

L'acide désoxyribonucléique (ADN) est un acide nucléique qui contient l'information génétique utilisée dans le fonctionnement des processus biologiques dans les organismes vivants et le séquençage de génomes entiers a révélé une énorme quantité d'informations sur les promoteurs de gènes et les éléments transposables (TEs). Dans les organismes eucaryotes, l'ADN et ses protéines associées forment une structure appelée chromatine condensée qui est un échafaudage d'ADN entouré de protéines dans le noyau. La chromatine est constituée par l'unité de base de conditionnement d'ADN appelée nucléosome. Le nucléosome qui est une particule du noyau contient environ 147 paires de bases d'ADN enveloppés dans près de 2 copies chacune des histones H2A, H2B, H3 et H4 comme un fil enroulé autour d'une bobine. Jusqu'à 80 pb d'ADN de liaison avec les histones H1 et H5 relient les particules du noyau.

La chromatine joue plusieurs rôles, par exemple, dans la prévention des dommages de l'ADN ou de l'orchestration, réplication de l'ADN et la régulation des gènes. Les propriétés de la chromatine varient entre les différentes régions du génome, les phases de cycle de vie, et des types de cellules dans le même organisme. La façon dont l'ADN est emballé détermine sa fonction. Une forme serrée de l'ADN est appelée hétérochromatine constitutive et qui comprend l'hétérochromatine facultative. L'hétérochromatine constitutive est généralement composée d'éléments répétitif, globalement distribués au niveau des centromères et télomères et est définitivement réduite au silence transcriptionnel, alors que l'hétérochromatine facultative n'est pas répétitive mais elle possède aussi une structure compacte comme l'hétérochromatine constitutive et elle peut perdre sa structure condensée et devenir transcriptionnellement active sous certaines conditions. Une forme légèrement tassée de l'ADN est appelée euchromatine dans laquelle les gènes sont transcriptionnellement actifs.

Les états chromatinien définissent l'état d'expression des gènes au niveau du génome entier. Alors que les génomes sont programmés pour exprimer un ensemble approprié de gènes à un moment spécifique en réponse à certains stimuli environnementaux. Le contrôle du génome par la chromatine n'est pas au niveau de la séquence d'ADN: les cellules portent la même séquence d'ADN mais des structures et propriétés différentes de leur chromatine ce qui affecte la lecture de la séquence d'ADN sous-jacent. En outre, le contrôle des génomes via par la chromatine peut être transitoire ou stable et transmis par divisions cellulaires, voire à travers les générations.

Les états chromatinien sont modulables en raison des modifications des queues d'histones de la chromatine qui sont soumis à des modifications post-traductionnelles. Jusqu'à présent, plus

de 60 modifications des histones ont été identifiées et connues pour leur grande contribution à la modification des états de la chromatine. La séquence d'ADN peut également faire l'objet de méthylation, qui est impliqués dans le transfert enzymatique d'un groupe méthyle sur la position de la cinquième pyrimidine cytosine ou le numéro 6 azote du cycle purine adénine. La méthylation de l'ADN sur les cytosines et adénines est largement distribuée dans les procaryotes et les eucaryotes. La modification principalement ou exclusivement sur les cytosines (5mC) peut être héritée par division cellulaire.

L'épigénétique est l'étude de la variation des états héréditaires chromatine sans modification de la séquence d'ADN sous-jacente. L'état de la chromatine est contrôlé par deux grands éléments épigénétiques: la méthylation de l'ADN et les modifications des histones. La somme des informations sur le génome entier de modifications de l'ADN et de la chromatine est appelé l'épigénome. Ces deux éléments épigénomiques principaux peuvent avoir des effets profonds sur l'état de la chromatine, qui régule ainsi des processus tels que l'expression des gènes.

Contrairement au génome, l'épigénome est variable entre les cellules, les différents stades de développement et en réponse à différents stimuli. Récemment, certaines études suggèrent que non seulement les deux éléments épigénétiques classiques mentionnés ci-dessus sont impliqués dans l'organisation du génome et le développement, mais d'autres aspects de la régulation épigénétique, tels que les petits ARN ont également été montrés jouer un rôle central dans la régulation des gènes. En outre, en 2011, il a été découvert que l'ARNm peut être soumis à la méthylation N6-méthyladénosine et il est proposé qu'il joue un rôle essentiel dans l'homéostasie énergétique humaine. Cela a ouvert le champ de l'épigénétique ARN.

