

Un nouvel outil pour l'évaluation des ressources génétiques : l'hybridation *in situ* appliquée au complexe des *Allium*

Agnès RICROCH *

Résumé: Selon l'approche de Jean Pernès, le complexe d'espèces du genre *Allium* sera décrit en relation avec la gestion de ses ressources génétiques (conservation *in situ* et *ex situ*). Les différentes stratégies de reproduction au sein de l'espèce cultivée *A. cepa* seront discutées selon leur avantage sélectif. Enfin sera présentée l'hybridation *in situ* sur chromosomes d'une sonde froide et fluorescente, technique nouvelle permettant d'obtenir de nouveaux marqueurs cytogénétiques. Cette technique, rapide et précise, est précieuse pour suivre l'introgression de chromosomes dans de nouveaux contextes génomiques et s'avère utile pour l'évaluation des ressources génétiques puisqu'elle fournit un profil chromosomique unique pour chaque génotype caractérisé.

Mots-clés : complexe d'espèces, *Allium*, ressources génétiques, hybridation, sonde froide et fluorescente, mode de reproduction

Le concept pernessien

Soucieux de valoriser des espèces végétales nécessaires au développement humain avant qu'elles ne disparaissent au profit d'espèces modernes à base génétique étroite, Jean Pernès élabora une stratégie de gestion des ressources génétiques en quatre opérations utilitaires de base : collecte des échantillons, évaluation, conservation et enregistrement des informations dans des banques de données.

Le concept **pernessien** de gestion des ressources génétiques qui en découle permet la compréhension d'une organisation dynamique des populations en complexe d'espèces. Ce concept novateur présente la triple originalité :

- d'apprécier les inter-relations des différentes composantes du complexe d'espèces en termes de contrôles des échanges (ou flux) géniques,
- d'évaluer l'utilité et la valeur des ressources génétiques,
- et de proposer des schémas de conservation *in situ*.

* Université Paris-Sud, Laboratoire d'Evolution et de Systématique Végétales, Bâtiment 362, 91405 Orsay cedex, France.

L'unité prise en compte dans l'évaluation de la diversité végétale selon Pernès est la population, ensemble d'individus reflétant la variabilité génétique **infraspécifique**, d'où l'attention toute particulière portée à l'échantillon collecté. L'évaluation multi-niveaux, agronomique et morphologique, de l'échantillon-population peut être complétée par des observations plus fines décelant les différences cachées entre les individus d'un même échantillon, d'où la nécessité de l'outil moléculaire dans les programmes de valorisation des ressources génétiques. La technique **d'électrophorèse** unidimensionnelle est un outil permettant d'illustrer le polymorphisme **isozymique** des populations naturelles et des variétés traditionnelles en mesurant le taux de polymorphisme, **l'hétérozygotie** et les distances génétiques. Une distribution géographique des allèles vient compléter celle des observations morphologiques et agronomiques. L'ensemble de ces descripteurs permet de tirer des informations sur l'organisation génétique de l'échantillon. La caractérisation à un niveau moléculaire contribue au marquage du génome de l'espèce considérée.

Un nouvel outil moléculaire : FISH

Les méthodes de la biologie moléculaire, qui connaissent une évolution rapide, ont une portée pratique pour la caractérisation des ressources génétiques. Notre propos est de présenter une toute nouvelle technique de marquage moléculaire de chromosomes par hybridation fluorescente *in situ* de sonde **biotinylée** (Fluorescent In Situ Hybridization). Cette technique est couramment employée pour cartographier le génome humain. Le marquage des sondes **biotinylées** et fluorescentes permet une détection des sites d'hybridation rapide (quelques heures), précise (0,25 μm contre 1 μm avec le tritium) et fiable (sur le chromosome).

A l'aide d'une sonde d'ADN ribosomal **d'*Helianthus argophyllus***, nous avons testé avec succès l'hybridation *in situ* de chromosomes **d'*Allium cepa*** (groupe des oignons) est de sélectionner pour la résistance aux stress environnementaux et aux maladies. La plupart des hybrides F1 issus de croisements **interspécifiques** avec *A. cepa* ont été obtenus avec *A. fistulosum* et aucun avec *A. schoenoprasum*.

La stérilité des hybrides FI **interspécifiques** est une barrière majeure à l'**introgression** des gènes ; le transfert de caractéristiques d'une espèce vers une autre peut être alors réalisé via des lignées d'addition. A un autre niveau, les analyses cytogénétiques permettent, outre la description du matériel, l'étude des relations entre les composantes du complexe d'espèces. En provoquant des hybridations entre ces composantes, il est alors possible d'étudier les confrontations chromosomiques chez les hybrides. Le résultat de ces hybridations traduit assez bien le degré d'éloignement des groupes d'individus (Pernès, 1985).

