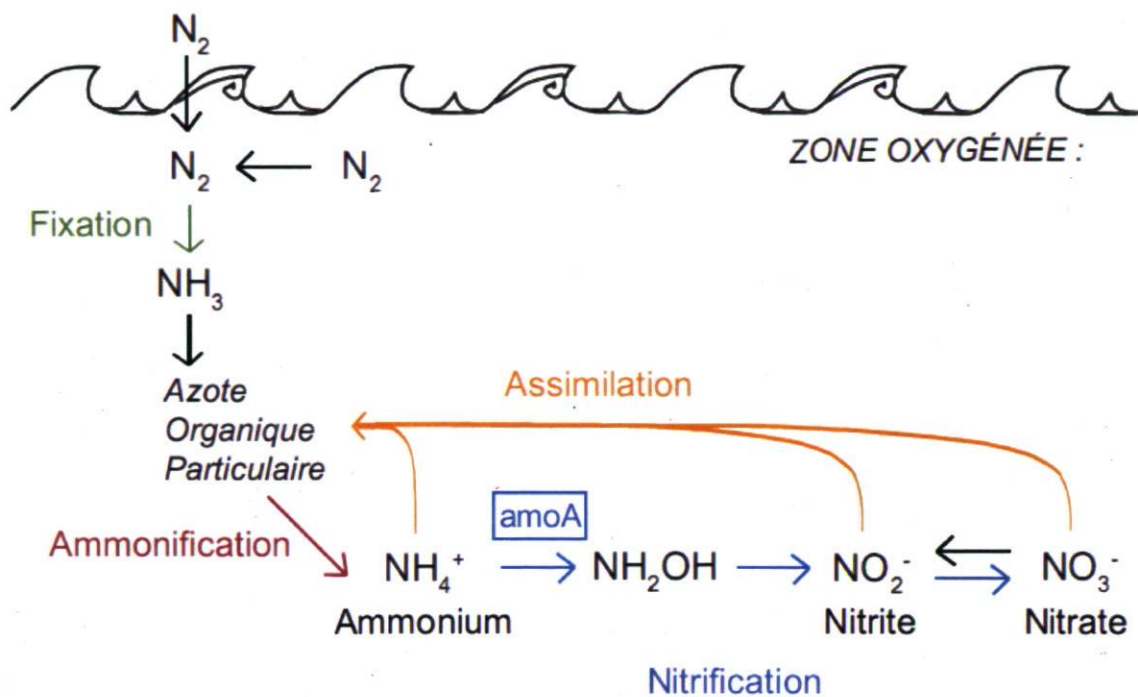


Figure 1.1. Arbre des trois domaines du vivant (Modifiée de Woese *et al.*, 1990).

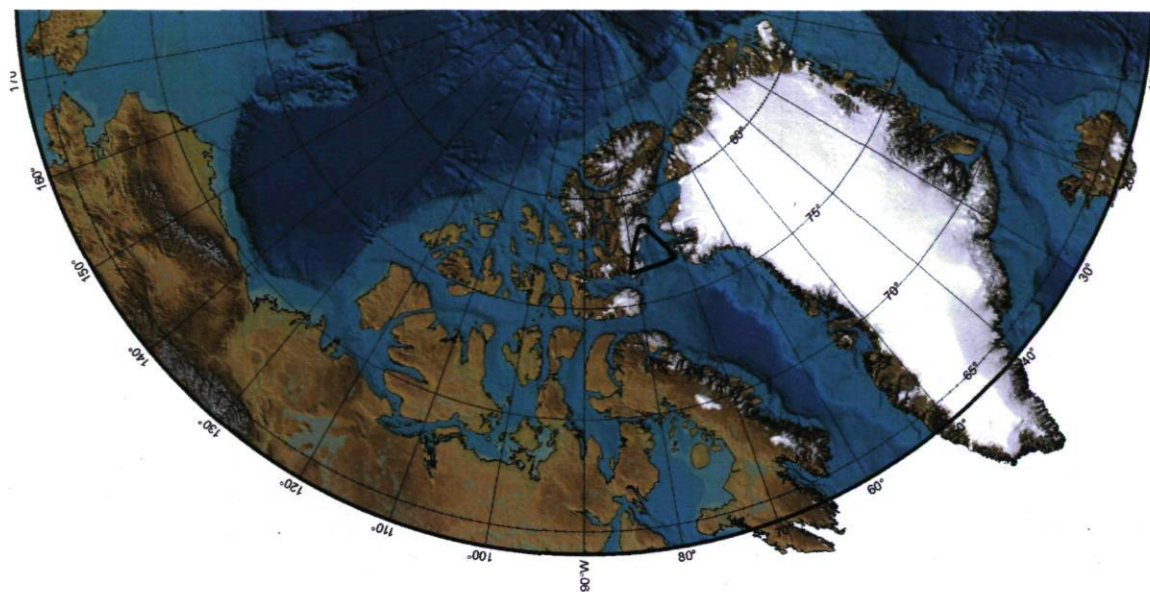


**Figure 1.2.** Diversité du domaine des Archaea. L'arbre est enraciné dans le domaine bactérien. Les triangles de couleur pâle sont des branches ayant exclusivement des espèces non cultivées, ceux de couleur foncée incluent au moins une espèce cultivée. La taille des triangles est proportionnelle au nombre de séquences analysées. Un total de 1344 séquences d'ARNr 16S y sont considérées. Les boîtes mettent en évidence des résultats d'études en génomique. Les lignes pointillées indiquent des positions phylogénétiques incertaines (Schleper *et al.*, 2005).

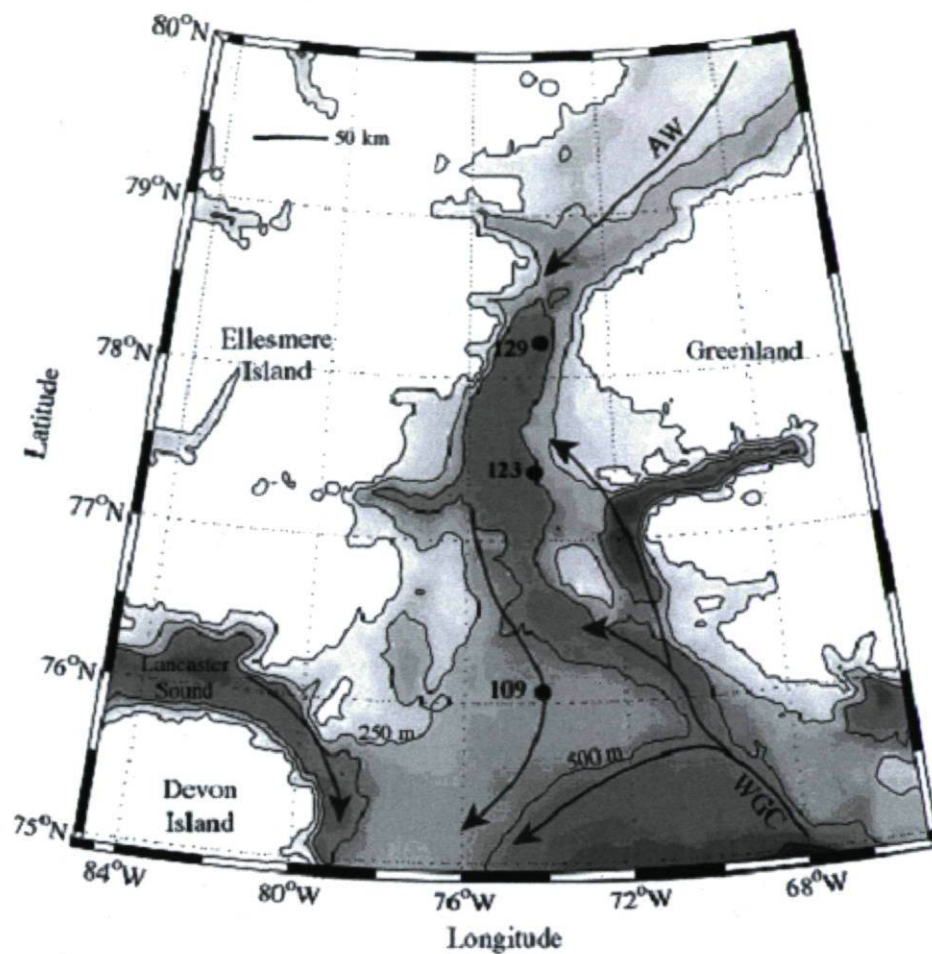


**Figure 1.3.** Cycle de l'azote en milieu aquatique oxygéné, mettant en évidence le segment de la nitrification (Adaptée d'après Karl *et al.*, 2002; Zehr *et al.*, 2003; Arrigo, 2005; Canfield *et al.*, 2005 et Ward, 2005).