Il a été proposé que le succès écologique du phytoplancton est également due à la dynamique adaptative conférée par des mécanismes de régulation épigénétiques car des mutations ponctuelles basées sur des processus peut être trop lent pour permettre une adaptation à l'environnement dynamique des océans. Dans les organismes marins, les mécanismes épigénétiques sont susceptibles d'être impliqués et jouent probablement un rôle encore plus vital par rapport à la régulation génétique pour une meilleure adaptation à l'environnement marin. Les phénomènes épigénétiques sont associées à des modes «doux» de l'héritage, contrairement aux modes dits «durs» de transmission des phénomènes génétiques. L'évolution des séquences d'ADN serait impliquée dans la stimulation des organismes dans l'océan pour s'adapter à leur environnement au cours des temps géologiques, tandis que les changements

épigénétiques pourraient conférer la plasticité phénotypique des individus sur des périodes beaucoup plus courtes.

Les exemples discutés ci-dessus et bien d'autres montrent une conservation remarquable chez les mammifères, les plantes et les invertébrés des mécanismes épigénétiques régulant l'expression des gènes. Cette conservation semble aller au-delà de ces espèces, y compris des organismes unicellulaires comme les diatomées. Les diatomées constituent le groupe d'algues le plus vigoureux et le phytoplancton eucaryote le plus abondant avec une grande diversité et une large diffusion tant en mer qu'en eau douce. Les diatomées sont censées contribuer à au moins 20% de la productivité primaire annuelle qui est l'équivalent de toutes les forêts tropicales combinées. On estime qu'environ 40% de la matière organique produite annuellement dans l'océan sont attribuées aux diatomées. Les diatomées sont également bien connues pour leur structure fine et élégante de la frustule de silice produite par l'utilisation de l'acide silicique dissous dans l'eau.

Il est connu que la régulation génétique est impliquée dans le traitement de l'adaptation à un environnement changeant. Cependant, la génétique seule ne peut expliquer tous les changements rapides observés en réponse à des environnements fluctuants. Il est de plus en plus considéré que l'épigénétique peut-être derrière certains des mécanismes qui permettent à certains organismes de faire face à des environnements hostiles. L'impact de la régulation épigénétique sur la diversité et l'adaptation est largement hypothétique, même si les phénomènes de changements épigénétiques sur les individus et les populations en réponse à des stress environnementaux ont été observés dans de nombreux organismes. Une «adaptation flexible» causée par des changements épigénétiques peut donc être approprié pour les réponses à court terme du changement de l'environnement.

L'étude des diatomées est un domaine fascinant qui attire des scientifiques de différentes disciplines. Les scientifiques des nano technologies tentent de comprendre et de reproduire les procédures de production de la frustule de silice fine des diatomées. La biomasse des diatomées antique était un contributeur majeur aux combustibles fossiles, et les diatomées d'aujourd'hui sont étudiées en tant que source de carbone renouvelable, carburants neutres pour l'avenir. Les diatomées sont également très attractive pour les écologistes: les diatomées sont des indicateurs environnementaux dans les eaux douces et sans doute peuvent jouer un rôle fondamental dans la compensation du réchauffement climatique.

J'essaie d'expliquer le succès des diatomées dans les océans contemporains par des approches génomique. La diatomée centrique *Thalassiosira pseudonana* et la diatomée pennée *Phaeodactylum tricornutum* ont été entièrement séquencé, ce qui révélé beaucoup d'informations sur les gènes et la structure du génome chez les diatomées. Les diatomées sont une partie de la branche stramenopile des eucaryotes qui est passé par des événements secondaires endosymbiotiques impliquant des algues vertes et rouges et a acquis des gènes bactériens par transfert horizontal de gènes (HGT) à partir de bactéries. L'origine différente des gènes intégrés dans le génome des diatomées fournit le "melting pot" du génome avec de nouveaux métabolismes, tels que le cycle de l'urée, qui était auparavant considéré comme existant seulement chez les animaux, mais dans la recherche *in Silico* et les données expérimentales ont montré que le cycle de l'urée existe dans les deux *T. pseudonana* et *P. tricornutum*. *P. tricornutum* est la première diatomée pennée entièrement séquencé avec un génome de petite taille (27Mo). Par rapport à d'autres diatomées, elle pousse facilement dans des conditions de laboratoire et plusieurs ressources moléculaire, tels que les bases de données EST (<http://www.diatomics.biologie.ens.fr/EST3/est3.php>) sont disponibles. Des outils de génétique inverse dans *P. tricornutum*, comme l'inactivation génique et la surexpression ont également été bien établis.