Ce type d'analyse souligne l'importance des formes spontanées comme réserve originale d'allèles. En prospectant dans les centres d'origine et de diversification, où la domestication s'est effectuée très lentement en dehors

de l'influence des technologies modernes, les sources de résistance aux maladies provenant d'espèces sauvages et de variétés ancestrales peuvent être valorisées en les transférant dans des cultivars modernes.

Des gènes de résistance sont ainsi présents chez *Allium fistulosum*, une espèce étroitement apparentée à *A. cepa*, provenant de l'Extrême-Orient où de nombreux types distincts se sont développés suite à une longue histoire de domestication. Ces gènes peuvent être transférés vers *Allium cepa* en utilisant des lignées d'addition (Peffley et Mangum, 1990). Les analyses cytogénétiques et *isozymiques* ont révélé la localisation des loci ADH-1 (alcool déshydrogénase) et PGM-1 (*phosphoglucomutase*), respectivement sur le chromosome 5 *subtelocentrique* et le chromosome 4 *submetacentrique* d'*A. fistulosum* (Peffley et Currah, 1988). Cependant, la détection de cette *introgression* est limitée sans marqueurs additionnels. Les sites d'hybridation *in situ* fournissent ces marqueurs supplémentaires pour l'étude de la structure du chromosome, la cartographie des gènes dans la constitution d'une carte physique, l'individualisation des chromosomes homologues et la détection des chromosomes spécifiques dans de nouveaux contextes génétiques (F1, F2, BC...).

Le matériel végétal utilisé est composé de :

- deux espèces parentales *A. cepa* ($2n = 2x = 16$) cv « New Mexico Yellow Grano » et *A. fistulosum* ($2n = 2x = 16$) cv « Heishiko »,
- et de deux hybrides *interspécifiques* diploïde [*A. cepa* x *A. fistulosum* ($2n = 2x = 16$)] cv « 81215 » et triploïde [(*A. cepa* x *A. cepa*) x *A. fistulosum* ($2n = 3x = 24$)] cv « Delta Giant » (Perkins *et al.*, 1958).

La sonde employée, le clone pHAR1, consiste en un plasmide pBR322 qui contient une séquence d'ADNr de 9,8 kb d'*Helianthus argophyllus* (Choumane et Heizmann, 1988). La méthode se déroule en 7 étapes :

- marquage à la **biotine** de la sonde (ADN exogène),
- traitement des préparations chromosomiques à la **RNase**,
- dénaturation de l'ADN des préparations chromosomiques (ADN endogène),
- dénaturation de l'ADN de la sonde (ADN exogène),
- hybridation de la sonde (15-16 h),
- marquage de l'antigène **avidine** par un agent fluorescent (fluorescéine),
- détection immunologique des sites d'hybridation par un anticorps **anti-avidine**, en parallèle avec un deuxième anticorps (antifade).

Les sites d'hybridation obtenus (Fig. 1) révèlent un marquage dans la région terminale du bras court, y compris la paire de chromosomes **satellifères** d'*Allium cepa* 6, sauf chez *A. fistulosum* (paire **satellifère** 5). On observe un marquage spécifique des séquences d'ADNr par une position différente des sites. Le nombre des sites d'hybridation des paires **satellifères** est variable, ce qui suppose que la plupart des séquences d'ADNr sont **délétées** dans un des chromosomes homologues de la paire 6 de *A. cepa* et 5 de *A. fistulosum* indiquant l'inactivité de la région organisatrice **nucléolaire**. Le marquage des paires 8 de *A. cepa* et de *A. fistulosum* indique la présence d'une région organisatrice **nucléolaire** additionnelle aux paires **satellifères**, et par le nombre de sites différents des chromosomes homologues, son activité. La position **distale** des sites intercalaires du chromosome 8F" suppose la transposition d'ADNr dans de nouveaux sites chromosomiques chez *A. fistulosum*. Le marquage des hybrides **interspécifiques** révèle une distinction non-ambiguë