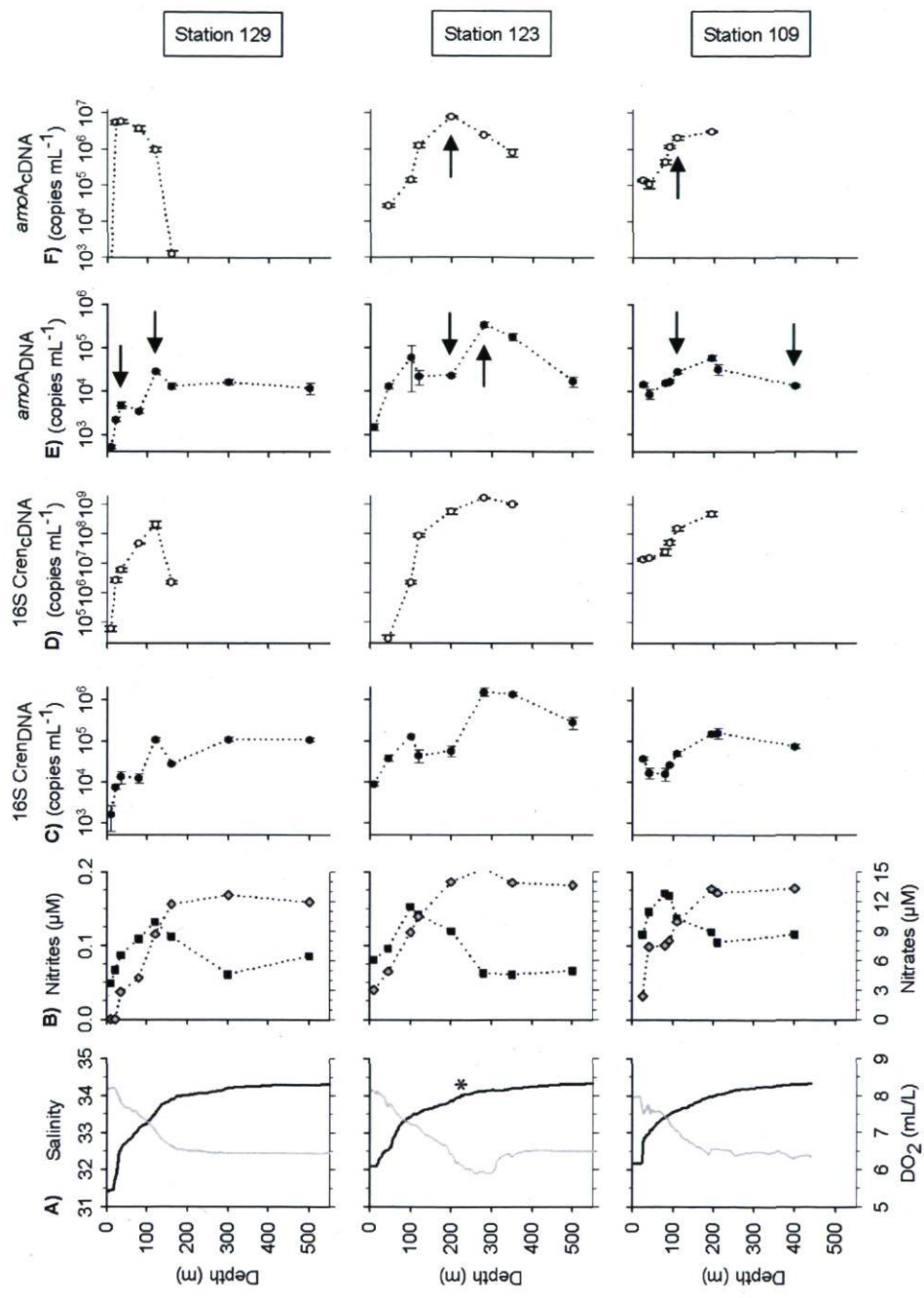




**Figure 1.4.** Carte délimitée par le cercle Arctique et le pôle Nord, mettant en évidence le site d'étude. Le triangle noir indique la localisation de la région des Eaux du Nord, au Nord de la Baie de Baffin et entre l'île d'Ellesmere et le Groenland (Modifiée de IBCAO, 2008).

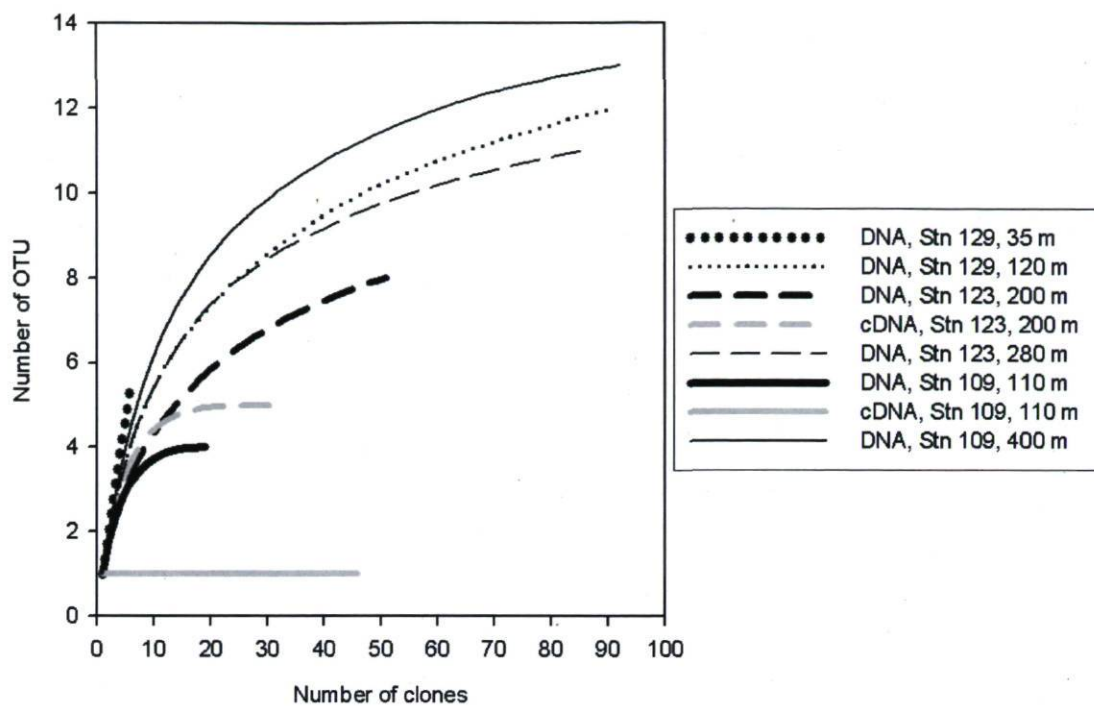


**Figure 2.1.** Map of the North Water area showing location of sampling stations 109 ( $76^{\circ} 15.28$  N,  $74^{\circ} 10.39$  W), 123 ( $77^{\circ} 20.48$  N,  $74^{\circ} 38.37$  W) and 129 ( $78^{\circ} 19.75$  N,  $74^{\circ} 0.87$  W). Bathymetry is indicated in isobaths and colour shaded. Main currents are illustrated with arrows. AW : Arctic Water; WGC : West Greenland Current (Adapted from Hamilton *et al.*, 2008).

