# Phylogénie moléculaire de polychètes Alvinellidae des sources hydrothermales actives de l'océan Pacifique

Molecular phylogeny of the active Pacific Ocean hydrothermal vents alvinellid Polychaetes

Jean-Pierre Féral  $^{(1)}$ , Hervé Philippe  $^{(2)}$ , Daniel Desbruyères  $^{(3)}$ , Lucien Laubier  $^{(4)}$ , Évelyne Derelle  $^{(1)}$ , Anne Chenuil  $^{(5, 6)}$ 

- (1) Observatoire Océanologique de Banyuls, URA CNRS 117, 66650 Banyuls-sur-Mer, France.
- (2) Laboratoire de Biologie Cellulaire 4, Université Paris-Sud, URA CNRS 1134, 91405 Orsay Cedex, France.
- (3) IFREMER, Centre de Brest, Département Environnement profond, BP 70, 29280 Plouzané, France.
- (4) Laboratoire de Physiologie des Êtres marins, Institut Océanographique, 195, rue Saint-Jacques, 75005 Paris, France.
- (5) Laboratoire de Biologie et Génétique Évolutives, CNRS, 91198 Gif-sur-Yvette Cedex, France.
- (6) Adresse actuelle: Génome et populations, URA CNRS 1493, Université Montpellier II, CC 63, 34095 Montpellier Cedex 5, France.

#### RÉSUMÉ

Des arbres phylogénétiques ont été construits par les méthodes de distances et de parcimonie, à partir des séquences alignées de l'extrémité 5' de l'ARNr 28S de 17 espèces de polychètes. Les Alvinellidae forment un taxon monophylétique, au sein du groupe également monophylétique des Terebellomorpha. À l'intérieur de la famille des Alvinellidae, le genre Alvinella (2/2 espèces étudiées) est un taxon monophylétique. De même, les espèces étudiées de Paralvinella (3/9 espèces) forment un groupe monophylétique. L'allongement de 30 nucléotides du domaine D<sub>1</sub> chez Alvinella s'explique par des insertions en blocs de 3 à 11 nucléotides dont la fixation n'a vraisemblablement pas beaucoup changé la conformation de l'ARNr. Cette fixation a pu être favorisée par la nécessité de colonisations rapides de nouveaux sites hydrothermaux actifs (isolement spatial des sources et courte période d'activité). La comparaison des séquences des polychètes étudiées avec celles d'invertébrés marins côtiers tempérés ou polaires montre que la richesse relative en cytosine-guanosine (C+G) du domaine D<sub>1</sub> des Alvinellidae n'a pas de lien direct avec les conditions particulières du milieu hydrothermal, infirmant une hypothèse récemment avancée.

**Mots clés**: phylogénie, ARNr 28S, sources hydrothermales, océan Pacifique, polychète, Terebellomorpha, Alvinellidae, Alvinella, Paralvinella.

#### ABSTRACT

Phylogenetic trees were inferred by both distance and parsimony methods, using aligned sequences of the 28S rRNA 5'-end of 17 Polychaete species. Alvinellidae is a monophyletic taxon among the Terebellomorpha which also make a monophyletic group. Alvinella (2/2 species studied) is also a monophyletic taxon. The Paralvinella species (3/9 species) we studied cluster also in a monophyletic group. The  $D_1$  domain of Alvinella is longer than in other Polychaetes (30 nucleotides more). This is explained by the insertions of groups of 3 to 11 nucleotides whose fixation probably did not change so much rRNA conformation. This fixation could have been promoted by the necessity of fast recolonizations of new active hydrothermal vents (spatial isolation and ephemeral habitat). The comparison of the studied sequences of Polychaetes with those of coastal temperate and polar marine

Note présentée par Lucien Laubier.

Note remise le 6 juin 1994, acceptée après révision le 13 juillet 1994.

Correspondance : J.-P. Féral.

invertebrates shows that the relatively high Cytosine-Guanosine (C+G) level found in alvinellid sequences is not directly related to the particular characteristics of the hydrothermal environment as recently hypothetized.  $\blacktriangle$ 

**Key words**: phylogeny, 28S rRNA, hydrothermal vents, Pacific Ocean, Polychaeta, Terebellomorpha, Alvinellidae, *Alvinella, Paralvinella*.

Abridged version (see p. 777)

es Alvinellidae ou « vers de Pompéi » sont une composante majeure des communautés hydrothermales profondes des dorsales océaniques et des bassins arrière-arc du Pacifique. On observe ces Annélides Polychètes aussi bien sur les parois des cheminées à la lisière du fluide hydrothermal où la température est comprise entre 20 et 80 °C (Alvinella spp, Paralvinella sulfincola) que vivant en épibiontes sur les tubes de Vestimentifères (Paralvinella grasslei, P. palmiformis, P. pandorae) ou encore dans les sédiments riches en hydrocarbures et couverts de mattes bactériennes [1-5]. L'habitat des Alvinellidae, discontinu dans le temps et l'espace, se situe à l'interface turbulente créée par le mélange entre le fluide hydrothermal acide, anoxique et surchauffé (> 300 °C) et l'eau de mer froide (2 °C) et oxygénée. Le biotope hydrothermal est localement de taille réduite ; dans la majorité des cas, les communautés couvrent des surfaces inférieures à 2 500 m<sup>2</sup>. L'habitat des Alvinellidae est donc isolé, mais aussi éphémère, l'activité d'un fumeur étant estimée à 10 à 50 ans [3, 6, 7].