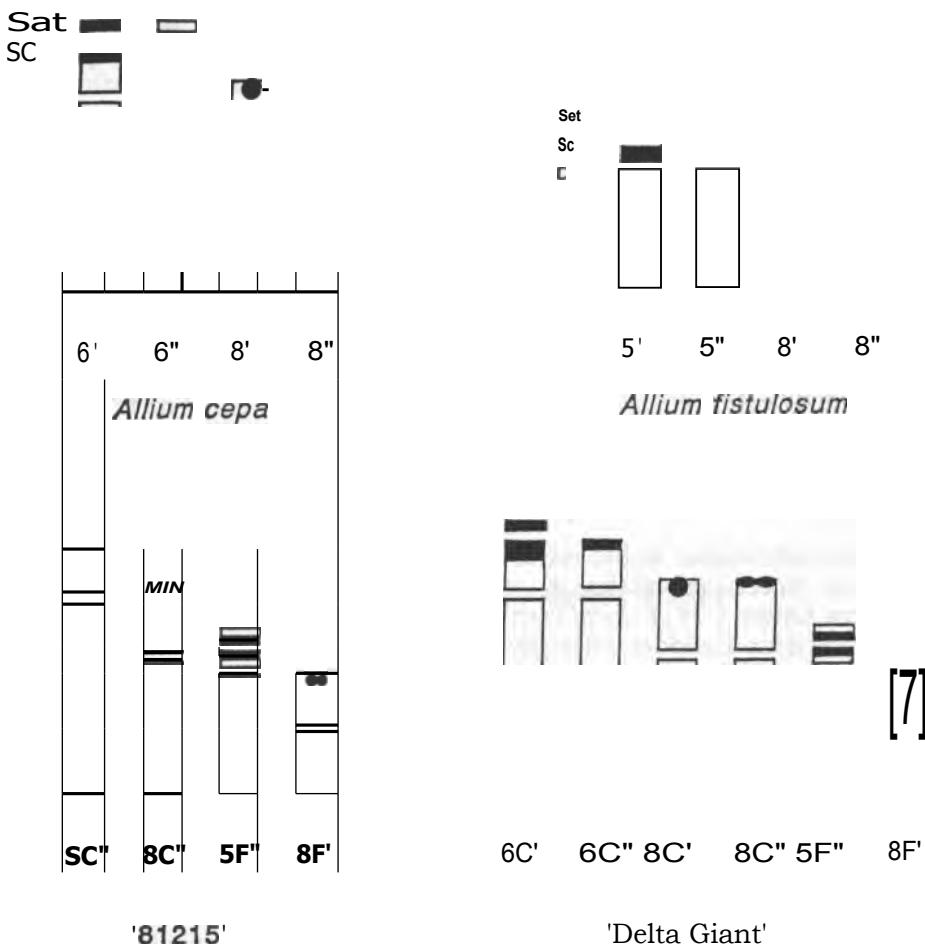


Fig. I. — Représentation schématique des sites d'hybridation *in situ* des chromosomes d'*Allium cepa* cv. «New Mexico Yellow Grano», d'*Allium fistulosum* cv. « Heishiko », d'un hybride diploïde interspécifique cv. « 81215 » et d'un hybride triploïde interspécifique cv. « Delta Giant ». Abréviations : Sat = satellite ; SC = constriction secondaire ; C = centromère.

entre les chromosomes paternels et maternels, et entre les chromosomes homologues.

L'emploi d'une sonde d'ADNr à séquences répétées en tandem révèle une localisation restreinte et prédominante des sites d'hybridation dans la région téloïérique (hétérochromatique) du bras court des Monocotylédones, excepté la localisation d'ADNr dans de nouveaux sites intercalaires (chez *A. fistulosum*). Par cette méthode, l'ADN ribosomal apparaît bien associé aux constrictions secondaires. Enfin, ces résultats ne sont pas en désaccord avec l'hypothèse que les régions organisatrices nucléolaires dans le complexe multispecifique des *Allium* sont mobiles par translocation.

Ce nouvel outil moléculaire permet par une détection du nombre de sites d'hybridation et de leur position un marquage précis des chromosomes en

révélant un profil unique, utile pour individualiser les chromosomes homologues. De par sa rapidité et sa précision, il peut être employé pour discriminer des espèces très proches et pour suivre l'introgression de chromosomes individuels chez les hybrides interspécifiques.

L'hétérochromatine en question

Cette méthode de génétique moléculaire est employée dans notre équipe pour étudier l'organisation structurelle des chromosomes. Les génomes eucaryotiques présentent un contenu et une organisation de l'ADN fortement variables, des petites régions d'ADN peuvent être altérées par des mécanismes tels que la délétion, la duplication, la mutation ou la translocation. La fixation de ces types d'altération peut créer des barrières génétiques contribuant à la réduction de la fertilité des hétérozygotes et conduisant à une possible spéciation. Beaucoup de ces séquences répétées sont localisées près des centromères ou télomères, ou moins fréquemment, dans les régions interstitielles des chromosomes. Elles sont associées aux régions hétérochromatiques.

L'hétérochromatine constitutive ne contient pas de séquences codantes mais des séquences répétées en tandem. Celles-ci peuvent être détectées par hybridation *in situ* à l'aide d'une sonde d'ADN satellite, comme nous l'avons montré avec la sonde d'ADN ribosomal de tournesol. L'hétérochromatine constitutive est variablement enrichie en ADN satellite selon les espèces. Les transposons sont souvent localisés dans l'hétérochromatine. La quantité et la composition de l'hétérochromatine ne sont pas uniformes chez les espèces ; des différences ont été décelées par des techniques de coloration de C-banding chez des espèces même très proches. Des lignées de maïs adaptées aux hautes altitudes présentent un petit nombre de « knobs » (segment hétérochromatique), suggérant que le contenu de l'ADN diminue avec l'altitude (Rayburn et Auger, 1989). Chez *Crepis*, il existe une corrélation entre la quantité et la distribution de l'hétérochromatine et la localisation altitudinale des espèces (Siljak-Yakovlev, 1986).