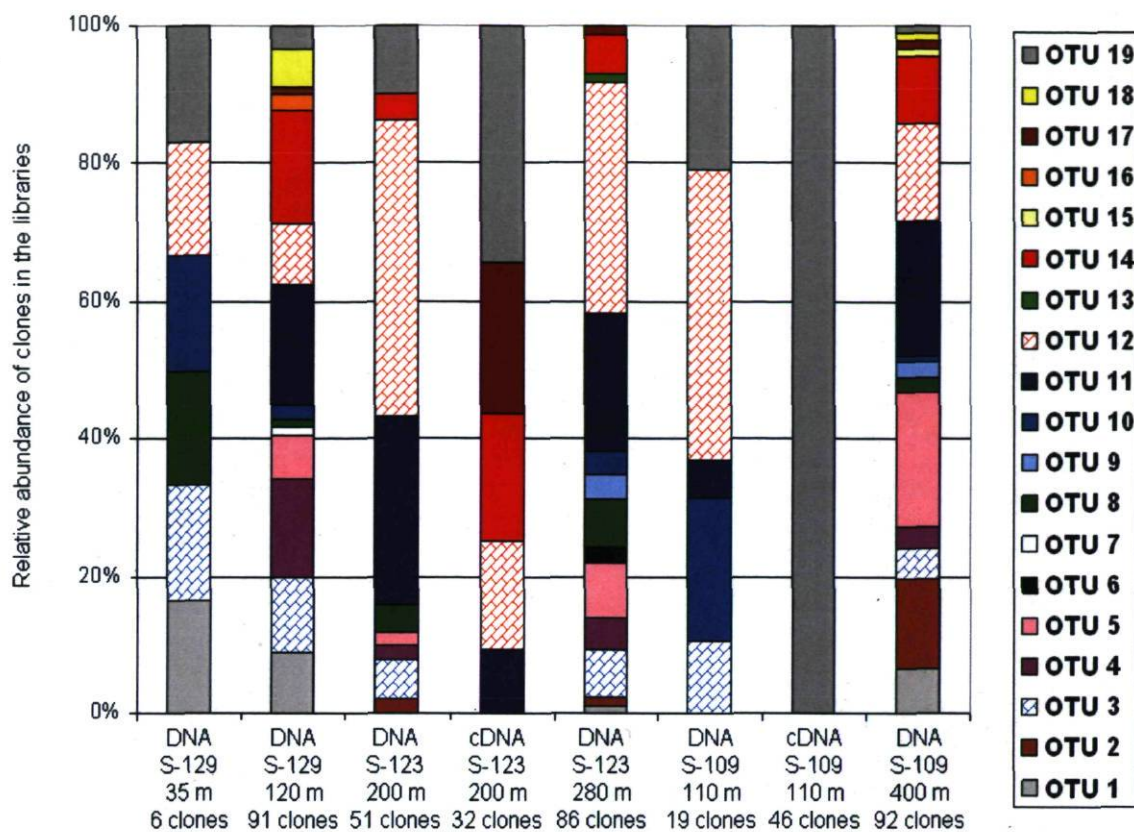


**Figure 2.2.** Depth profiles for stations 129, 123 and 109. Column A) Salinity : black line; Dissolved oxygen concentration ( $DO_2$ ) : gray line; Column B) Nitrite concentration : black squares; Nitrate concentration : gray diamonds; Column C) Number of crenarchaeal 16S gene copy  $ml^{-1}$  (16S Cren<sub>DNA</sub>) and Column E) Number of archaeal *amoA* gene copy  $ml^{-1}$  (*amoA*<sub>DNA</sub>) : black circles; Column D) Number of crenarchaeal 16S gene transcript copy  $ml^{-1}$  (16S Cren<sub>cDNA</sub>) and Column F) Number of archaeal *amoA* gene transcript copy  $ml^{-1}$  (*amoA*<sub>cDNA</sub>) : no-filled circles. All numbers of copy  $ml^{-1}$  are on logarithmic scale. The asterisk corresponds to the slight increase in salinity associated with the decline in  $DO_2$ . Arrows indicate samples for which clone libraries were constructed for diversity analysis in this study.



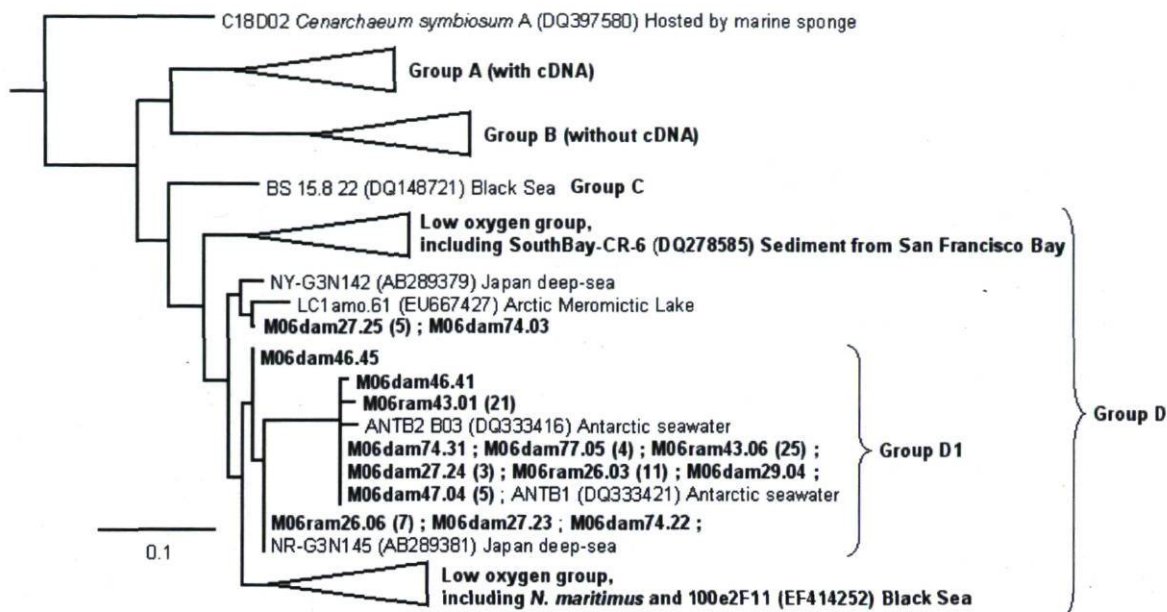


**Figure 2.3.** Rarefaction curves for clone libraries of the archaeal *amoA* gene (DNA and cDNA) assemblages, calculated at a similarity level of  $\geq 97\%$ . DNA: black; cDNA: grey; Station 129: dots (normal and bold); Station 123: dashes (normal and bold); Station 109: lines (normal and bold).



**Figure 2.4.** Community structure of archaeal *amoA* gene (DNA and cDNA) from clone libraries at stations 129, 123 and 109. Sampling depths and number of clones per library are indicated. Caption shows the corresponding color of each OTU. S: station.

(a)

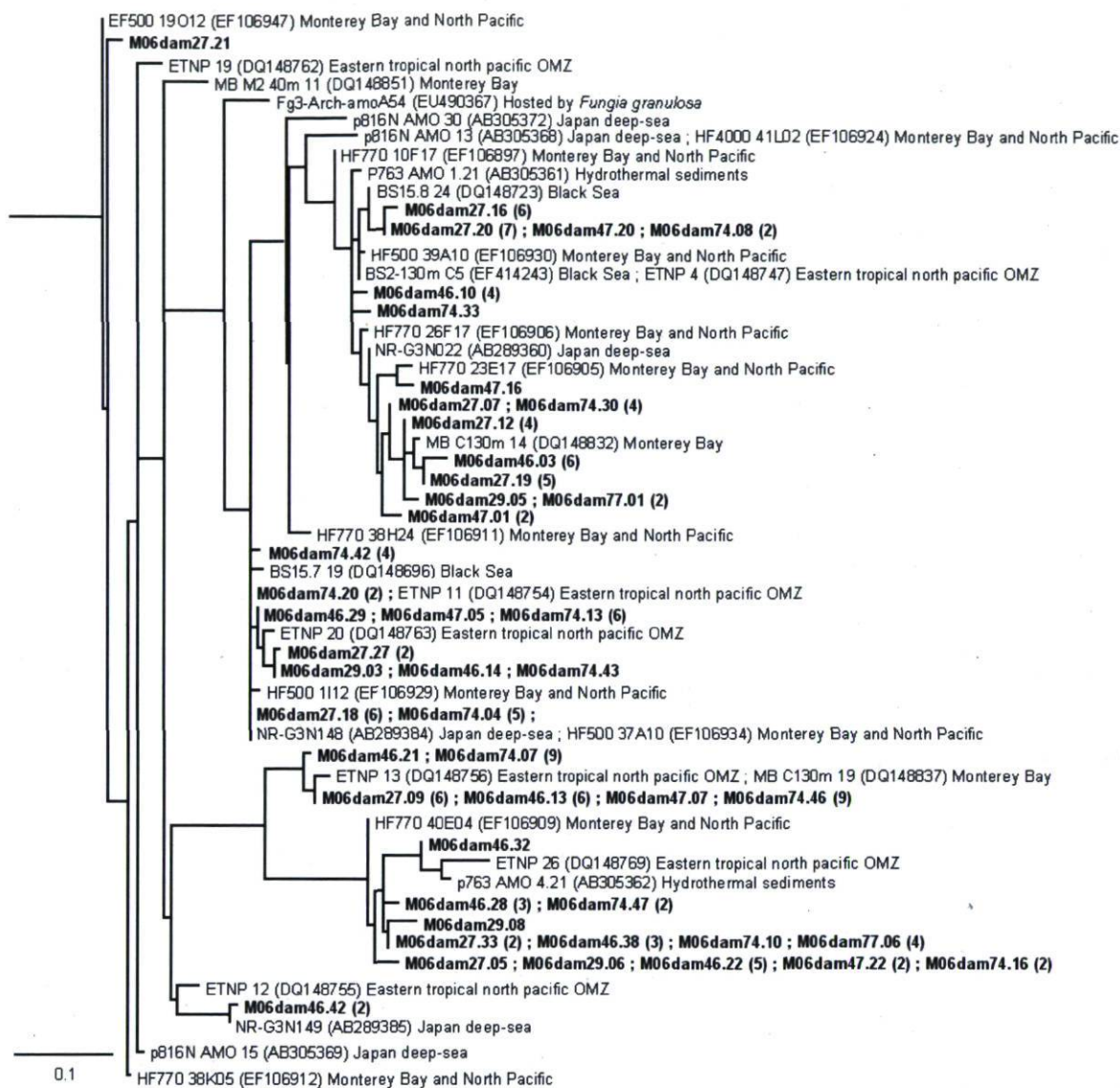


(b)





(c)



**Figure 2.5.** Phylogenetic relationships using maximum likelihood among archaeal *amoA* sequences from the North Water and previously reported sequences. A condensed tree shows relationships between groups (a), details of group A (b) and group B (c). Scale bar represents 10% sequence divergence. Sequences from this study are named following a code in which M stands for the cruise name (MMBOAS : Molecular Microbial Biodiversity of Arctic Seas), 06 stands for the sampling year, d or r stand for DNA or RNA (cDNA), am stands for *amoA*, the following number indicate the sample (26: RNA from station 123 at 200 m; 27: DNA from station 129 at 120 m; 29: DNA from station 129 at 35 m; 43: RNA from station 109 at 110 m; 46: DNA from station 123 at 280 m; 47: DNA from station 123 at 200 m; 74: DNA from station 109 at 400 m; 77: DNA from station 109 at 110 m) and the last number indicates the sequence number in the library. One representative sequence per library is shown and numbers in parenthesis indicate how many clones are represented by that sequence. The outgroup was the soil fosmid clone 54d9 (AJ627422) and has been removed from the tree for clarity.