Actuellement deux genres sont reconnus sur des critères morphologiques par les taxonomistes : Alvinella, qui compte 2 espèces sympatriques de la ride du Pacifique oriental, et Paralvinella dont 9 espèces ont été décrites et qui est largement distribué dans le Pacifique. Ces organismes, d'abord regroupés dans une sous-famille nouvelle d'Ampharetidae, les Alvinellinae [1], sont actuellement considérés comme formant une famille de Terebellomorpha, les Alvinellidae [2]. Ils partagent des caractères plésiomorphes avec les Terebellidae et les Ampharetidae, ainsi que des caractères apomorphes avec les Trichobranchidae [3]. Ainsi, la présence d'un uncinus aviculaire ayant un processus subrostral réduit est un caractère qui rapproche les Alvinellidae des Trichobranchidae, alors que la présence de tentacules buccaux rétractiles dans la cavité buccale est un caractère d'Ampharetidae. La présence de notopodes sur tous les segments du corps est un caractère clairement plésiomorphe chez les Terebellomorpha qui n'a jamais été signalé chez les Trichobranchidae ou les Ampharetidae actuels [3].

Nous avons construit des arbres phylogénétiques, fondés sur des séquences d'ARN ribosomique 28S. Cet ARN,

par sa structure en mosaïque de domaines conservés et variables, est un puissant indicateur phylogénétique [8]. La possibilité d'utiliser des séquences partielles d'ARNr, ou d'ADNr, pour aborder des problèmes à différents niveaux phylogénétiques a déjà été montrée chez les annélides [9-12]. Nous avons choisi d'étudier l'extrémité 5' de l'ARN de la grande sous-unité du ribosome (domaine divergent D<sub>1</sub> et régions adjacentes C<sub>1</sub> et C<sub>2</sub>) [13] qui contient, chez les Annélides, un signal phylogénétique suffisamment puissant pour des comparaisons à relativement courte distance taxonomique [10, 11], les domaines les plus divergents étant trop variables pour être alignés sur toute leur longueur [11].

Par ailleurs, Dixon *et al.* [9] ont émis l'hypothèse de l'existence d'une relation entre un taux élevé de cytosine-guanosine (C+G) chez *Paralvinella* et une résistance au milieu hydrothermal (en particulier haute température). Nous avons donc analysé la composition en bases des séquences de plusieurs espèces d'annélides vivant dans les sites hydrothermaux et dans des milieux aux caractéristiques différentes.

# Matériel et méthodes

Les 17 espèces utilisées figurent dans le *Tableau I*, avec leur origine. Notre échantillon comporte des espèces représentant les deux genres connus d'Alvinellidae et des Terebellomorpha non hydrothermaux. Un groupe externe de non-Terebellomorpha a été constitué en fonction de la disponibilité des séquences des domaines pris en compte dans cette étude. Les ARN totaux ont été extraits d'individus congelés ou frais, et séquencés par la méthode de Qu *et al.* [8, 14]. Les amorces utilisées correspondent aux positions 393-412 et 51-70 de l'ARNr 28S de la séquence de la souris [13]. L'extrémité 5' des séquences, ainsi que le domaine D<sub>1</sub> et les secteurs adjacents (domaine C<sub>1</sub> et une partie du domaine C<sub>2</sub>) ont été alignés. L'ensemble est homologue des positions 1 à 392 de la séquence de la souris [13].

Des arbres ont été construits en utilisant les séquences alignées des domaines conservés et/ou variables (homologie avec les positions 3 à 356 de la souris) au moyen de logiciels utilisant les méthodes de distances ou de parcimonie (PHYLIP [15], MUST [16]). Les régions

Figure 1. Alignement (5'-3') des séquences du domaine  $D_1$  et des domaines conservés adjacents  $C_1$  et  $C_2$  (partiel) de l'ARNr 285  $\triangleright$  d'annélides polychètes. Les positions sont comptées à partir de l'extrémité 5'. Les nucléotides qui diffèrent de ceux de la première ligne (Terebellides stroemi) sont les seuls indiqués par une lettre. Les autres, identiques, sont notés par un tiret. Les délétions sont marquées par des astérisques. Le code standard I.U.B. est utilisé pour les nucléotides incomplètement identifiés. Les régions hypervariables, non prises en compte dans les constructions phylogénétiques, sont surlignées. Les zones d'insertions par blocs extérieures à ces régions sont repérées par des barres obliques ; elles sont omises dans certaines analyses.