Bien que le séquençage du génome entier est essentiel pour une meilleure compréhension des diatomées et de leur succès écologique, la séquence primaire est seulement une base pour comprendre comment le programme génétique est lu. Une autre couche d'informations héréditaires qui est superposée à la séquence d'ADN est l'information épigénétique. Dans les dernières décennies, des progrès remarquables ont été accomplis dans le domaine de l'épigénétique chez les mammifères et les plantes ce qui a permis de mieux comprendre le rôle essentiel de la régulation épigénétique dans l'organisation de la chromatine, l'expression des gènes et le développement. J'ai étudié la caractérisation de l'épigénome chez la diatomée modèle *P. tricornutum* pendant ma thèse.

Le manuscrit de thèse est organisé en cinq chapitres. Le chapitre I est une introduction aux autres, car il donne un aperçu de ce qui est connu dans la littérature à propos de la méthylation, les modifications des histones, les petits ARN et de leurs interactions éventuelles dans les organismes modèles où ils ont été examinés, et ce que nous savons jusqu'à présent sur l'épigénétique de diatomées.

Chapitre II «La méthylation dans le modèle de diatomées pennées *Phaeodactylum tricornutum*» est principalement au sujet du profilage de méthylation de l'ADN dans *P. tricornutum*. Le document intitulé " Insights into the Role of Methylation in Diatoms by Genome-Wide Profiling in *Phaeodactylum tricornutum*" a été soumis pour publication. Dans le chapitre II, j'ai utilisé McrBC-ChIP pour le profilage de la méthylation de l'ADN à l'échelle du génome chez *P. tricornutum*. En général, *P. tricornutum* a une faible méthylation de l'ADN par rapport aux éléments transposables (TEs) largement méthylés et quelques gènes méthylés.

Dans le chapitre III intitulé «La cartographie de H3K4me2, H3K9me2 et H3K27me3 et leur corrélation avec méthylation de l'ADN et les petits ARN chez *Phaeodactylum tricornutum*», je discute les résultats des expériences ChIP-seq pour trois différentes modifications des histones, H3K4me2 H3K9me2 et H3K27me3 qui ont été menées pour étudier la distribution de ces histones à l'échelle du génome de *P. tricornutum*. ARN-seq profilage a été menée en parallèle pour corréler l'expression des gènes avec des profils de modification des histones. La présence in vivo de ces histones modifiées a été vérifiée par western blot. Les résultats montrent que H3K4me2 marque principalement les gènes tandis qu'une grande partie de H3K9me2 et H3K27me3 dépositions sont sur les TEs. Ainsi, H3K27me3 présente une évolution inhabituelle soulevant des questions sur son rôle dans *P. tricornutum* et si cette distribution particulière est conservée chez les diatomées et les organismes unicellulaires en général. Dans d'autres modèles bien étudiés, comme chez *A. thaliana* et *D. melanogaster*, H3K27me3 principalement se dépose sur les gènes réprimés pas les TEs. Une petite partie des zones marquées par H3K27me3 sont également gènes réprimés, ce qui est conforme aux études antérieures. Les annexes du chapitre III est un document soumis - "immunoprécipitation de la chromatine couplée à la détection par PCR quantitative en temps réel pour étudier in vivo les interactions protéines d'ADN chez deux diatomées modèle *Phaeodactylum tricornutum* et *Thalassiosira pseudonana*", qui décrit le protocole immunoprecipitation chromatine que j'ai utilisé pour ChIP-seq expériences.