## **Bibliographie**

- ACIA, 2005. Arctic Climate Impact Assessment. Cambridge University Press, New York, 1042 p.
- Acinas, S.G., Marcelino, L.A., Klepac-Ceraj, V. et M.F. Polz, 2004. Divergence and redundancy of 16S rRNA sequences in genomes with multiple *rrn* operons. *J. Bacteriol.* 186(9) :2629-2635.
- Agogu , H., Brink, M., Dinasquet, J. et G.J. Herndl, 2008. Major gradients in putatively nitrifying and non-nitrifying Archaea in the deep North Atlantic. *Nature* 456 :788-791.
- Aljanabi, S.M. et I. Martinez, 1997. Universal and rapid salt-extraction of High quality genomic DNA for PCR-based techniques. *Nucleic Acids Res.* 25(22) :4692-4693.
- Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W. et D. Lipman, 1997. Gapped BLAST and PSI-BLAST : a new g n ration of protein database search programs. *Nucleic Acids Res.* 25 :3389-3402.
- Arrigo, K.R., 2005. Marine microorganisms and global nutrient cycles. *Nature* 437(7057):349-355.
- B cle, J. Carmack, E.C. et R.G. Ingram, 2002. Water column structure and circulation Under the North Water during spring transition : April-July 1998. *Deep-Sea Res. II* 49 :4907-4925.
- Barns, S.M., Delwiche, C.F., Palmer, J.D. et N.R. Pace, 1996. Perspectives on archaeal diversity, thermophily and monophyly from environmental rRNA sequences. *PNAS* 93(17):9188-9193.
- Brochier-Armanet, C., Boussau, B., Gribaldo, S. et P. Forterre, 2008. Mesophilic crenarchaeota: proposal for a third archaeal phylum, the Thaumarchaeota. *Nat. Rev. Microbiol.* 6(3):245-252.
- Canfield, D.E., Kristensen, E. et B. Thamdrup, 2005. The nitrogen cycle. *Adv. Mar. Biol.* 48:205-267.
- Carmack, E.C., 2007. The alpha/beta ocean distinction: A perspective on freshwater fluxes, convection, nutrients and productivity in high-latitude seas. *Deep-Sea Res. II* 54(23-26) :2578-2598.



- Church, M.J., DeLong, E.F., Ducklow, H.W., Karner, M.B., Preston, C.M. et D.M. Karl, 2003. Abundance and distribution of planktonic Archaea and Bacteria in the waters west of the Antarctic Peninsula. *Limnol. Oceanogr.* 48(5):1893-1902.
- Church, M.J., Short, C.M., Jenkins, B.D., Karl, D.M. et J.P. Zehr, 2005. Temporal Patterns of Nitrogenase Gene (*nifH*) Expression in the Oligotrophic North Pacific Ocean. *Appl. Environ. Microbiol.* 71(9):5362-5370.
- Coolen, M.J.L., Abbas, B., van Bleijswijk, J., Hopmans, E.C., Kuypers, M.M.M., Wakeham, S.G. et J.S. Sinninghe-Damsté, 2007. Putative ammonia-oxidizing Crenarchaeota in suboxic waters of the Black Sea: a basin-wide ecological study using 16S ribosomal and functional genes and membrane lipids. *Environ. Microbiol.* 9(4):1001-1016.
- DeLong, E.F., 1992. Archaea in costal marine environments. *Proc. Natl Acad. Sci.* 89: 5685-5689.
- DeLong, E.F., Preston, C.M., Mincer, T., Rich, V., Hallam, S.J., Frigaard, N.U., Martinez, A., Sullivan, M.B., Edwards, R., Rodriguez-Brito, B., Chisholm, S.W. et D.M. Karl, 2006. Community genomics among stratified microbial assemblages in the ocean's interior. *Science* 311(5760):496-503.
- Dugdale, R.C. et J.J. Goring, 1967. Uptake of new and regenerated forms of nitrogen in primary productivity. *Limnol. Oceanogr.* 12:196-206.
- Francis, C.A., Roberts, K.J., Beman, J.M., Santoro, A.E. et B.B. Oakley, 2005. Ubiquity and diversity of ammonia-oxidizing archaea in water columns and sediments of the ocean. *PNAS.* 102(41):14683-14688.
- Fuhrman, J.A., McCallum, K. et A.A. Davis, 1992. Novel major archaeobacterial group from marine plankton. *Nature* 356 :148-149.
- Galand, P.E., Lovejoy, C. et W.F. Vincent, 2006. Remarkably diverse and contrasting archaeal communities in a large arctic river and the coastal Arctic Ocean. *Aquat. Microbiol. Ecol.* 44 :115-126.
- Galand, P.E., Lovejoy, C., Hamilton, A.K., Ingram, R.G., Pedneault, E. et E.C. Carmack, 2009. Archaeal diversity and a gene for ammonia oxidation are coupled to oceanic circulation. *Environ. Microbiol.* 11(4) :971-980.



- Granéli, W. et E. Granéli, 1991. Automatic potentiometric determination of dissolved oxygen. *Mar. Biol.* 108 :341-348.
- Grasshoff, K., 1976. *Methods of seawater analyses*. Verlag Chemie, Weinheim, New York.
- Hallam, S.J., Mincer, T., Schleper, C., Preston, C.M., Roberts, K., Richardson, P.M. et E.F. DeLong, 2006. Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLOS Biol.* 4(4) :520-536.
- Hamilton, A.K., Lovejoy, C., Galand, P.E. et R.G. Ingram, 2008. Water masses and biogeography of picoeukaryote assemblages in a cold hydrographically complex system. *Limnol. Oceanogr.* 53 :922-935.
- Herfort, L., Schouten, S., Abbas, B., Veldhuis, M.J.W., Coolen, M.J.L., Wuchter, C., Boon, J.P., Herndl, G.J. et J.S. Sinninghe-Damsté, 2007. Variations in spatial and temporal distribution of Archaea in the North Sea in relation to environmental variables. *FEMS Microbiol. Ecol.* 62 :242-257.
- Hillebrand, H., Dürselen, C.D., Kirschtel, D., Pollinger, U. et T. Zohary, 1999. Biovolume calculation for pelagic and benthic microalgae. *J. Phycol.* 35: 403-424.
- Huber, H., Hohn, M.J., Rachel, R., Fuchs, T., Wimmer, V.C. et K.O. Stetter, 2002. A new phylum of Archaea represented by a nanosized hyperthermophilic symbiont. *Nature* 417: 63-67.
- International Bathymetric Chart of the Arctic Ocean (IBCAO), page consultée le 4 août 2008. IBCAO current map of Arctic Ocean bathymetry, [En ligne], URL: <http://www.ngdc.noaa.gov/mgg/bathymetry/arctic/currentmap.html>
- Ingram, R.G., Bâcle, J., Barber, D.G., Gratton, Y. et H. Melling, 2002. An overview of physical processes in the North Water. *Deep-Sea Res. II.* 49 : 4893-4906.
- Kalanetra, K.M., Bano, N. et J.T. Hollibaugh, 2009. Ammonia-oxidizing Archaea in the Arctic Ocean and Antarctic coastal waters. *Environ. Microbiol.* doi:10.1111/j.1462-2920.2009.01974.x
- Karl, D., Michaels, A., Bergman, B., Capone, D., Carpenter, E., Letelier, R., Lipschultz, F., Pearl, H., Sigman, D. et L. Stal, 2002. Dinitrogen fixation in the world's oceans. *Biogeochemistry.* 57-58(1):47-98.
- Kirchman, D.L., 2000. *Microbial ecology of the oceans*, Wiley-Liss, New York, 542 p.