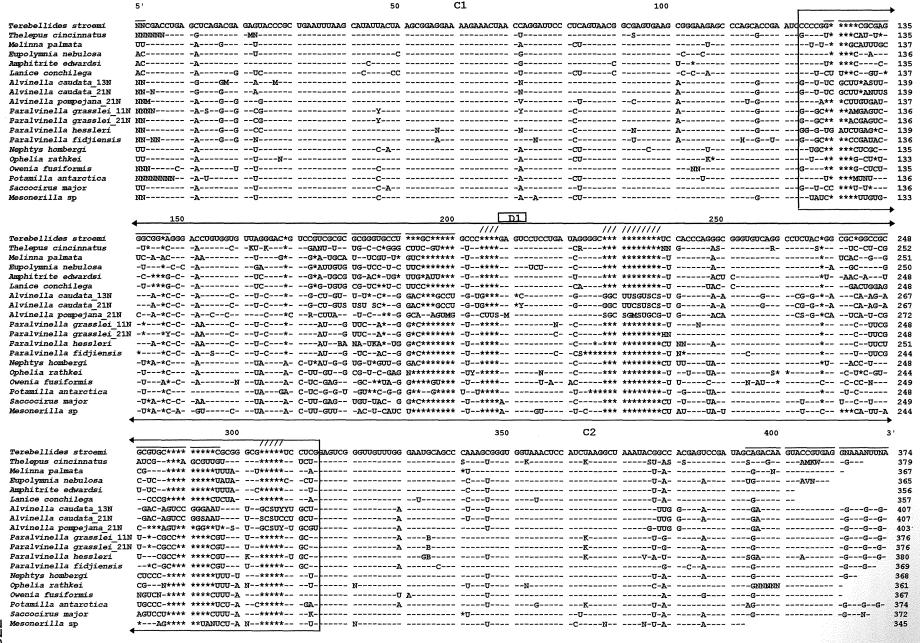


Tableau I Origine des spécimens et des séquences utilisés

Taxon	Famille	Lieu de récolte	prélèvement	Collecteur	Identification
Terebellides stroemi Sars, 1835	Trichobranchidae	Méditerranée occidentale	plongée scaphandre	JC. Roca	JC. Duchêne
<i>Potamilla antarctica</i> Kinberg, 1867	Sabellaridae	Antarctique, Terre-Adélie	plongée scaphandre	JP. Féral	JC. Duchêne
<i>Thelepus cincinnatus</i> Fabricius, 1780	Terebellidae	Antarctique, Terre-Adélie	plongée scaphandre	JP. Féral	JC. Duchêne
<i>Paralvinella hessleri</i> Desbruyères et Laubier 1989	Alvinellidae	Bassin arrière- arc des Mariannes	Alvin (Marianna et BAB exp.)	R.R. Hessler	D. Desbruyères
<i>Paralvinella grasslei</i> Desbruyères et Laubier 1982	Alvinellidae	Dorsale Est- Pacifique 11° N	Alvin (MVT)	D. Desbruyères	D. Desbruyères
<i>Paralvinella grasslei</i> Desbruyères et Laubier 1982	Alvinellidae	Dorsale Est- Pacifique 21°N	Alvin (MVT)	D. Desbruyères	D. Desbruyères
<i>Paralvinella fidjiensis</i> Desbruyères et Laubier 1993	Alvinellidae	Bassin nord-fidjien	Nautile (STARMER)	D. Desbruyères	D. Desbruyères
<i>Alvinella pompejana</i> Desbruyères et Laubier 1980	Alvinellidae	Dorsale Est- Pacifique 21°N	Alvin (MVT)	D. Desbruyères	D. Desbruyères
<i>Alvinella caudata</i> Desbruyères et Laubier 1986	Alvinellidae	Dorsale Est- Pacifique 13°N	Alvin (MVT)	D. Desbruyères	D. Desbruyères
<i>Alvinella caudata</i> Desbruyères et Laubier 1986	Alvinellidae	Dorsale Est- Pacifique 21°N	Alvin (MVT)	D. Desbruyères	D. Desbruyères

ARNr séquencé dans le cadre du présent article (O. O. Banyuls-sur-Mer).

Taxon	Famille	Lieu de récolte	Origine	
Owenia fusiformis della Chiaje, 1841	Oweniidae	Manche	[11]	
Ophelia rathkei McIntosh, 1908	Opheliidae	Manche	[11]	
Nephtys hombergi Savigny, 1818	Nephtyidae	Manche	[11]	
Saccocirus major Pierantoni, 1907	Saccocirridae	Manche	[11]	
Mesonerilla sp.	Nerillidae	Manche	[11]	
Melinna palmata Grube, 1870	Ampharetidae	Manche	[11]	
Eupolymnia nebulosa Montagu, 1818	Terebellidae	Méditerranée occidentale	[10]	
Lanice conchilega Pallas, 1766	Terebellidae	Méditerranée occidentale	[10]	
Amphitrite edwardsi Quatrefages, 1865	Terebellidae	Manche	[10]	

Séquences publiées.

hypervariables, surlignées, quasi inalignables, n'ont pas été utilisées dans les analyses (Fig. 1). Des arbres ont également été construits en ne tenant pas compte des insertions en bloc qui caractérisent surtout les Alvinella. La robustesse des nœuds a été mesurée par la méthode du bootstrap [17].

#### Résultats

# Analyse des séquences

Si les domaines conservés  $C_1$  et  $C_2$  sont facilement alignables, le domaine  $D_1$  des annélides est très variable (*Fig. 1*). Les séquences des espèces étudiées ne présentent environ que 16 % de sites constants. Le domaine  $D_1$  s'aligne sur 191 positions pour une moyenne de 147 nucléotides. Le  $D_1$  de Terebellomorpha non hydrothermaux possède  $148 \pm 4$  ( $\pm$  intervalle de

confiance au risque  $\alpha=0,05$ ) nucléotides, celui des annélides non Terebellomorpha  $146\pm 2$  nucléotides. Les  $D_1$  de *Paralvinella grasslei* et *P. hessleri*, formés respectivement de 147 et 151 nucléotides, sont compris dans cet intervalle. *Paralvinella fidjiensis* a un  $D_1$  un peu plus court (142 nucléotides), mais cette valeur se situe entre les extrêmes observés. En revanche, les deux espèces du genre *Alvinella* ont un  $D_1$  significativement plus long d'une trentaine de nucléotides (176 et 178 nucléotides). Cet allongement est dû en particulier à des insertions par blocs de 3 à 11 nucléotides (*Fig. 1*). Les différents spécimens d'*Alvinella caudata*, comme ceux de *Paralvinella grasslei*, récoltés à des stations éloignées de plus de 1 000 km ( $11-13^\circ$ N et  $23^\circ$ N) ont des séquences virtuellement identiques, aux ambiguïtés de séquençage près (*Fig. 1*).

Le taux de (C+G) est de 47 à 54 % dans les parties conservées de l'ARNr des Alvinellidae comme dans celles des autres annélides étudiées (Tableau II). Ce taux augmente dans le domaine divergent  $D_1$  (Tableau II). Il est de  $70 \pm 4$  % ( $\alpha = 0,05$ ) en moyenne chez les Alvinellidae (valeurs extrêmes, 67 à 74 %) et de  $68 \pm 4$ % chez les Terebellomorpha tempérés (61 à 77 %) ou chez le Terebellomorpha antarctique (68 %). Chez Potamilla antarctica, annélide antarctique non Terebellomorpha, le taux de (C+G) observé dans  $D_1$  est également compris dans cet intervalle (Tableau II).