Les variations de la quantité d'ADN total peuvent être considérées comme l'expression d'un équilibre dynamique entre la prolifération des séquences répétées et la sélection contre l'augmentation de la quantité d'ADN (Schweitzer *et al.*, 1989). Cette augmentation aurait pour effet de modifier le volume et le rythme des divisions cellulaires (Price, 1976 ; Rees *et al.*, 1981). Ceci pourrait expliquer que la diminution de la taille des organismes et l'accélération de leur cycle de vie sont associées à une diminution de la quantité d'ADN. Chez *Crepis*, le passage d'un type pérenne vers un type annuel est accompagné d'une réduction du nombre de base de chromosomes et de la quantité d'ADN (Siljak-Yakovlev, 1982).

La dynamique de ces séquences est considérée par certains auteurs comme un phénomène rapide à l'échelle de l'évolution biologique. Cela aurait deux conséquences, d'une part la quantité d'ADN nucléaire serait corrélée positivement à celle de l'hétérochromatine, et d'autre part le pourcentage des bases C + G des séquences répétées par rapport à celui de l'ADN total

nucléaire augmenterait avec la teneur en **hétérochromatine** (Godelle, 1990). Ces différences de composition en ADN reflétant des origines diverses demeurent encore non élucidées.

Modèles de complexes d'espèces

Le complexe **multispecifique** des *A. Ilium* se prête bien à ce type d'études cytogénétiques grâce à la visibilité excellente des **chromosomes**, mais aussi parce qu'il renferme des espèces ayant subi des pressions de sélection liées au processus de domestication depuis au moins 5 000 ans. Les écotypes qui en résultent sont adaptés à des habitats très localisés, donc à des cycles de culture très précis répondant à des exigences **thermo** et photopériodiques rigoureuses. Les variations de la quantité d'ADN total reliées à celles de la quantité **d'hétérochromatine** pourraient être en partie responsables de ces adaptations climatiques (Jones et Rees, 1988).

Au laboratoire, nous utilisons une sonde fluorescente et **biotinylée** de séquences répétées pour détecter des modifications de la nature de l'ADN d'espèces proches, appartenant à des complexes aussi divers que ceux de *Crepis* et de *Pennisetum*. L'utilisation des mêmes sondes à séquences répétées permettra d'éclaircir les hypothèses sur l'équilibre dynamique entre amplification et sélection contre l'augmentation de la quantité d'ADN répété. Aux côtés d'autres techniques (**RFLP**), cette méthode contribuera au marquage du génome, notamment celui du mil, plante modèle de notre groupe. Sauver, évaluer et valoriser les ressources génétiques des plantes à l'aide d'outils nouveaux et prometteurs sont les objectifs que nous nous donnons, en respect de la pensée de Jean **Pernès** : « Gérer les ressources génétiques c'est assurer un avenir à l'humanité, c'est donc un pari incertain que force nous est de tenir » (Pernès, 1984).

Bibliographie

- CHOUMANE** W. and HEIZMANN P., 1988 — Structure and variability of nuclear ribosomal genes in the genus of *Helianthus*. *Theor. Appl. Genet.*, 76 : 481-489.
- GODELLE** B., 1990 — *Quantité d'ADN et composition des bandes C dans le complexe Crepis praemorsa : un modèle d'évolution de l'hétérochromatine*. Rapport de DEA, 32 p.
- JONES** R.N. and **REES** H., 1988 — Nuclear DNA variation in *Allium*. *Heredity*, 23 : 591-605.
- PEFFLEY** E.B. and **CURRAH** L., 1988 — The chromosomal locations of enzyme-coding genes **Adh-1** and **Pgm-1** in *Allium fistulosum* L. *Theor. App!. Genet.*, 75: 945-949
- PEFFLEY** E.B. and **MANGUM** P.D., 1990 — **Introgression of Allium fistulosum** L. into *A. cepa* L. : **cytogenetic** evidence. *Theor. App!. Genet.*, 79: 113-118