- Klein, B., LeBlanc, B., Mei, Z.P., Beret, R., Michaud, J., Mundy, C.J., von Quillfeldt, C.H., Garneau, M.E., Roy, S., Gratton, Y., Cochran, J.K., Bélanger, S., Larouche, P., Pakulski, J.D., Rivkin, R.B. et L. Legendre, 2002. Phytoplankton biomass production and potential export in the North Water. *Deep-Sea Res. II* 49(22-23):4983-5002.
- Könneke, M., Bernhard, A.E., de la Torre, J.R., Walker, C.B., Waterbury, J.B. et D.A. Stahl, 2005. Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* 437:543-546.
- Lam, P., Jensen, M.M., Lavik, G., McGinnis, D.F., Müller, B., Schubert, C.J., Amann, R., Thamdrup, B. et M.M.M Kuypers, 2007. Linking crenarchaeal and bacterial nitrification to anammox in the Black Sea. *PNAS* 104(17):7104-7109.
- Lewis, E.L., Ponton, D., Legendre, L. et B. LeBlanc, 1996. Springtime sensible heat, nutrients and phytoplankton in the North Water Polynya, Canadian Arctic. *Cont. Shelf Res.* 16: 1775-1792.
- Lomas, M.W. et F. Lipschultz, 2006. Forming the primary nitrite maximum: nitrifiers or phytoplankton? *Limnol. Oceanogr.* 51(5):2453-2467.
- Lovejoy, C., Carmack, E.C., Legendre, L. et N.M. Price, 2002. Water column interleaving: a new physical mechanism determining protist communities and bacterial states. *Limnol. Oceanogr.* 47:1819-1831.
- Mei, Z.P., Legendre, L., Gratton, Y., Tremblay, J.E., LeBlanc, B., Mundy, C.J., Klein, B., Gosselin, M., Larouche, P., Papakyriakou, T.N., Lovejoy, C. et C.H. von Quillfeldt, 2002. Physical control of spring-summer phytoplankton dynamics in the North Water, April-July 1998. *Deep-Sea Res. II* 49:4959-4982.
- Melling, H., Gratton, Y. et R.G. Ingram, 2001. Ocean circulation within the North Water Polynya of Baffin Bay. *Atmos. Ocean.* 39:301-325.
- Miller, C.B., 2004. *Biological oceanography*, Blackwell Publishing, Oxford, p.22
- Mincer, T., Church, M.J., Taylor, L.T., Preston, C., Karl, D.M. et E.F. DeLong, 2007. Quantitative distribution of presumptive archaeal and bacterial nitrifiers in Monterey Bay and the North Pacific subtropical gyre. *Environ. Microbiol.* 9(5):1162-1175.



- Nicholas, K.B. et H.B.J. Nicholas, 1997. GeneDoc: Analysis and visualization of genetic variation. <http://www.cris.com/~Ketchup/genedoc.shtml>.
- Noble, R.T. et J.A. Fuhrman, 1998. Use of SYBR Green I for rapid epifluorescence counts of marine viruses and bacteria. *Aquat. Microb. Ecol.* 14:113-118.
- Nusch, E., 1980. Comparison of different methods for chlorophyll and phaeopigment determination. *Arch. Hydrobiol. Beih.* 14:14-36.
- Parker, M.S. et E.V. Armbrust, 2005. Synergistic effects of light, temperature, and nitrogen source on transcription of genes for carbon and nitrogen metabolism in the centric diatom *Thalassiosira pseudonana* (Bacillariophyceae). *J. Phycol.* 41:1142-1153.
- Pedros-Alio, C., 2006. Marine microbial diversity: can it be determined? *Trends Microbiol.* 14(6):257-263.
- Pedros-Alio, C., 2007. Dipping into the rare biosphere. *Science* 315:192-193.
- Perry, J.J., Staley, J.T. et S. Lory, 2004. *Microbiologie*, Dunod, Paris, 891p.
- Porter, K.G. et Y.S. Feig, 1980. The use of DAPI for identifying and counting aquatic microflora. *Limnol. Oceanogr.* 25:943-948.
- Pouliot, J., Galand, P.E., Lovejoy, C. et W.F. Vincent, 2009. Vertical structure of archaeal communities and the distribution of ammonia monooxygenase A gene variants in two meromictic high Arctic lakes. *Environ. Microbiol.* 11(3):687-699.
- Prescott, L.M., Harley, J.P. et D.A. Klein, 1995. *Microbiologie*, 2e éd., De Boeck, Bruxelles, 1014p.
- Preston, C.M., Wu, K.Y., Molinski, T.F. et E.F. DeLong, 1996. A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *PNAS* 93:6241-6246.
- Prosser, J.I. et G.W. Nicol, 2008. Relative contributions of Archaea and Bacteria to aerobic ammonia oxidation in the environment. *Environ. Microbiol.* 10(11):2931-2941.
- Redfield, A.C., Ketchum, B.H. et F.A. Richards, 1963. The influence of organisms on the composition of sea water. Dans, *The Sea; ideas and observations*, 2, 26-77. Interscience Publishers, New York, 554 pp.
- Schleper, C., Jurgens, G. et M. Jonuscheit, 2005. Genomic studies of uncultivated Archaea. *Nat. Rev. Microbiol.* 3(6): 479-488.



- Schloss, P.D. et J. Handelsman, 2005. Introducing Dotur, a computer program for defining operational taxonomic units and estimating species richness. *Appl. Environ. Microbiol.* 71:1501-1506.
- Stirling, I., 1980. The biological importance of polynyas in the Canadian Arctic. *Arctic* 33:303-315.
- Suttle, C.A., 2007. Marine viruses – major players in the global ecosystem. *Nat. Rev. Microbiol.* 5(10):801-812.
- Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F. et D.G. Higgins, 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876-4882.
- Thurman, H.V. et E.A. Burton, 2001. *Introductory oceanography*, 9<sup>e</sup> éd., Prentice Hall, New Jersey, 554 p.
- Tremblay, J.E., Gratton, Y., Fauchot, J. et N.M. Price, 2002. Climatic and oceanic forcing of new, net, and diatom production in the North Water. *Deep-Sea Res. II.* 49(22-23):4927-4946.
- Ward, B.B., 2005. Molecular approaches to marine microbial ecology and the marine nitrogen cycle. *Annu. Rev. Earth Planet. Sci.* 33(1):301-333.
- Wassmann, P., Reigstad, M., Haug, T., Rudels, B., Carroll, M.L., Hop, H., Gabrielsen, G.W., Falk-Petersen, S., Denisenko, S.G., Arashkevich, E., Slagstad, D. et O. Pavlova, 2006. Food webs and carbon flux in the Barents Sea. *Prog. Oceanogr.* 71(2-4):232-287.
- Wayne, L.G., Brenner, D.J., Colwell, R.R., Grimont P.A.D., Kandler, O., Krichevsky, M., Moore, L.H., Moore, W.E.C., Murray, R.G.E., Stackebrandt, E., Starr, M.P. et H.G. Trüper, 1987. Report of the ad hoc committee on reconciliation of approaches to bacterial systematics. *Int. J. Syst. Bacteriol.* 37:463-464.
- Woese, C.R. et G.E. Fox, 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *PNAS.* 74(11):5088-5090.
- Woese, C.R., Kandler, O. et M.L. Wheelis, 1990. Towards a natural system of organisms : Proposal for the domain Archaea, Bacteria, and Eucarya. *PNAS* 87 :4576-4579.

- Wuchter, C., Abbas, B., Coolen, M.J.L., Herfort, L., van Bleijswijk, J., Timmers, P., Strous, M., Teira, E., Herndl, G.J., Middelburg, J.J., Schouten, S. et J.S.S. Damsté, 2006. Archaeal nitrification in the ocean. *PNAS* 103(33):12317-12322.
- Yamamoto-Kawai, M., Carmack, E. et F. McLaughlin, 2006. Nitrogen balance and Arctic throughflow. *Nature* 443:43.
- Zehr, J.P., Jenkins, B.D., Short, S.M. et G.F. Steward, 2003. Nitrogenase gene diversity and microbial community structure: a cross-system comparison. *Environ. Microbiol.* 5(7):539-554.
- Zhu, F., Massana, R., Not, F., Marie, D. et D. Vaultot, 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *FEMS Microbiol. Ecol.* 52:79-92.

## Annexe 1

Caractéristiques biophysicochimiques des échantillons recueillis à la (a) station 129 (b) station 123 et (c) station 109 dans les Eaux du Nord. DO<sub>2</sub> : concentration d'oxygène dissous ; PAR : rayonnement photosynthétique actif ; Nox : concentration de nitrate plus nitrite ; VLP : pseudo particules virales (de l'anglais « virus like particules ») ; Cren : Crenarchaeota ; nd : aucune donnée.