Dans le chapitre IV "Approche de génétique inverse pour disséquer les mécanismes épigénétiques chez *Phaeodactylum tricornutum*", j'ai étudié les gènes putatifs codant pour des protéines impliquées dans la machinerie de méthylation d'ADN-C5, N6, les mécanismes de méthylation de l'adénine, la déméthylation et méthylation des histones avec un accent sur les protéines du groupe Polycomb (GPC) qui jouent un rôle crucial dans le dépôt de H3K27me3. Dans le chapitre IV, l'analyse fonctionnelle sur trois putatifs C-5 méthyltransférases (46156,

45072 et 47357) a été réalisée en mutant ces gènes dans *P. tricornutum* par RNAi. Les résultats préliminaires montrent que les niveaux globaux de méthylation de l'ADN chez les lignées mutantes 45072 et 47357 sont plus faibles que le type sauvage. J'ai étudié les fonctions de E putatif (z) qui est le composant clé de PcG par approche de génétique inverse. Les cellules de E (z) 32817 des *P. tricornutum* mutants ont tendance à être arrondie ovale ce qui est différent des cellules de type sauvage. Basé sur les résultats des expériences menées dans ma thèse, il semblerait que E (z) dans *P. tricornutum* pourrait être impliqué dans la détermination du sort de la cellule quant à son morphotype. J'ai aussi exploré la distribution des protéines PcG dans des algues unicellulaires dont trois espèces de diatomées. Il est connu que les protéines du groupe Polycomb sont associées à H3K27me3. La distribution originale de H3K27me3 dans *P. tricornutum* suscite mon intérêt pour étudier la distribution des protéines du groupe Polycomb dans les organismes unicellulaires. Ceci est d'un intérêt particulier parce que la marque H3K27me3 a été initialement considéré n'existant que dans les organismes multicellulaires. J'ai montré que les protéines PcG ont une distribution distincte chez *P. tricornutum* rapport à d'autres organismes multicellulaires bien étudiés. La diffusion à grande échelle des protéines du groupe Polycomb dans les algues unicellulaires en outre a confirmé le fait que non seulement elles existent dans les organismes unicellulaires, mais elles peuvent aussi avoir des fonctions différentes.

Le chapitre V "Discussion et perspectives», résume la thèse et donne un point de vue sur l'étude de l'épigénétique des diatomées et sa relation à l'environnement.

Dans cette thèse, j'ai exploré le monde épigénétique d'un organisme phytoplanctonique avec une importance évolutive et écologique pour nous donner l'occasion de comprendre comment un épigénome est organisé en une seule algue unicellulaire marine, quels sont les interactions entre les différentes composantes épigénétiques sont, et quelles sont les corrélations avec d'autres modèles bien étudiés. J'ai été un pionnier de la recherche de l'épigénétique chez les diatomées en enquêtant sur l'épigénome de *P. tricornutum*, en particulier les aspects de la distribution des différentes marques d'histones générées par ChIP-seq combiné avec la méthylation de l'ADN. Dans ma thèse, le chapitre II porte sur l'étude de la méthylation de l'ADN chez *P. tricornutum*. La méthylation de l'ADN est le composant le plus largement étudié dans le domaine de l'épigénétique. Elle est largement distribuée dans les différentes branches de la vie comme les mammifères et les plantes, mais essentiellement absente chez certaines espèces modèles comme *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster* et *Caenorhabditis elegans*. Feng et al Zilberman ont étendu la liste

des methylomes des espèces modèles à différentes espèces et ont étudié les caractéristiques de la distribution de la méthylation de l'ADN dans un contexte évolutif. Cependant, tous les methylomes de génomes entiers qui ont été signalés jusqu'à présent étaient seulement confinés aux Unikont et Archaeplastida. Il était donc intéressant d'étudier le methylome d'une diatomée dérivé d'un groupe eucaryote important, les straménopiles, et de l'importance écologique des diatomées ainsi que de l'évolution. Le chapitre II présente la cartographie de la méthylation à l'échelle du génome chez *P. tricornutum* dans des conditions normales en utilisant la technique McrBC-ChIP. En général, il a été trouvé que *P. tricornutum* a une méthylation de l'ADN faible avec des TEs méthylés de manière dense et quelques gènes méthylés. Les éléments transposables de type Copia ont tendance à être plus dénaturés par rapport à d'autres groupes d'ETs qui n'est pas surprenant car ces éléments sont susceptibles d'être actifs, et donc facilement mobilisable dans le génome. La méthylation de ces ETs peuvent contrôler leur mobilité ce qui permet de maintenir la stabilité du génome.