- PERKINS D.Y., KEHR A.E., BROWN R.T., TIMS E.C., MILLER J.C., 1958 — "Delta Giant", a new long season shallot. Circular n° 52. Louisiana State University and Agricultural and Mechanical College, Agricultural Experiment Station.
- PERNÈS J., 1984 — Gestion des ressources génétiques des plantes. Tome 2 : Manuel Technique et Documentation. ACCT, Paris. 346 p.
- PRICE H.J., 1976 — Evolution of DNA content in higher plants. *Bot. Review*, 42 : 27-52.
- RAYBURN A.L. and AUGER J.A., 1989 — Genome size variation in *Zea mays* ssp *mays* adapted to different altitudes. *Theor. Appl. Genet.*, 79 : 470-474.
- REES H., NARAYAN R.K.J., 1981 — Chromosomal DNA in higher plants. *Phil. Trans. R. Soc. Lond.*, 292: 569-578.
- SCHWEITZER D., STREHL S., HAGEMANN S., 1989 — Plant repetitive DNA elements and chromosomal structure. *Chromosomal Today*, 10 : 33-43.
- SILJAK-YAKOVLEV S., CARTIER D., 1982 — Comparative analysis of C-band karyotypes in *Crepis praemorsa* subsp. *praemorsa* and subsp. *dinarica*. *Pl. Syst. Evol.*, 141 : 85-90.
- SILJAK-YAKOVLEV S., CARTIER D., 1986 — Heterochromatin patterns in some taxa of *Crepis praemorsa* complex. *Caryologia*, 39 : 27-32.

Approche moléculaire et cellulaire du génome

Abbrechre wo u ii
génoue

Phylogeny of adenylyl and guanylyl cyclases

Antoine DANCHIN *

Résumé: LAMP cyclique et le GMP cyclique sont des messagers secondaires quasi-universels (ils sont toutefois absents des végétaux), qui relaient, à l'intérieur de la cellule, les mécanismes régulateurs mis en marche par les variations de l'environnement. L'étude phylogénétique des enzymes qui en permettent la synthèse révèle l'existence de trois classes indépendantes, comme s'il y avait eu convergence évolutive. Mais l'observation la plus frappante est qu'une de ces classes est présente aussi bien chez les eucaryotes que chez les procaryotes, ce qui indique que le rôle initial des nucléotides cycliques pourrait avoir été différent. Cela semble confirmé par le fait que les *cyclases* correspondantes sont voisines d'enzymes permettant la synthèse de l'ATP et le transport des ions. Il est alors possible d'imaginer que le rôle initial des nucléotides cycliques était peut-être de permettre la synthèse des nucléosides triphosphates.

Mots-clés : AMP cyclique, GMP cyclique, évolution, calmoduline, entérobactéries, synthèse d'ATP.

Abstract : Comparison of the primary sequence of adenylyl and guanylyl cyclase genes from phylogenetically distant organisms permits us to propose that they form three classes. The first class comprises enzymes isolated from enterobacteria and related bacteria. The second class comprises the prokaryotic toxic adenylyl cyclases, activated by the host calmodulin, synthesised by *Bordetella pertussis* and *Bacillus anthracis*. Finally, the third class comprises adenylyl cyclases and guanylyl cyclases isolated from organisms ranging from prokaryotes to higher eucaryotes. This suggests that cAMP synthesis predated separation between eucaryotes and prokaryotes. The extreme diversity displayed between the different classes raises the question of the origin of present day cAMP : are cyclases a case of phylogenetic convergence ? A tree can be constructed that links all enzymes to each other, but, even if one takes this tree for granted it appears that cAMP synthesis has probably been reinvented at least once.

Key words: cyclic AMP, cyclic GMP, evolution, calmodulin, enterobacteria, ATP synthesis.

cAMP is the product of an enzyme, adenylyl cyclase, that generates cAMP and PP_i from ATP. The study of cAMP-mediated effects permitted identification of the structure, function and regulation of many adenylyl cyclases. The recent discovery that several prokaryotic adenylyl cyclase genes synthe-

* Unité de Régulation de l'Expression Génétique, Institut Pasteur, 28 rue du Docteur Roux, 75724 Paris cedex 15, France.

sized proteins related to their *eucaryotic* counterparts (Beuve *et al.*, 1990) has demonstrated that *adenylyl cyclases* do not constitute a homogeneous class of enzymes, but comprise at least three well defined types of enzymes, the *enterobacterial* type, the toxin type (corresponding to *cyclases* secreted by pathogens such as *Bacillus anthracis* and *Bordetella pertussis*) and a class, grouping the vast majority of *cyclases*, that comprises both *prokaryotic* and *eucaryotic* enzymes.

The *enterobacterial* class (class I)

The first *adenylyl cyclase* gene to be isolated and sequenced was the *Escherichia coli cya* gene. Cloning of *cya* genes from other bacterial types was obtained by direct complementation of *cya*-deficient *E. coli* mutants. In *Erwinia chrysanthemi* both the gene environment and the proteins are similar in size and overall organization to the *E. coli* counterparts (Danchin and Lenzen, 1988). A similar type of work was subsequently performed by several authors on a truncated *Salmonella typhimurium* gene, but the complete sequence was not obtained (Fandl *et al.*, 1990; Holland *et al.*, 1988; Thorner *et al.*, 1990). Finally, the *cya* gene from two other *enterobacteria*, *Yersinia intermedia* and *Y. pestis* (the agent of plague), showed that all major features found in the *E. coli* gene and protein organization were conserved (Glaser, Sismeiro and Danchin, unpublished). Knowledge of the catalytic domain sequence from five genes permitted construction of a phylogenetic tree where *S. typhimurium* and *E. coli* are clustered together, as are the two *Yersinia* species (Fig. 1).