### Analyse phylogénétique

La topologie des arbres de distances et de parcimonie est identique, que l'on prenne ou non en compte les insertions par blocs (positions 205-208, 228-238, 304-308) chez *Alvinella (Fig. 1)*. Les longueurs de branches restent semblables dans les arbres de distances (FITCH et NJ), sauf celles qui portent les *Alvinella*, qui bien sûr s'allongent si on utilise ces sites (*Fig. 2*). La topologie est indépendante de la méthode choisie.

Les Alvinellidae apparaissent comme un groupe monophylétique au sein des Terebellomorpha, eux-mêmes monophylétiques (Fig. 2). Ce dernier nœud est soutenu par une valeur de bootstrap de 70 % aussi bien en utilisant une méthode de distances (NJBOOT) que de parcimonie (DNABOOT). En incluant les insertions par blocs, le regroupement des Alvinellidae est soutenu par des valeurs de bootstrap de 90 % en distances et en parcimonie (Fig. 2). Ces valeurs sont respectivement 90 % et 87 % quand on supprime les blocs d'insertion de l'analyse. À l'intérieur de ce groupe, 2 taxa s'individualisent avec des valeurs de bootstrap élevées (100 % par les méthodes

Tableau II

Taux de cytosine-guanine (C + G) dans les domaines séquencés de l'ARN ribosomique 28S d'annélides

	% de (C+G)		
Espèces et origines	C <sub>1</sub>	D <sub>1</sub>	C <sub>2</sub> (partiel)
Terebellides stroemi (Méditerranée)	51	77	50
Thelepus cincinnatus (Antarctique)	50	68	48
Melinna palmata (Manche) Eupolymnia nebulosa	48	61	53
(Méditerranée)	51	67	48
Amphitrite edwarsi (Manche)	50	62	50
Lanice conchilega (Méditerranée)	52	65	51
Alvinella caudata (DEP)	50	68	48
Alvinella pompejana (DEP)	50	67	47
Paralvinella grasslei (DEP)	54	72	50
Paralvinella hessleri (Mariannes)	53	68	47
Paralvinella fidjiensis (Fidji)	50	74	53
Nephtys hombergi (Manche)		63	51
Ophelia rathkei (Manche)		63	53
Owenia fusiformis (Manche)		67	52
Potamilla antarctica (Antarctique)		68	49
Saccocirus major (Manche)	47	61	50
Mesonerilla sp. (Manche)	50	53	50

(DEP) = dorsale Est-Pacifique.

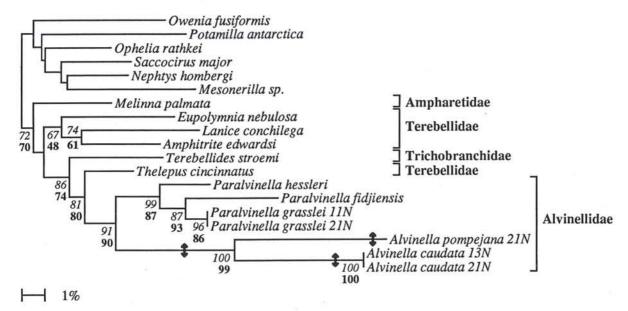


Figure 2. Arbre construit à partir d'une matrice de distances (programme FITCH du logiciel Phylip) résultant de l'alignement de la Figure 1. 319/429 positions sont utilisées. 139 sont variables. Le chiffre en caractère italique à chaque nœud indique le nombre de fois (en %) où le groupe d'espèces situé à sa droite est apparu au cours de 1000 tirages de bootstrap (programme NJBOOT du logiciel MUST). Le chiffre en gras donne cette valeur dans le cas où une méthode de parcimonie a été utilisée pour construire la phylogénie (programme DNABOOT du logiciel PHYLIP avec 89 sites informatifs). La distance entre deux espèces est proportionnelle à la somme des projections des longueurs des branches sur l'axe des abcisses. L'unité sur l'axe des ordonnées est arbitraire. L'arbre est raciné sur le groupe externe. Les flèches doubles indiquent la longueur des « branches Alvinella » de l'arbre construit sans les zones d'insertions par blocs (voir Fig. 1).

de distances et 99 % par les méthodes de parcimonie) : *Alvinella* qui est un taxon monophylétique et *Paralvinella* pour lequel les trois espèces étudiées forment aussi un groupe monophylétique.

Les autres Terebellomorpha sont regroupés en fonction des familles reconnues par les morphologistes, à l'exception de *Terebellides stroemi*, un Trichobranchidae qui se trouve parmi les représentants des Terebellidae (*Eupolymnia nebulosa, Amphitrite edwardsi, Lanice conchilega* et *Thelepus cincinnatus*). Cependant, les valeurs de bootstrap à la base de ce groupe Terebellidae/Trichobranchidae sont faibles ( $\leq 50\%$ ). Selon cette topologie, il n'est pas possible de déterminer avec certitude si les Terebellidae sont le groupe frère des Alvinellidae. En revanche, les Ampharetidae (*Melinna palmata*) seraient les Terebellomorpha qui en seraient les plus éloignés.