(a)

Station	129	129	129	129	129	129	129	129
Depth (m)	10	22	35	80	120	160	300	500
Salinity	31.4	31.8	32.6	33.1	33.5	33.9	34.2	34.3
Temperature (°C)	0.2	0.4	0.8	-0.1	-0.9	-0.7	-0.4	-0.4
Density (kg m <sup>-3</sup> )	25.2	25.5	26.1	26.6	26.9	27.2	27.5	27.5
DO <sub>2</sub> (mL L <sup>-1</sup> )	8.2	8.2	7.8	7.5	7.0	6.6	6.4	6.4
PAR	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg L <sup>-1</sup> )	0.6	1.4	1.2	0.2	0.1	0.0	0.0	0.0
Nitrite (µM)	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1
Nitrate (µM)	0.0	0.0	2.7	4.2	8.6	11.7	12.6	11.9
Nox (µM)	0.6	1.4	1.2	0.2	0.1	0.0	0.0	0.0
Silicate (µM)	nd	nd	nd	nd	nd	nd	nd	nd
Phosphate (µM)	nd	nd	nd	nd	nd	nd	nd	nd
Virus (10 <sup>9</sup> VLP mL <sup>-1</sup> )	12.1	24.7	24.3	4.6	15.8	4.5	4.4	3.2
Bacteria (10 <sup>7</sup> cells mL <sup>-1</sup> )	9.4	21.7	22.8	16.3	26.5	14.0	11.5	0.9
Bacteria Grazers (pgC L <sup>-1</sup> )	711.10	496.76	521.09	477.31	187.29	157.19	151.77	134.53
Protists Grazers (pgC L <sup>-1</sup> )	799.84	454.19	703.73	4305.43	6037.03	4959.94	1824.54	989.23
DNA Archaea 16S (10 <sup>4</sup> copies mL <sup>-1</sup> )	2.08 (0.17)	3.24 (0.72)	2.81 (0.77)	2.06 (0.11)	18.84 (3.08)	7.33 (0.6)	12.79 (2.94)	9.82 (0.47)
DNA <i>amoA</i> (10 <sup>3</sup> copies mL <sup>-1</sup> )	0.51 (0.06)	2.22 (0.26)	4.77 (0.85)	3.46 (0.36)	28.72 (1.13)	12.95 (2.11)	16.08 (2.24)	11.86 (3.42)
DNA Cren 16S (10 <sup>3</sup> copies mL <sup>-1</sup> )	1.56 (0.96)	7.36 (1.03)	13.09 (4.4)	12.17 (3.04)	108.68 (8.58)	27.25 (1.55)	107.55 (13.13)	106.06 (9.96)
cDNA <i>amoA</i> (10 <sup>3</sup> copies mL <sup>-1</sup> )	0.71 (0.17)	5546.78 (607.74)	5846.12 (698.21)	3753.22 (658.52)	992.05 (139.95)	1.27 (0.34)	nd	nd
cDNA Cren 16S (10 <sup>3</sup> copies mL <sup>-1</sup> )	60.09 (12.1)	2754.85 (540.41)	6267.94 (961.03)	48304.28 (2102.59)	218232.7 (59414.89)	2342.69 (283.63)	nd	nd
Ratio DNA Cren 16S : DNA Archaea 16S	0.1	0.2	0.5	0.5	0.6	0.6	0.8	1.1
Ratio DNA <i>amoA</i> : DNA Archaea 16S	0.0	0.1	0.2	0.1	0.2	0.2	0.1	0.1
Ratio DNA <i>amoA</i> : DNA Cren 16S	0.3	0.3	0.4	0.3	0.3	0.3	0.2	0.1
Ratio cDNA <i>amoA</i> : DNA <i>amoA</i>	472.6	3308.5	1777.8	2845.9	128.2	3.1	0.0	0.0
Ratio cDNA Cren : DNA Cren	38.6	374.5	478.7	4939.5	2008.1	50.4	0.0	0.0



(b)

Station	123	123	123	123	123	123	123	123
Depth ( m)	10	45	100	120	200	280	350	500
Salinity	32.1	32.5	33.4	33.5	33.9	34.1	34.2	34.3
Temperature (°C)	-0.6	-0.4	-0.7	-1.0	-0.3	0.2	-0.2	-0.4
Density (kg m <sup>-3</sup> )	25.8	26.1	26.9	27.0	27.2	27.4	27.5	27.6
DO <sub>2</sub> (mL L <sup>-1</sup> )	8.1	7.7	7.2	7.1	6.4	5.9	6.3	6.5
PAR	8.7	0.2	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll a (µg L <sup>-1</sup> )	0.7	0.4	0.1	0.1	0.1	0.1	0.0	0.0
Nitrite (µM)	0.1	0.1	0.2	0.1	0.1	0.1	0.1	0.1
Nitrate (µM)	3.0	4.9	8.9	10.4	14.0	15.3	13.9	13.6
Nox (µM)	9.4	0.7	0.1	0.1	0.1	0.1	0.0	0.0
Silicate (µM)	5.8	7.7	8.8	11.2	13.7	16.7	13.8	12.5
Phosphate (µM)	0.8	0.8	1.0	1.0	1.1	1.2	1.1	1.0
Virus (10 <sup>9</sup> VLP mL <sup>-1</sup> )	8.3	12.9	13.5	7.1	6.2	6.7	4.9	2.5
Bacteria (10 <sup>7</sup> cells mL <sup>-1</sup> )	42.0	18.7	17.3	17.9	24.7	18.7	10.3	9.7
Bacteria Grazers (pgC L <sup>-1</sup> )	387.14	65.93	162.95	208.68	75.64	19.38	139.59	53.30
Protists Grazers (pgC L <sup>-1</sup> )	1345.51	380.33	406.23	2307.36	1682.45	1521.98	1476.20	1022.90
DNA archaea 16S (10 <sup>4</sup> copies mL <sup>-1</sup> )	1.85 (0.39)	8.47 (0.81)	57.76 (6.54)	14.13 (1.17)	16.28 (1.66)	126.5 (25.92)	125.66 (9.02)	31.59 (3.42)
DNA <i>amoA</i> (10 <sup>3</sup> copies mL <sup>-1</sup> )	1.49 (0.26)	12.76 (1.68)	59.94 (50.45)	21.43 (7.8)	22.83 (1.63)	330.51 (43.93)	175.67 (29.23)	16.77 (4.2)
DNA cren 16S (10 <sup>3</sup> copies mL <sup>-1</sup> )	8.67 (1.01)	37.31 (6.21)	128.17 (8.4)	43.54 (14.74)	56.68 (17.07)	1550.54 (357.27)	1373.82 (150.94)	279.02 (89.74)
cDNA <i>amoA</i> (10 <sup>3</sup> copies mL <sup>-1</sup> )	nd	27.16 (2.79)	139.37 (18.15)	1259.76 (196.62)	7972.29 (302.68)	2466.31 (17.25)	792.88 (193.46)	nd
cDNA cren 16S (10 <sup>3</sup> copies mL <sup>-1</sup> )	nd	28.06 (9.13)	2288.59 (221.89)	89440.77 (12232.52)	588323.68 (122123.22)	1695047.66 (24064.92)	1035795.47 (7352.88)	nd
Ratio DNA cren 16S : DNA archaea 16S	0.5	0.4	0.2	0.3	0.4	1.2	1.1	0.9
Ratio DNA <i>amoA</i> : DNA archaea 16S	0.1	0.2	1.1	0.2	0.1	0.3	0.1	0.1
Ratio DNA <i>amoA</i> : DNA cren 16S	0.2	0.3	5.1	0.5	0.3	0.2	0.1	0.1
Ratio cDNA <i>amoA</i> : DNA <i>amoA</i>	0.0	60.5	0.6	128.8	578.9	11.6	15.8	0.0
Ratio cDNA cren : DNA cren	0.0	0.8	17.9	2054.4	10379.5	957.3	638.9	0.0

(c)