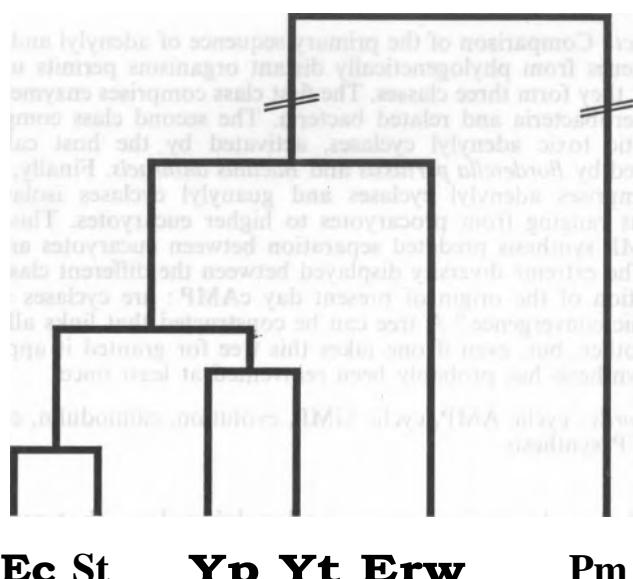


Fig. 1. — Phylogeny of class I bacterial adenylate cyclases. Ec : *Escherichia coli*; Erw : *Erwinia chrysanthemi*; Pm : *Pasteurella multocida*; St : *Salmonella typhimurium*; Yi : *Yersinia intermedia*; Yp : *Yersinia pestis*. Note that this does not necessarily reflect the phylogeny of the harbouring bacteria.

In all cases the 95 kDa polypeptide chain of the enzyme forms a bifunctional structure. The catalytic domain is amino-terminal, while the glucose-sensitive regulatory domain is carboxy-terminal. No significant similarity with known proteins was found, showing that these enzymes form a specific class. Finally, the gene from *Pasteurella multocida* (the causal agent of fatal septicemia in mammals) was found to be very similar to the enterobacterial type (Mock *et al.*, 1991). This indicated that the regulation of the protein activity (including glucose-mediated regulation) was preserved, but the expression of the gene may be different in *Pasteurellaceae* from the expression in *Enterobacteriaceae*.

The calmodulin-activated toxic class (class II)

B. pertussis, a Gram negative prokaryote, causes whooping cough by secreting in the host many proteins having toxic activities. Among these proteins an adenylyl cyclase activity was discovered (Hewlett and Wolff, 1976). In 1980 a further observation was made : the enzyme is activated by a host protein, calmodulin, which is not known to occur in bacteria (Wolff *et al.*, 1980 ; Goldhammer and Wolff, 1982). Two years later Leppla demonstrated that another toxic adenylyl cyclase, secreted by a Gram positive organism, *B. anthracis*, the etiological agent of anthrax, was also dependent on the presence of host calmodulin (Leppla, 1982). Several attempts to isolate the gene failed until a simple idea, introduction of a plasmid directing synthesis of calmodulin into an *E. coli* strain deficient in adenylyl cyclase activity, permitted the cloning, sequencing, and expression of adenylyl cyclase genes coding for the calmodulin-dependent cyclases (Glaser *et al.*, 1988a ; Escuyer *et al.*, 1989).

It was thus discovered that *B. pertussis* adenylyl cyclase is synthesized as a large bifunctional polypeptide of 1706 amino acid residues (Glaser *et al.*, 1988). The N-terminal 400 first residues of the protein display calmodulin-activated ATP-cyclizing activity, while the rest of the molecule was shown to be responsible for secretion as well as for the hemolytic activity of the pathogen (Ehrmann *et al.*, 1991 ; Glaser *et al.*, 1988a, 1989 ; Roger *et al.*, 1991). Sequence analysis indicated that the cyclase domain was fused to a polypeptide similar to the polypeptide chain of *E. coli* hemolysin toxin (Goebel and Hedgpeth, 1982). Accordingly, the name cyclolysin was coined for the toxic adenylyl cyclase from *B. pertussis* (Glaser *et al.*, 1988b). Glaser and coworkers analysed the adjacent regions for functionally related genes and discovered that the protein is secreted using a mechanism similar to that permitting release in the external medium of *E. coli* hemolysin or cognate proteins in related media. Three genes located downstream from the cyclolysin gene were identified as required for secretion (Glaser *et al.*, 1988b).