## Discussion

#### Variabilité du domaine D<sub>1</sub>

Le domaine D<sub>1</sub> est très variable chez les Alvinellidae, comme chez les autres annélides [11]. Certains secteurs n'ont pas pu être alignés avec suffisamment de sécurité pour être utilisés pour construire les arbres. La longueur de D<sub>1</sub> est à peu près constante, environ 150 nucléotides, chez les Terebellomorpha étudiés à l'exception des *Alvinella* chez lesquelles le domaine D<sub>1</sub> possède environ 30 nucléotides supplémentaires. Ils sont insérés par blocs de 3 à 11 nucléotides (*Fig. 1*) et correspondent à des événements d'insertion par bloc de nucléotides, plutôt qu'à des événements successifs d'insertions ponctuelles. La validité de cette interprétation est soutenue par le fait que ces blocs sont souvent situés dans des

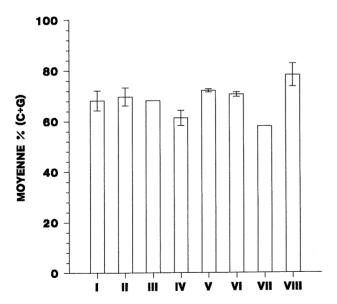


Figure 3. Représentation des moyennes des pourcentages de cytosine-guanosine (C+G) dans le domaine  $D_1$  des annélides polychètes prises en compte dans le présent travail et chez d'autres organismes. La barre verticale représente l'intervalle de confiance :  $\bar{x} \pm t \frac{\alpha}{2}$ .  $s_{\bar{x}}$ , au risque  $\alpha = 0.05$ , avec  $s_{\bar{x}} = \frac{s_x}{\sqrt{n}}$ . Terebellomorpha tempérés (I), hydrothermaux (II), antarctique (III). Autres polychètes (IV). Echinoïdes tempérés (V), antarctiques (VI). Mytilidae hydrothermal (VII). Vertébrés (VIII).

régions relativement bien conservées. Ce changement important de structure primaire se produit dans des régions qui ont, selon toute vraisemblance, peu d'importance fonctionnelle [18]. Les insertions en bloc se situent à la base de la boucle terminale du domaine D<sub>1</sub> [19]. Elles devraient être accompagnées de changements compensés dans une autre partie de la molécule d'ARNr, ou plus vraisemblablement, dans une ou plusieurs protéines ribosomales. Ce phénomène ne concerne qu'un des deux genres d'Alvinellidae. L'augmentation du nombre de nucléotides dans D<sub>1</sub> n'indique sans doute pas, par ellemême, un temps de divergence ancien. Il est possible de faire l'hypothèse d'un événement qui a provoqué la séparation d'une population à l'origine des Alvinella. Un facteur explicatif de la fixation des insertions serait l'effet de fondation induit par la nécessité de coloniser de nouvelles (jeunes) sources hydrothermales avant que leur population d'origine ne disparaisse avec l'extinction des fumeurs. Dans ce cadre, l'importance d'une mutagenèse induite par les hauts niveaux de radioactivité naturelle mesurés dans l'environnement hydrothermal et dans les organismes eux-mêmes [20] devrait aussi être éprouvée. Notons que de telles insertions n'existent pas chez Paralvinella, dont l'habitat diffère sensiblement de celui d'Alvinella (éloignement relatif des sorties chaudes et écran constitué par l'eau de mer).

Malgré cette grande variabilité, des individus récoltés dans des sites éloignés de plus de 1 000 kilomètres apparaissent identiques quand on compare la séquence de leur ARNr. Cela est en accord avec les résultats de l'analyse des distances génétiques de populations d'Alvinellidae qui montrent qu'il existe un flux génique entre les différentes populations, en particulier chez *Paralvinella grasslei* [21]. Cependant, le petit nombre de données existant sur le développement des Alvinellidae et sur leur dispersion n'aident pas à comprendre comment une espèce se déplace d'une source hydrothermale à l'autre sur des distances de quelques dizaines de mètres à quelques kilomètres. Comment ont-ils colonisé l'environnement linéaire, mais discontinu, sur plusieurs milliers de kilomètres que représentent les rides océaniques ? Cela se produit malgré la possibilité d'une pseudocopulation et d'une fécondation interne [22], ainsi qu'une flottabilité négative des œufs et un développement à phase larvaire raccourcie, sinon inexistante, et donc un recrutement probable près des adultes [23]. La dissémination de source en source peut faire l'objet d'hypothèse telle que l'épizoochorie, dont le support pourrait être des organismes non inféodés strictement à l'environnement hydrothermal [24, 25]. Une autre hypothèse pourrait faire intervenir le transport de stades larvaires ou juvéniles par les courants canalisés le long du rift médian de la dorsale.

# Taux de cytosine-guanosine (C+G)

Bien que l'hypothèse en ait été faite récemment, il semble peu probable qu'une relation directe existe entre la composition en (C+G) et les facteurs environnementaux [9], y compris ceux des sources hydrothermales dont le milieu est très réducteur et à haute teneur en métaux, et surtout où la température est élevée (20 à 80 °C et même 105 °C occasionnellement dans le cas des Alvinellidae [6]). Les domaines conservés C<sub>1</sub> et C<sub>2</sub> de l'ARNr d'Annélide, contraints par leur structure primaire,

ont un taux de (C+G) comparable à celui observé chez d'autres organismes (environ 50 %, voir Tableau II). Le taux de (C+G) du domaine D<sub>1</sub> des annélides étudiées est plus élevé et beaucoup plus variable (53 à 77 %). Chez les Terebellomorpha, les valeurs extrêmes concernent des espèces récoltées dans des mers tempérées, Melinna palmata (61 %) et Terebellides stroemi (77 %). Le taux de (C+G) des Alvinellidae, bien qu'élevé (68 à 74 %) est compris dans la variabilité de la richesse en (C+G) des Terebellomorpha (Tableau II et Fig. 3). Chez Potamilla antarctica, une espèce récoltée en Terre-Adélie, ce taux est de 68 %. Bien qu'il ne s'agisse ici que de résultats concernant le seul domaine D<sub>1</sub>, notre point de vue est confirmé par un taux de (C+G) de 58 % du domaine D<sub>1</sub> chez un Mytilidae des sources hydrothermales, Bathymodiolus sp récolté aux Fidji. La teneur en (C+G) d'échinoïdes tempérés (72,1  $\pm$  0,6 %,  $\alpha = 0.05$ ) ou antarctiques (70.6  $\pm$  0.9 %) que nous avons séquencés [données non publiées] va également dans ce sens (Fig. 3). Rappelons les valeurs de 81 % observées dans le domaine D<sub>1</sub> de xénope, vertébré pœcilotherme, et de 76 et 78 % chez celui de souris et d'homme, vertébrés homéothermes (Fig. 3). Les différences entre Terebellomorpha hydrothermaux, tempérés et antarctiques ne sont pas significatives. Les autres annélides, échantillon relativement petit et hétérogène, ont un taux de (C+G) qui se situe au voisinage de la limite inférieure du seuil de signification chez les Terebellomorpha.