En outre, les données de l'ARN-seq ont permis d'évaluer la corrélation entre l'expression des gènes et la méthylation de l'ADN. Même si la méthylation du gène n'est pas significative chez *P. tricornutum* (seulement 3% des gènes) par rapport aux espèces de plantes modèles comme *Arabidopsis thaliana* (environ un tiers des gènes sont méthylés) et de mammifères, j'ai trouvé que les gènes avec une méthylation interne ont tendance à être modérément exprimés. Grâce la disponibilité des banques d'EST de 16 différentes conditions de culture, nous avons eu une bonne occasion d'étudier la corrélation entre la méthylation des gènes et leur expression dans différentes conditions. Un groupe de gènes méthylés qui ont méthylation lourde et complète attiré notre attention: ces gènes méthylés n'ont aucune expression dans la condition normale, mais certains d'entre eux d'exprimer sous certaines conditions spécifiques. En d'autres termes, ces gènes sont sous le contrôle étroit et seulement exprimer lorsque certains signaux de l'environnement se produisent. Je spécule que l'expression de ces gènes est probablement causée par déméthylation ou l'hypométhylation de ces gènes dans des conditions spécifiques qui doit être confirmé en vérifiant leur état de méthylation de l'ADN. Si c'est le cas, les gènes fortement et complètement méthylés montrent qu'ils sont contrôlés et régulés par l'hyperméthylation en réponse à des signaux environnementaux.

Il est intéressant de noter qu'*Ectocarpus siliculosus*, une autre straménopiles a été trouvée en analysant la séquence de son génome ne contenant pas les protéines codant pour l'ADN methyltransferases. Il a été également montré que la méthylation d'ADN n'était pas détectable en utilisant une analyse HPLC. Il sera intéressant à l'avenir d'étudier pourquoi ces deux

organismes étroitement liés sur le plan évolutif ont des modèles de méthylation de l'ADN distincts. Le méthylome de *P. tricornutum*, qui est le premier méthylome de straménopiles, contribuera par ailleurs à notre compréhension de l'évolution de la méthylation de l'ADN chez les eucaryotes.

À l'avenir, les méthylomes supplémentaires de différents organismes appartenant à différentes lignées peuvent aider à mieux comprendre le rôle de l'évolution de la méthylation de l'ADN qui est considéré comme une marque épigénétique ancienne. En outre, l'analyse de la méthylation de l'ADN à l'échelle du génome dans des conditions différentes mettra en lumière si la méthylation du gène peut dynamiquement réguler la transcription des gènes. Dans l'ensemble, mon travail a jeté les bases pour une enquête plus approfondie de la méthylation de l'ADN dans différentes conditions environnementales chez *P. tricornutum* mais aussi chez d'autres diatomées et algues.

Bien que la répartition de la méthylation de l'ADN à l'échelle du génome entier chez *P. tricornutum* a été profilée, la méthode ChIP-McrBC ne permet pas distinguer les différents contextes de méthylation de l'ADN: CG, CHG et CHH. Le séquençage après traitement au bisulfite peut générer des cartes à haute résolution de la méthylation de l'ADN et distinguer les différents contextes CG, CHG et CHH, est désormais considérée comme la méthode de référence pour l'étude de la méthylation de l'ADN. Initialement, les mammifères ont été considéré n'ayant que le contexte CG de méthylation de l'ADN. Mais récemment, dans les cellules souches humaines, d'autres contextes de méthylation ont été également été détectés. Chez les plantes, les trois contextes de méthylation de l'ADN ont été détectées sur TE, mais seulement la methylation CG se trouve sur les gènes. La méthylation CG est le seul contexte de la méthylation de l'ADN des gènes qui est conservée parmi tous les organismes où les méthylomes ont été examinés. Jusqu'à présent, seule la méthylation des gènes CHG a été retrouvé chez *Chlamydomonas*. À des fins de validation, j'ai effectué le séquençage bisulfite e séquençage sur 28 loci répartis uniformément sur le génome. Seulement la méthylation CG a été trouvé parmi tous les loci que j'ai examinés. Il semble donc que la méthylation CG est le contexte dominant de la méthylation dans l'ADN chez *P. tricornutum*. Cependant, parmi les loci choisis, la plupart d'entre eux sont des genes. Le séquençage bisulfite à l'échelle du génome devrait donc être effectué pour nous donner une meilleure idée du contexte de méthylation de l'ADN chez *P. tricornutum*. Cependant, le séquençage bisulfite a aussi ses inconvénients: il ne peut pas distinguer 5MC et 5MC-OH et parce que le C non méthylé est converti en U et devient finalement T, c'est un défi difficile à relever pour la cartographie des

régions contenant des séquences répétées. Ainsi, le séquençage bisulfite avec l'approche McrBC-ChIP peut compenser les inconvénients de l'autre et permet de fournir ainsi des profils de méthylation de l'ADN plus précis.