A different organization was found for *B. anthracis* adenylyl cyclase. It is encoded in a plasmid, together with another toxin (lethal factor, LF), and a carrier protein (protective antigen, PA), which is necessary for internalisation of both adenylyl cyclase and LF in host target cells (Beall *et*

al., 1962 ; Gordon *et al.*, 1988 ; Leppla *et al.*, 1984 ; Thorne *et al.*, 1960). The gene product is comprised of four segments. The first one is similar to signal peptides of other bacteria, permitting secretion of the protein. The second segment permits binding with PA. It displays regions of significant similarity with a cognate region in LF. The third domain is adenyl cyclase. It is finally followed by a domain of unknown function, with no known counterparts (Fig. 2).



Bacillus anthracis

Fig. 2. — Adenylate cyclase toxins are organized as independent modules : CYA, cyclase ; HLY, hemolysin ; —PA, binding domain for protective antigen.

The catalytic regions of the *B. pertussis* and *B. anthracis* enzymes are similar in four conserved regions that are involved in catalysis, calmodulin binding and activation. The first region contains a peptide, G(G/A)KS similar in sequence to the nucleotide binding region found in many ATP or GTP binding proteins (Fig. 3). It was therefore proposed to be part of the catalytic site (Gilles *et al.*, 1990 ; Glaser *et al.*, 1989).

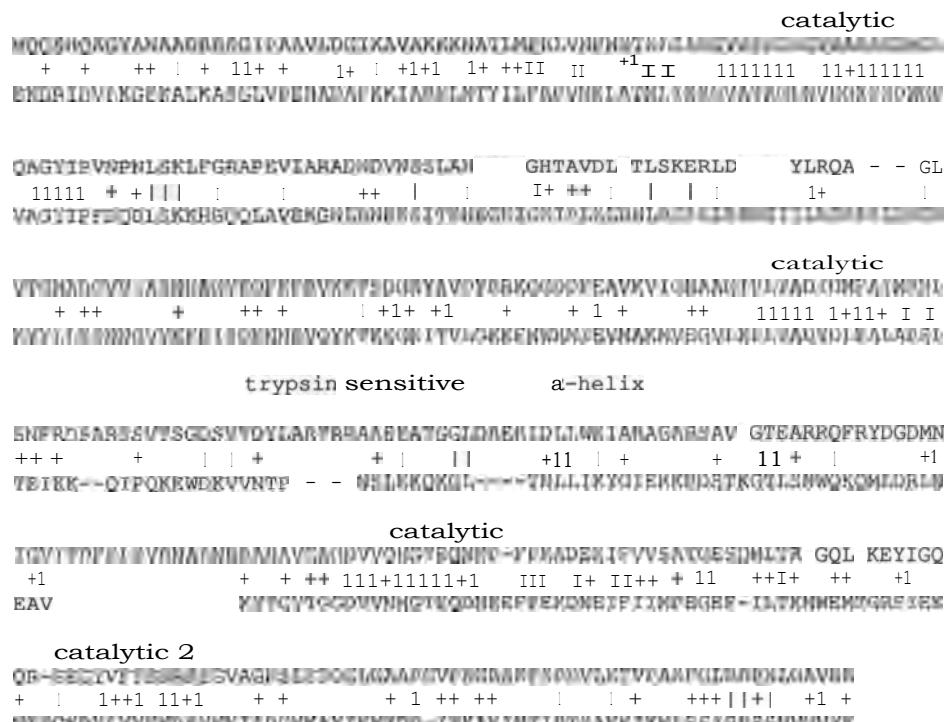


Fig. 3. — Comparison between the catalytic domains of *B. pertussis* and *B. anthracis* cyclases.
+ indicates conservative replacements.

Experiments with enzymes modified after *in vitro* mutagenesis substantiated this interpretation (Glaser *et al.*, 1991 ; Sarfati *et al.*, 1990). A second region was also demonstrated to be involved in several aspects of catalysis, and it was proposed that the aspartate residues present in this region, PLTADID, were involved in ribose and magnesium phosphate binding (Glaser *et al.*, 1991). Finally, the distal part of the cyclase domain has been found to be involved in catalysis. An amphiphilic helix, centrally located around residue tryptophane 242 is necessary for calmodulin dependent activation, thus suggesting that the protein must be folded around calmodulin in order to be active.