La Figure 3 illustre la grande homogénéité des taux de (C+G) à l'intérieur des groupes monophylétiques (Terebellomorpha : I, II et III ; Echinoïdes : V et VI ; Vertébrés : VIII) alors même que les espèces composant ces groupes habitent dans des environnements très différents. La contrainte historique, due au taux de (C+G) de l'ancêtre commun et à la faible vitesse d'évolution des gènes ribosomiques, explique cette homogénéité. Les différences entre les groupes (par exemple Terebellomorpha versus Mytilidae) ne peuvent pas être dues aux caractéristiques des milieux où vivent les espèces.

Il serait par ailleurs nécessaire d'étudier aussi l'impact des « caractéristiques environnementales du gène » étudié

# ,

ABRIDGED VERSION

A lvinellid Polychaetes are strictly associated with deep sea active hydrothermal vents in the Pacific Ocean. Taxonomists consider alvinellids as a family of Terebellomorpha. They recognize 2 genera: Alvinella (2 species) and Paralvinella (9 species) [1-5]. These annelids live in waters between 20 to 80°C, while the abyssal surroundings are only at 1 or 2°C. This habitat is isolated. It is also ephemeral considering that the activity of a vent lasts only 10 to 50 years [3, 6, 7].

The aim of the present work is to understand the phylogenetic relationships among vent annelids and between vent annelids and annelids living in other environments (*Table I*) by means of the analysis of partial 28S ribosomal RNA sequences. It is also to improve the recent hypothesis that Cytosine-Guanosine (C+G) enrichment of alvinellid rDNA is an adaptation to extreme thermal environment [9].

Sequences of the 5'-end, including the  $D_1$  divergent domain and the conserved flanking regions  $C_1$  and  $C_2$  of the 28S rRNA of

(contraintes de l'environnement génomique [26]). Il a en effet été montré, aussi bien chez les vertébrés que chez les plantes [27-29], qu'un gène donné avait tendance à avoir un taux de (C+G) proche de celui de l'isochore auquel il appartient.

#### Place des Alvinellidae par rapport aux autres Terebellomorpha

La comparaison des séquences du domaine D<sub>1</sub> et des domaines conservés contigus (extrémité 5' de l'ARN ribosomique de la grande sous-unité) a permis de mettre en évidence la monophylie des Alvinellidae, d'Alvinella et de Paralvinella. La robustesse de ce résultat est soutenue par des valeurs de bootstrap supérieures à 90 %. Néanmoins, plusieurs problèmes subsistent. (1) L'accélération de la vitesse d'évolution, dans les deux groupes d'Alvinellidae, traduite par la longueur importante des branches correspondantes (Fig. 2), a pu créer, au cours de l'analyse, une association artéfactuelle de taxons à l'extrémité de ces grandes branches [29]. (2) Toutes les espèces de Paralvinella ne sont pas représentées dans notre analyse et il n'est pas possible d'exclure la paraphylie de ce groupe. (3) Le petit nombre d'espèces formant un groupe externe proche (Thelepus et Terebellides) réduit aussi la fiabilité des inférences phylogénétiques [30].

À l'exception de *Terebellides stroemi*, l'échantillon de Terebellomorpha utilisé dans ce travail se regroupe selon les familles dans lesquelles les morphologistes les classent (*Fig. 2*). Cette situation, qui n'est peut-être qu'un artefact dû à la présence d'un seul Trichobranchidea (*T. stroemi*) dans notre analyse, ne permet pas d'affirmer laquelle des familles, Terebellidae ou Trichobranchidae, est le groupe frère des Alvinellidae.

D'autres espèces de *Paralvinella* vivent dans les sites hydrothermaux de la dorsale de Juan de Fuca qui prolonge, vers le nord, la dorsale Est-Pacifique. La dérive sud-ouest de la plaque d'Amérique du Nord a détruit la plaque de Kula et scindé la plaque de Farallon, provoquant l'interruption de la dorsale Est-Pacifique, il y aurait environ 25 millions d'années (Oligocène) [31]. Ces espèces permettrajent de dater certaines branches de notre arbre.  $\blacksquare$ 

5 alvinellid species, 2 other Terebellomorpha and 1 sabellarid annelid species (*Table I*) were analysed after direct heterologous sequencing [8, 13, 14]. Other sequences (3 Terebellomorpha and 5 other annelids) were already published [10, 11] (*Table I*).

Parts of the Polychaete D<sub>1</sub> divergent domain (which consists of 142-155 nucleotides) are much more variable than in other invertebrates. In Alvinellids the length of this domain is of the same size as in other Polychaetes except in the case of *Alvinella* species whose D<sub>1</sub> domain consists of 176-178 nucleotides (*Fig. 1*). No difference in the sequenced domains was found between specimens reported to the same species of *Alvinella* or *Paralvinella* from venting locations more than 1 000 kilometres apart. This agrees with results of population genetic studies which have demonstrated that alvinellid populations are connected all along the rifts [20]. The very few data on alvinellid larval development do not help in understanding how they move from one hydrothermal site to another, from some tens of metres to kilometres away, through the cold abyssal water and how they colonised a linear environment of several thousand kilometres. This happened in spite of a

possible pseudocopulation and an internal fecundation [22] and of a negative egg buoyancy that indicates a probable recruitment near the parents [23]. However, the observation of hydrothermal organisms outside the vent locations [24, 25] suggests the possibility of dispersal by epizoochory. Another hypothesis is the dispersion of young stages *via* the median rift current.