Station	109	109	109	109	109	109	109	109
Depth (m)	25	40	80	90	110	195	210	400
Salinity	32.5	33.0	33.4	33.5	33.6	34.0	34.0	34.3
Temperature (°C)	0.6	-1.1	-0.8	-0.9	-0.9	-0.6	-0.6	-0.4
Density (kg m <sup>-3</sup> )	26.0	26.5	26.9	26.9	27.0	27.3	27.3	27.6
DO <sub>2</sub> (mL L <sup>-1</sup> )	7.6	7.7	7.4	7.3	7.0	6.6	6.6	6.3
PAR	3.2	0.7	0.0	0.0	0.0	0.0	0.0	0.0
Chlorophyll <i>a</i> (µg L <sup>-1</sup> )	1.3	0.0	0.4	0.1	0.0	0.0	0.0	0.0
Nitrite (µM)	0.1	0.1	0.2	0.2	0.1	0.1	0.1	0.1
Nitrate (µM)	2.3	7.4	7.5	8.0	9.9	13.3	12.9	13.3
Nox (µM)	4.5	0.7	0.5	0.1	0.0	0.0	0.0	0.0
Silicate (µM)	4.1	10.9	6.2	8.3	8.8	14.6	14.1	14.8
Phosphate (µM)	0.5	1.0	0.9	0.9	0.9	1.1	1.0	1.0
Virus (109 VLP mL <sup>-1</sup> )	24.3	13.9	18.9	8.6	8.1	4.9	3.7	2.9
Bacteria (10 <sup>7</sup> cells mL <sup>-1</sup> )	20.4	40.6	21.0	27.3	17.0	20.1	13.7	14.0
Bacteria Grazers (pgC L <sup>-1</sup> )	563.28	236.37	216.11	84.15	72.68	46.88	80.81	171.15
Protists Grazers (pgC L <sup>-1</sup> )	627.88	1039.49	1814.02	1325.76	2382.67	1264.23	976.92	638.23
DNA archaea 16S (10 <sup>4</sup> copies mL <sup>-1</sup> )	2.44 (0.96)	6.54 (1.13)	10.86 (1)	13.13 (1.74)	18.67 (2.36)	36.03 (15.37)	23.82 (4.25)	11.97 (0.87)
DNA <i>amoA</i> (10 <sup>3</sup> copies mL <sup>-1</sup> )	14.19 (1.78)	8.6 (2.15)	15.67 (2.35)	17.03 (1.36)	28.75 (2.83)	58.59 (9.06)	32.73 (8.98)	13.08 (0.85)
DNA cren 16S (10 <sup>3</sup> copies mL <sup>-1</sup> )	37.35 (2.49)	16.5 (4.62)	15.81 (5.47)	25.9 (0.65)	49.39 (7.51)	149.78 (22.77)	159.21 (47.69)	73.22 (9.32)
cDNA <i>amoA</i> (10 <sup>3</sup> copies mL <sup>-1</sup> )	137.92 (12.22)	111.16 (26.98)	451.81 (65.22)	1196.58 (92)	2087.49 (255.17)	3124.05 (108.14)	nd	nd
cDNA cren 16S (10 <sup>3</sup> copies mL <sup>-1</sup> )	13986.88 (1088.6)	15825.37 (610.22)	26165.25 (6248.26)	53688.24 (9930.04)	157410.29 (21770.58)	504861.07 (90143)	nd	nd
Ratio DNA cren 16S : DNA archaea 16S	1.5	0.3	0.2	0.2	0.2	0.6	0.7	0.6
Ratio DNA <i>amoA</i> : DNA archaea 16S	0.4	0.1	0.1	0.1	0.2	0.2	0.1	0.1
Ratio DNA <i>amoA</i> : DNA cren 16S	0.3	0.5	1.0	0.7	0.7	0.4	0.2	0.2
Ratio cDNA <i>amoA</i> : DNA <i>amoA</i>	151.0	1110.9	2902.1	2631.3	1534.7	6.8	0.0	0.0
Ratio cDNA cren : DNA cren	374.5	959.0	1654.7	2073.0	3795.5	2342.8	0.0	0.0

## Annexe 2

Description des OTU mentionnés à la figure 2.4 et liste des séquences que chacun regroupe. Les clades et la description du code utilisé pour nommer les séquences sont mentionnés à la figure 2.5.

OTU	Clade	Nombre de séquences	Nom de séquence		
1	B	16	M06-dam-27.18 (6) M06-dam-46.14	M06-dam-27.27 (2) M06-dam-74.04 (5)	M06-dam-29.03 M06-dam-74.43
2	B	14	M06-dam-46.29 M06-dam-74.20 (2)	M06-dam-47.05 M06-dam-74.42 (4)	M06-dam-74.13 (6)
3	B	26	M06-dam-27.07 M06-dam-29.05 M06-dam-47.16	M06-dam-27.12 (4) M06-dam-46.03 (6) M06-dam-74.30 (4)	M06-dam-27.19 (5) M06-dam-47.01 (2) M06-dam-77.01 (2)
4	B	21	M06-dam-27.16 (6) M06-dam-47.20	M06-dam-27.20 (7) M06-dam-74.08 (2)	M06-dam-46.10 (4) M06-dam-74.33
5	B	32	M06-dam-27.09 (6) M06-dam-47.07	M06-dam-46.13 (6) M06-dam-74.07 (9)	M06-dam-46.21 M06-dam-74.46 (9)
6	B	2	M06-dam-46.42 (2)		
7	B	1	M06-dam-27.21		
8	B	12	M06-dam-27.05 M06-dam-46.32	M06-dam-29.06 M06-dam-47.22 (2)	M06-dam-46.22 (5) M06-dam-74.16 (2)
9	B	5	M06-dam-46.28 (3) M06-dam-74.47 (2)		
10	B	11	M06-dam-27.33 (2) M06-dam-74.10	M06-dam-29.08 M06-dam-77.06 (4)	M06-dam-46.38 (3)
11	A	69	M06-dam-27.14 (16) M06-dam-74.18 (18)	M06-dam-46.37 (17) M06-dam-77.09	M06-dam-47.19 (14) M06-ram-26.05 (3)
12	A	86	M06-dam-27.28 M06-dam-46.35 M06-dam-46.20 (12) M06-dam-47.02 (8) M06-dam-47.12 M06-dam-74.34 M06-dam-77.11 (6) M06-ram-26.01	M06-dam-27.02 (7) M06-dam-46.17 M06-dam-46.34 M06-dam-47.06 (4) M06-dam-47.14 (5) M06-dam-74.02 (8) M06-dam-77.04 M06-ram-26.04 (2)	M06-dam-29.02 M06-dam-46.39 M06-dam-46.33 (13) M06-dam-47.11 (4) M06-dam-74.14 (4) M06-dam-77.07 M06-ram-26.10 (2)
13	D	1	M06-dam-46.41		
14	A	37	M06-dam-27.35 (15) M06-dam-74.17 (9)	M06-dam-46.05 (5) M06-ram-26.18 (6)	M06-dam-47.03 (2)
15	A	1	M06-dam-74.24		
16	A	2	M06-dam-27.13 (2)		
17	D	10	M06-dam-27.23 M06-ram-26.06 (7)	M06-dam-46.45	M06-dam-74.22
18	D	6	M06-dam-27.25 (5) M06-dam-74.03		
19	D	71	M06-dam-27.24 (3) M06-dam-74.31 M06-ram-43.01 (21)	M06-dam-29.04 M06-dam-77.05 (4) M06-ram-43.06 (25)	M06-dam-47.04 (5) M06-ram-26.03 (11)



## Annexe 3

Séquences d'acides aminés sélectionnées, obtenues par la traduction des séquences nucléotidiques mentionnées à la figure 2.5, et mise en évidence de leurs distinctions.

	10	20	30	40	50	60	70
MO6dam27.18	LMVWLRRC	THYLP	IVVVAVN	STLLTINS	SDYIF	YTDWAWT	SFVVF
MO6dam29.03	V						
MO6dam46.14							
MO6dam74.43	V						
MO6dam46.29							
MO6dam74.13	V						
MO6dam74.20							
MO6dam74.42							
MO6dam27.07						I	T
MO6dam27.12	V					I	T
MO6dam77.01	V					I	T
MO6dam46.10						I	T
MO6dam27.09							GA
MO6dam46.21	V					I	GA
MO6dam46.42	V				V	I	GA
MO6dam27.21	V					I	GA
MO6dam27.05						I	GA
MO6dam47.22	V					I	GA
MO6dam27.14			A		SQ	LT	C
MO6dam46.37	V		A		SQ	LT	C
MO6dam46.41	V		A		YT	SQ	LI
MO6dam27.35	V		A			SQ	LA
MO6dam46.05			A			SQ	LA
MO6dam27.13	V		A		A	SQ	LC
MO6dam27.23			A		YT	SQ	LI
MO6dam46.45	V		A		YT	SQ	LI
MO6dam27.25	V		A		T	SQ	LI

	70	80	90	100	110	120	130	140
MO6dam27.18	TASVYGL	IMQVY	TWVAV	WAFAL	GYVDP	PIVHP	MWIEP	SCMLLD
MO6dam29.03								
MO6dam46.14								
MO6dam74.43								
MO6dam46.29								
MO6dam74.13								
MO6dam74.20								
MO6dam74.42								
MO6dam27.07	T							
MO6dam27.12	T							
MO6dam77.01	T							
MO6dam46.10	T							
MO6dam27.09	GA							
MO6dam46.21	GA							
MO6dam46.42	GA							
MO6dam27.21	GA							
MO6dam27.05	GA		A					
MO6dam47.22	GA		A					
MO6dam27.14	T	A	T	I	G	Y	T	V
MO6dam46.37	T	A	T	I	G	Y	T	V
MO6dam46.41	T	A	T	I	G	Y	T	V
MO6dam27.35	T	A	T	I	G	Y	T	V
MO6dam46.05	T	A	T	I	G	Y	T	V
MO6dam27.13	T	A	T	I	G	Y	T	V
MO6dam27.23	T	A	T	I	G	Y	T	V
MO6dam46.45	T	A	T	I	G	Y	T	V
MO6dam27.25	T	A	T	I	G	Y	T	V