Un autre élément épigénétique important est la modification des histones. Des études à l'échelle du génome du paysage épigénomique ont été menées chez la levure, *C. elegans*, la drosophile, *Arabidopsis*, le riz, le maïs, et les cellules humaines. Toutes ces études montrent qu'il existe une complexité combinatoire relativement faible des marques de la chromatine. Chez *D. melanogaster*, l'analyse intégrative de la distribution des marques de la chromatine a révélé cinq types principaux de chromatine. Chez *Arabidopsis*, 12 points le long de la séquence génomique ont permis de définir quatre principaux états de chromatine: les gènes, les gènes actifs, des éléments répétés silencieux et des régions intergéniques. Les modifications des histones H3K4me₂, H3K9me₂ et H3K27me₃ ont été choisies pour l'étude du génome car H3K4me₂ est généralement associée aux gènes, alors que H3K9me₂ est associée aux TEs et H3K27me₃ est associé aux gènes réprimés chez les métazoaires. Chez *P. tricornutum*, nous avons constaté que ces trois modifications des histones couvrent près de la moitié du génome ce qui n'est pas le cas chez les métazoaires. Le taux de couverture élevé des trois marques des histones pourrait être dû au fait que le génome de *P. tricornutum* est beaucoup plus petit avec une taille plus courte du gène et régions intergéniques.

Comme prévu, H3K4me₂ marque les gènes, sans discrimination pour tout niveau d'expression génique alors que H3K9me₂ est principalement associée aux TEs chez *P. tricornutum*. A notre grande surprise, H3K27me₃, qui est étroitement associée à la protéine Polycomb, a une distribution nouvelle chez *P. tricornutum*: elle est principalement associée aux TEs et les gènes réprimés pas comme on l'observe dans d'autres organismes. H3K27me₃ et le groupe des protéines Polycomb ont été considérés seulement exister dans les organismes multicellulaires. Notre travail est le premier génome d'un organisme unicellulaire photosynthétique où la distribution de H3K27me₃ a été étudiée. Les résultats non seulement démontrent l'existence de H3K27me₃ chez *P. tricornutum*, mais a également soulevé la question de ses fonctions potentielles dans les organismes unicellulaires.

La majorité des régions marquées par H3K27me₃ sont des TEs, ce qui implique que H3K27me₃ peut probablement réguler l'activité des éléments transposables et peut être une marque des régions d'hétérochromatine facultative. A côté des TEs, certains gènes sont marqués par H3K27me₃. Ces gènes sont réprimés et exprimés de façon différentielle dans des

conditions différentes, une notion qui est cohérente avec les études précédentes chez les plantes et les métazoaires. Il semble que H3K27me3 chez *P. tricornutum* peut avoir une fonction similaire sur les gènes que l'on trouve dans les organismes multicellulaires. H3K27me3 est considéré comme associé au développement et à l'identité cellulaire chez les organismes multicellulaires. Bien que dans la seule espèce unicellulaires, la différenciation cellulaire n'existe pas, la transition d'un type de cellule à l'autre, comme gamétophyte et sporophyte est très commun. Il s'agit donc d'une question intéressante à approfondir dans l'étude des fonctions de H3K27me3 dans les organismes unicellulaires.

Un autre phénomène que j'ai trouvé intéressant est que H3K27me3 et H3K9me2 ont tendance à co-marquer les gènes fortement méthylés et les éléments transposables. La plupart des gènes méthylés qui sont fortement marqués par H3K9me2 et H3K27me3 ne sont pas exprimés sauf sous conditions particulières. Il semble que ces marques épigénétiques peuvent coopérer pour réguler les gènes et les TEs chez *P. tricornutum*. Je spécule que dans un tel petit génome, les marques des histones très probablement ont tendance à se chevaucher les unes avec les autres. Dans un tel génome aussi compact, les modèles de marques combinatoires plutôt que de simples marques épigénétiques peuvent servir comme un code épigénétique pour exercer de profondes répercussions sur les états de chromatine.