It was recently stated [9] that a relationship exists between a (C+G) enrichment in alvinellid rRNA and an increase of the thermal stability or a greater chemical resistance of this molecule in highly reduced and metal rich environments such as hydrothermal vents. We find that alvinellid (C+G) levels are of 47-54 % in the conserved studied domains as in other invertebrates and vertebrates. (C+G) content increases in the D<sub>1</sub> domain (Table II). The average is 70 % in Alvinellids (67-74 %, extreme values), while it is 68 % in other Terebellomorpha (61-77 %); it is of the same order of magnitude as some Antarctic Polychaetes (68 %) or temperate (71-73 %) and polar (68-71 %) Echinoids (Fig. 3). Moreover, the (C+G) level of the hydrothermal vent Mytilidae Bathymodiolus sp from the Fidji back-arc basin is 58 %. If any, the relationship between (C+G) level and extreme thermal environment is not direct. Therefore, the evolution of (C+G) level cannot be explained by environmental factors. Moreover, the existence of isochores has been demonstrated in vertebrates and plants [26-28]. A linear correlation may also exist in invertebrates between the (C+G) level of a gene and the (C+G) level of the isochore harboring this gene. This potential impact of a "genome compositional constraint" should also be taken into account.

Phylogenetic trees inferred by distance and parsimony methods using PHYLIP [15] and MUST [16] computer packages, show that the alvinellids are monophyletic (bootstrap values of

90 % [17]) and that they belong to the monophyletic taxon of Terebellomorpha (Fig. 2). They are clustering at least in two taxa supported by very high bootstrap values (99 %). Alvinella is a monophyletic group. Paralvinella may be monophyletic, but the fact that only 3 out of 9 species were sequenced must be considered (possible paraphyly?). Compared to other Polychaetes, phylogenetic distances inside alvinellids suggest a high rate of mutation fixation. But actually the increases in length of about 30 nucleotides in Alvinella D1 domain that are inserted by blocks of 3 to 11 nucleotides (Fig. 1) are probably due to a succession of events such as mutations (may be induced by the high levels of natural radioactivity found in the vent locations [20]) and fixations. Thus, the fixations of the nucleotides are not necessarily a sign of an acceleration of the evolution speed, but they could have been promoted by very different facts: (1) the insertions do certainly not change too much the conformation of the rRNA molecule, the ribosome being still functionnal [18] and (2) due to the necessity of an individual's fitness to colonize new active hydrothermal vents before the activity of its own vent location ends, a new colony each time derives from just a few individuals. This resulted in a founder effect. Except Terebellides stroemi (Trichobranchidae) which clusters among Terebellidae, the inferred groups agree with the morphological taxonomy. Terebellidae appears to be a possible alvinellid sister group. Paralvinella monophyly would be confirmed by the resolution of the molecular phylogeny of all Alvinellidae [29]. Our sample should be completed with other Paralvinella spp, particularly with those of the Northern part of the East Pacific Rise (Juan de Fuca) separated from the Southern part 25 M yrs ago [31]. This will allow to date some branches of our tree.

Remerciements: nous remercions tous ceux qui nous ont apporté leur aide par leurs commentaires et de fructueuses discussions (D. Jollivet, É. Poulin, un arbitre anonyme) ainsi que A. Patel et D. McKey qui ont amélioré le texte anglais. Nous sommes également reconnaissants envers les responsables des campagnes « Mariana BAB expedition » (R.R. Hessler et P.A. Lonsdale), « MVT » (R.A. Lutz et R. Vriejenhoek) pour la fourniture du matériel biologique, envers J.-C. Roca pour la collecte d'annélides en plongée, ainsi qu'envers J.-C. Duchêne et C. Jouin-Toulmond pour l'identification de certains polychètes. Ce travail a été soutenu par le CNRS (GDR ECOPROPHYCE) et par l'IFRTP (programme « MacroBenthos »).

#### RÉFÉRENCES

- Desbruyères D., Laubier L. 1980. Alvinella pompejana gen. sp. nov., Ampharetidae aberrant des sources hydrothermales de la ride Est-Pacifique. Oceanol. Acta 3: 267-74.
- 2. Desbruyères D., Laubier L. 1986. Les Alvinellidae, une famille nouvelle d'annélides polychètes inféodés aux sources hydrothermales sous-marines: systématique, biologie et écologie. *Can. J. Zool.* 64: 2227-45.
- 3. Desbruyères D., Laubier L. 1991. Systematics, phylogeny, ecology and distribution of the Alvinellidae (Polychaeta) from deep-sea hydrothermal vents. *Ophelia* 5 (suppl.): 31-45.
- 4. Desbruyères D., Laubier L. 1993. New species of Alvinellidae (Polychaeta) from the North Fidji back-arc basin hydrothermal vents (South Western Pacific). *Proc. Biol. Soc.* 106: 225-36.
- 5. Tunnicliffe V., Desbruyères D., Jollivet D., Laubier L. 1993. Systematic and ecological characteristics of *Paralvinella sulfincola* Desbruyères and Laubier, a new Polychaete (Family Alvinellidae) from Northeast Pacific hydrothermal vents. *Can. J. Zool.* 71: 286-97.
- 6. Chevaldonné P., Desbruyères D., Childress J.J. 1992. Some like it hot... and some even hotter. *Nature* 359: 593-4.
- 7. Lalou C., Reyss J.L., Brichet E. 1993. Actinide-series desequilibrium as a tool to establish the chronology of deep-sea hydrothermal activity. *Geochimica Cosmochimica* 57: 1221-31.