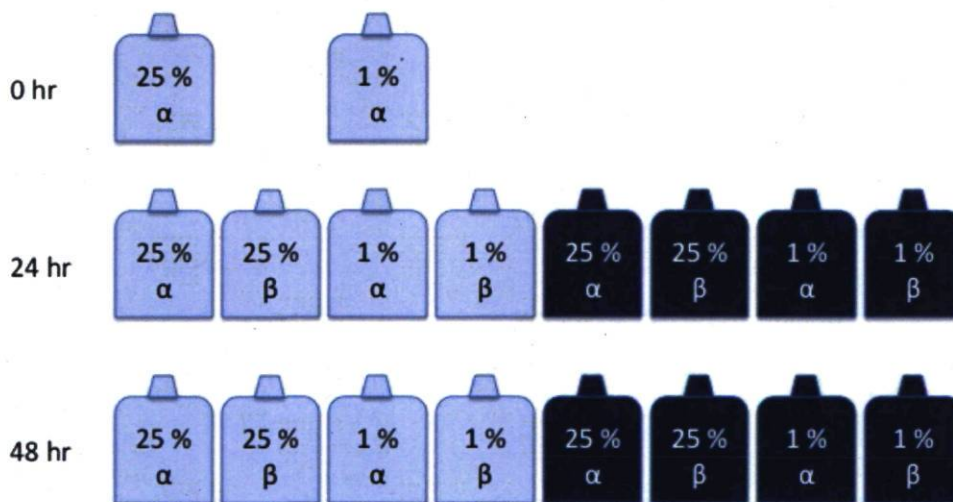
	140	150	160	170	180	190	200	210
MO6dam27.18	LFNMQV	QLMLI	ADPL	ETAFK	YPRPT	LPPYMT	PIEP	QVGR
MO6dam29.03								
MO6dam46.14								
MO6dam74.43								
MO6dam46.29								
MO6dam74.13								
MO6dam74.20								
MO6dam74.42								
MO6dam27.07								
MO6dam27.12								
MO6dam77.01								
MO6dam46.10								
MO6dam27.09		M					SV	L
MO6dam46.21		M					SV	L
MO6dam46.42							SV	L
MO6dam27.21							SV	L
MO6dam27.05							SV	L
MO6dam47.22							SV	L
MO6dam27.14		N	ITV				A	LT
MO6dam46.37		N	ITV				A	LT
MO6dam46.41		N	ITV				A	LT
MO6dam27.35		N	ITV				A	SV
MO6dam46.05		N	ITV				A	SV
MO6dam27.13		N	ITV				A	SV
MO6dam27.23		N	ITV				A	SV
MO6dam46.45		N	ITV				A	SV
MO6dam27.25		N	ITV				A	SV

## Annexe 4

### Expérience sur l'effet de l'éclairement sur l'expression du gène *amoA*

Les échantillons utilisés pour observer l'effet de l'éclairement sur l'expression du gène *amoA* ont été recueillis en septembre 2006, à la station 301, soit à la frontière entre la région des Eaux du Nord et le passage du Nord-Ouest. Deux profondeurs d'échantillonnage ont été sélectionnées, à 25 et 80 m, c'est-à-dire avec respectivement 25 et 1 % de lumière incidente.

Le plan d'échantillonnage (figure 1) implique 18 bouteilles de 1 L, dont 10 transparentes et 8 opaques. L'eau recueillie par le système CTD-rosette a été transférée à ces bouteilles, incluant des duplicatas, puis manipulée et analysée suivant le protocole décrit au chapitre 2, sauf exceptions. Un échantillon par profondeur représente l'état initial, filtré sans délai. Les autres échantillons ont été placés dans un incubateur offrant des conditions stables et adaptées d'éclairement (filtre bleu) et de température (4 °C), pour une durée de 24 ou 48 heures, ce préalablement à la filtration. Un seul filtre de 0.2 µm par échantillon a été utilisé, aucun pré-filtre. Pour chaque échantillon, les résultats disponibles sont l'expression du gène *amoA* (Q-PCR sur ADNc) et les concentrations de virus, de nitrite, de nitrate, de silice et de phosphore. Plusieurs données bio-physico-chimiques caractérisant l'état initial de l'expérience sont également disponibles.



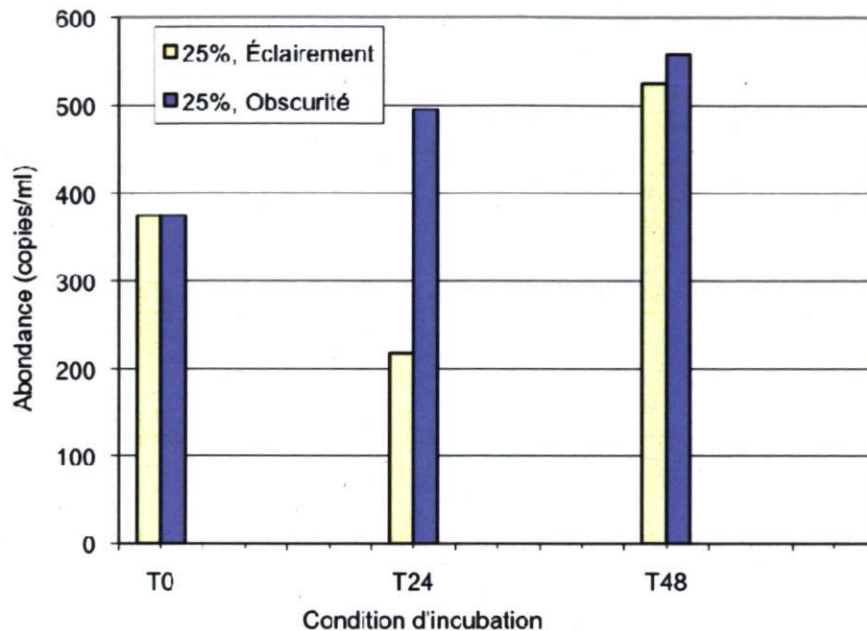
**Figure 1.** Schéma du plan d'échantillonnage pour observer l'effet de l'éclairement sur l'expression du gène *amoA*.

L'hypothèse de départ était que l'éclairement inhibe significativement l'expression du gène *amoA*. Les résultats préliminaires obtenus à l'issue de cette expérience (figure 2) ont démontré d'une part que l'expression du gène *amoA* était moindre sous des conditions d'éclairement par rapport à des conditions d'obscurité. D'autre part, sous des conditions d'obscurité, l'expression du gène *amoA* après incubation de 24 ou 48 heures était supérieure à l'expression initiale. Sous des conditions d'éclairement, l'expression du gène

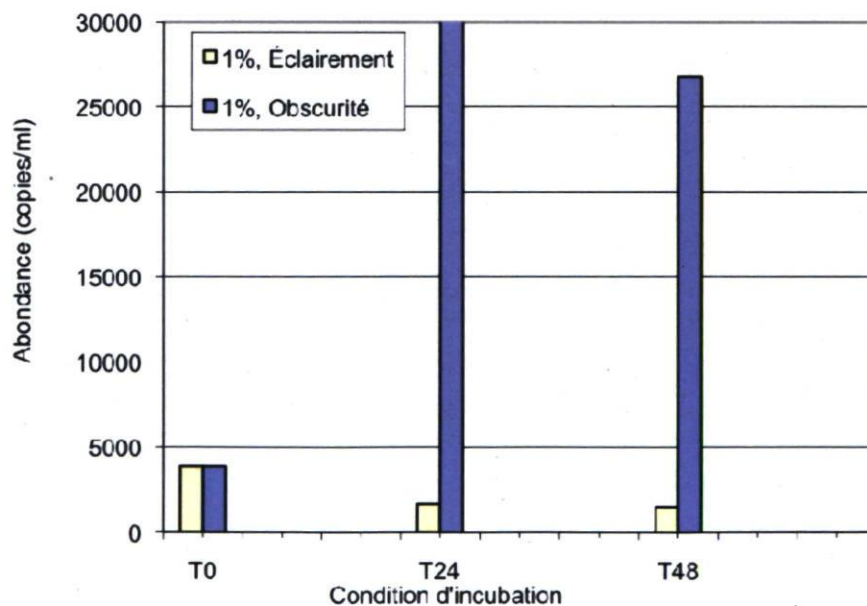


*amoA* après incubation de 24 ou 48 heures était inférieure ou supérieure à l'expression initiale, mais aucune diminution d'expression significative n'a été observée tel que supposé par l'hypothèse de départ.

(a)



(b)



**Figure 2.** Expression du gène *amoA* à partir d'un état initial à (a) 25 % de lumière incidente et (b) 1 % de lumière incidente, puis suite à des conditions d'incubation sous éclairément ou obscurité pour des périodes de 24 ou 48 heures.