P. tricornutum non seulement possède des histones méthyltransférases sur lysine mais présente également d'autres enzymes putatives responsables d'autres modifications des histones, comme l'acétylation de lysine. Il sera très intéressant de mener des études de ChIP-seq afin d'explorer la distribution à l'échelle du génome entier des différentes modifications des histones. Des études récentes montrent que non seulement les modifications des histones mais aussi des variants d'histones peuvent influencer la transcription. H3 de *P. tricornutum* possède différentes variantes (Pt11841 H3-1a, 1b Pt50872 H3-H3 et Pt50695-1c). Dans les études futures d'épigénétique chez *P. tricornutum*, les variants d'histones devrait donc également être pris en considération.

Bien que le génome de *P. tricornutum* est petit, il contient des systèmes complexes de régulation épigénétique qui font défaut dans la levure bourgeonnante, tels que la méthylation de l'ADN et les protéines Polycomb. À la lumière du fait que les espèces modèles classiques telles que *S. cerevisiae*, *D. melanogaster* et *C. elegans* n'ont pas ou très faible méthylation de l'ADN, la corrélation des modifications des histones et la méthylation de l'ADN, ne peut être évaluée. Cependant, *P. tricornutum*, en dépit d'être un organisme unicellulaire affiche des

caractéristiques épigénétiques complexes. Etre polymorphique encore plus attrayant pour les études épigénétiques. Je souhaite donc à l'avenir de voir son émergence en tant que modèle pour les études épigénétiques dans les organismes unicellulaires. Une étape cruciale dans le futur proche est d'effectuer des analyses de spectrométrie de masse des nucléosomes chez *P. tricornutum* afin de déterminer tous les types de modifications des histones existants chez *P. tricornutum*.

La distribution peu orthodoxe de H3K27me3, qui est étroitement associée à des protéines du groupe Polycomb, pose la question de leur rôle dans les organismes unicellulaires dont il est question dans le chapitre suivant. Dans le chapitre IV, j'ai étudié plus en détail les protéines Polycomb chez plusieurs espèces en m'appuyant sur des analyses génomiques comparatives. Les résultats ont confirmé que le système PcG n'existe pas seulement chez les organismes multicellulaires, mais aussi dans les organismes unicellulaires parce que toutes les algues unicellulaires eucaryotes que j'ai examinées contiennent des composants Prc2 putatifs. En outre, j'ai exploré la fonction de PcG par mutation du gène homologue de E(z) (32 817) dans *P. tricornutum*. Les résultats préliminaires ont montré que les lignées mutantes de *P. tricornutum* Ez (32 817) ont diminué par rapport aux niveaux H3K27me3 de type sauvage. Il est intéressant d'établir le profil de la distribution sur l'ensemble du génome de H3K27me3 chez les lignées mutantes. En outre, le profilage d'autres modifications de la chromatine dans les lignées mutantes de E(z) tels que la méthylation de l'ADN et H3K4me3 qui antagonise H3K27me3, peut aider à mieux comprendre les interactions entre les différentes modifications de la chromatine chez *P. tricornutum*.

J'ai également étudié les fonctions des ADN méthyltransférases putatifs chez *P. tricornutum* par la méthode de manipulation génétique inverse. Selon les résultats préliminaires, deux méthyltransférases putatives (Pt45072 et Pt47357) après mutagenèse ont montré un niveau de méthylation de l'ADN du génome diminué par rapport au type sauvage. À l'avenir, il sera utile d'explorer la distribution à échelle du génome entier de la méthylation de l'ADN dans ces mutants par du séquençage bisulfite et de continuer à disséquer les fonctions de ces ADN méthyltransférases. Cela permettra de fournir des réponses sur le contexte de la méthylation de l'ADN. Dans son ensemble, mon travail est la première étude de l'épigénome d'une Stramenopile, a montré de nouvelles caractéristiques qui révèlent certains mécanismes conservés et différents par rapport à d'autres eucaryotes. J'espère que mon travail sur l'épigénome de *P. tricornutum* sera la base pour des études futures sur l'épigénétique des

diatomées et une grande opportunité pour l'épigénomique comparatifs et l'élucidation des relations entre la dynamique de l'évolution du génome et la régulation épigénétique.

Comme mon travail de thèse le démontre, l'épigénétique chez les diatomées est un sujet de recherche fascinant qui est susceptible de conduire à de nouvelles perspectives quant à son rôle au cours de l'évolution. J'espère que mes études ouvrent la voie à de nouvelles recherches visant à explorer des phénomènes épigénétiques dans un contexte environnemental.