- 8. Qu L.H., Nicoloso M., Bachellerie J.-P. 1988. Phylogenetic calibration of the 5' terminal domain of the large rRNA achieved by determining twenty eukaryotic sequences. *J. Mol. Evol.* 28: 113-24.
- 9. Dixon D.R., Dixon L.R.J., Tunnicliffe V. 1992. (C+G) (Cytosine-Guanine)-enrichment of ribosomal DNA of deep sea hydrothermal vent polychaetes: adaptation to extreme thermal environments. 4th International Polychaete Conference, Angers, France, 27/7/92-1/8/92 (summary).
- 10. Lenaers G., Bhaud M. 1992. Molecular phylogeny of some polychaete annelids: an initial approach to the Atlantic-Mediterranean speciation problem. *J. Mol. Evol.* 35: 429-35.
- 11. Chenuil A. 1993. Étude des relations de parenté entre les principaux groupes d'invertébrés protostomiens par amplification, séquençage et comparaison de portions du gène de l'ARN 28S. Thèse de doctorat, Université Montpellier II.
- 12. Williams N.A., Dixon D.R., Southward E.C., Holland P.W.H. 1993. Molecular evolution and diversification of the vestimentiferan tube worms. J. Mar. Biol. Ass. U.K. 73: 437-52.
- 13. Hassouna N., Michot B., Bachellerie J.-P. 1984. The complete nucleotide sequence of mouse 28S rRNA gene. Implications for the process of size increase of the large subunit rRNA in higher eukaryotes. *Nucleic Acids Res.* 8: 3563-83.
- 14. Qu L.H., Michot B., Bachellerie J.-P. 1983. Improved method for structure probing in large RNAs: a rapid heterologous sequencing approach is coupled to the direct mapping of nuclease accessible sites. Implications to

- the 5' terminal domains of eukaryotic 28S rRNA. *Nucleic Acids Res.* 11: 5903-20.
- 15. Felsenstein J. 1990. PHYLIP version 3.3, University of California, Berkeley, USA.
- 16. Philippe H. 1993. MUST, a computer package of Management Utilities for Sequences and Trees. *Nucleic Acids Res.* 21: 5264-72.
- 17. Felsenstein J. 1985. Confidence limits on phylogenies : an approach using the bootstrap. *Evolution* 39: 783-91.
- 18. Gerbi S.A. 1985. Evolution of ribosomal DNA. In : Intyre R.J., ed. *Molecular Evolutionary Genetics*. New York : Plenum Press, 419-57.
- 19. Cherry R., Desbruyères D., Heyraud M., Nolan C. 1992. High levels of natural radioactivity in hydrothermal vent polychaetes. *C. R. Acad. Sci. Paris, Ser. III* 315: 21-6.
- 20. Qu L.H. 1986. Structuration et évolution de l'ARN ribosomique 28S chez les eucaryotes. Étude systématique de la région 5'-terminale. Thèse de doctorat d'État. Université Paul-Sabatier, Toulouse.
- 21. Jollivet D. 1993. Distribution et évolution de la faune associée aux sources hydrothermales à 13°N sur la dorsale du Pacifique oriental : le cas des Polychètes Alvinellidae. Thèse de doctorat de l'Université de Bretagne occidentale.
- 22. Zal F., Desbruyères D., Jouin-Toulmond C. 1994. Sexual dimorphism in *Paralvinella grasslei*, a polychaete annelid from deep-sea hydrothermal vents. C. R. Held. Séanc. Acad. Sci. Paris 317: 42-8.

- 23. Chevaldonné P., Jollivet D. 1993. Videoscopic study of deep-sea hydrothermal vent alvinellid polychaete populations: biomass estimation and behaviour. *Mar. Ecol. Prog. Ser.* 95: 251-62.
- 24. Tunnicliffe V., Jensen R.G. 1987. Distribution and behaviour of the spider crab *Macroregonia macrochira* Sakai (Brachyura) around the hydrothermal vents of the Northeast Pacific. *Can. J. Zool.* 65: 2443-9.
- 25. Van Dover C.L., Franks P.J.S., Ballard R.D. 1987. Prediction of hydrothermal vent locations from distributions of brachyuran crabs. *Limnol. Oceanogr.* 32: 1006-10.
- 26. Bernardi G., 1986. Compositional constraints and genome evolution. J. Mol. Evol. 24: 1-11.
- 27. Bernardi G., 1991. Compositional properties of nuclear genes from cold-blooded vertebrates. J. Mol. Evol. 33: 57-67.
- 28. Montero L.M., Salinas J., Matassi G., Bernardi G. 1990. Gene distribution and isochore organization in the nuclear genome of plants. *Nucleic Acids Res.* 18: 1859-67.
- 29. Felsenstein J. 1978. Cases in which parsimony and compatibility methods will be positively misleading. *Syst. Zool.* 27: 401-10.
- 30 Lecointre G., Philippe H., Le H.L.V., Le Guyader H. 1993. Species sampling has a major impact on phylogenetic inference. *Mol. Phyl. Evol.* 2: 205-24.
- 31. Tunnicliffe V. 1988. Biogeography and evolution of hydrothermal vent fauna in the Eastern Pacific Ocean. *Proc. R. Soc. Lond.* B 233: 347-66.

Numéros d'accès des séquences de l'article : phylogénie moléculaire des annélides Alvinellidae... à la banque de données de Heidelberg (EMBL). X80645 : Alvinella caudata. X80646 : Alvinella pompejana. X80647 : Melinna palmata. X80648 : Mesonerilla sp. X80649 : Nephtys hombergi. X80650 : Owenia fusiformis. X80651 : Ophelia rathkei. X80652 : Potamilla antarctica. X80653 : Paralvinella fidjensis. X80654 : Paralvinella grasslei. X80655 : Paralvinella hessleri. X80656 : Saccocirus major. X80657 : Thelepus cincinnatus. X80658 : Terebellides stroemi.