The « universal » class (class III)

Adenylyl cyclases from higher eucaryotes have long remained elusive enzymes, because the purification of the corresponding catalytic subunit is of extreme difficulty. In addition their activity is subject to a complex regulatory pattern, so that the catalytic center will often display poor or no activity when expressed alone. Techniques as the one that has been used for the cloning of adenylyl cyclase toxins are therefore difficult to undertake, because they may require the simultaneous presence of several regulatory subunits. The tedious but efficient method of Southern hybridization on cDNA libraries using oligonucleotides designed after the sequence of known peptides of the protein, has therefore been used. The availability of affinity chromatography using forskolin (Pfeuffer and Metzger, 1982) permitted purification of small amounts of the protein (Pfeuffer *et al.*, 1985a ; Coussen *et al.*, 1985 ; Pfeuffer *et al.*, 1985b, 1989), that finally resulted in the cloning of one adenylyl cyclase cDNA from bovine brain by Gilman and coworkers in 1989 (Krupinski *et al.*, 1989).

The difficulty encountered with the cloning of the higher eucaryotes adenylyl cyclase genes was fortunately not so crucial in organisms such as *Saccharomyces cerevisiae*. Cloning was attempted by Masson *et al.* who first cloned the wild type allele (Masson *et al.*, 1984). Subsequently Wigler *et al.* succeeded in determining the complete gene sequence after direct complementation of a *S. cerevisiae* mutant (Kataoka *et al.*, 1985). In parallel, Masson *et al.* cloned the gene directly in a *cya*-deficient *E. coli* strain (Masson *et al.*, 1986). These experiments demonstrated that this eucaryotic cyclase was completely different from the enterobacterial class, not only because of the sequence difference in the catalytic center, but also because the organization of the gene was different : the catalytic domain is COOH-terminal in *S. cerevisiae* cyclase, whereas it is NH₂ terminal in *E. coli*. This suggested that the new protein constituted the first member of a third adenylyl cyclase class. The yeast enzyme remained the only instance of its class until Garbers and coworkers recognized that the guanylyl cyclase genes that had been cloned from several metazoans, were clearly derived from a common ancestor (Garbers, 1990 ; Lowe *et al.*, 1989 ; Schulz *et al.*, 1989 ; Singh *et al.*, 1988). Another member of the class was subsequently discovered by Young *et al.* (1989) who cloned the adenylyl cyclase gene from *Schizosaccharomyces pombe* by hybridization using the catalytic domain gene

sequence from *S. cerevisiae* as a probe (Young *et al.*, 1989). Finally, the first higher eucaryote **adenylyl cyclase** gene, isolated in Gilman's laboratory from bovine brain, displayed features clearly reminiscent of this class (Krupinski *et al.*, 1989). Beuve *et al.* (1990) then described an **adenylyl cyclase** gene, isolated from the Gram negative procaryote *Rhizobium meliloti*, which was also a member of the same class (Beuve *et al.*, 1990). This was remarkable because procaryotes are supposed to have separated from eucaryotes at least 2.5 billions years ago. Since then many other genes or cDNAs for **adenylyl** or **guanylyl cyclases** belonging to this class have been isolated and sequenced, not only from eucaryotes (*S. pombe*, *Trypanosoma brucei* and *T. equiperdum*, *Plasmodium falciparum*, *Neurospora crassa*, human olfactory cells [Alexandre *et al.*, 1990 ; Bakalyar and Reed, 1990 ; Garbers, 1990 ; Kore-Eda *et al.*, 1991 ; Parma *et al.*, 1991 ; Read and Mikkelsen, 1991 ; Schulz *et al.*, 1989 ; Thorpe and Garbers, 1989 ; Young *et al.*, 1989]), but also from other procaryotes : a Gram positive organism, *Brevibacterium liquefaciens* was found to be a member of the same class (Peters *et al.*, 1991). It was also recently discovered that the Gram negative sliding bacteria, *Stigmatella aurantiaca*, harboured two **cyclase** genes, each of them corresponding to the same **adenylyl cyclase** class (B. Lubochinsky, V. de la Fuente, O. Sismeiro, A. Danchin, in preparation).

As shown in Figure 4, several amino acid motifs are specific of the class III proteins, such as **F-DIFT-L**, **QVV-LLND**, **VK(E)T-GDM**, **R-GIH-G**, **GDTVN-ASR**. Analysis of the upstream part of the domain suggested that the third motif could be involved in base discrimination, because it is different in **adenylyl cyclases** (VKT) and **guanylyl cyclases** (VET) (Peters *et al.*, 1991). However the set of individual **adenylyl cyclase** examples encompasses a series of organisms that is much wider, from the point of view of phylogeny, than the **guanylyl cyclase** set, the former containing organisms from widely separated organisms while the latter was obtained only from metazoans. The difference could therefore reflect a coincidence. Experiments meant to permit selection of evolution from **adenylyl** to **guanylyl** activity however demonstrated that the **GIH...GDTVN** region is likely to be involved in substrate discrimination (Beuve and Danchin, 1992).

Are **cyclases** the result of convergent evolution ?

In view of the apparent antiquity of **cAMP**, the hypothesis that it was originally involved in energy scavenging or production rather than regulation seems worth considering. It is thus of interest to note that there is a weak, but significant, similarity between this general class of **cyclases** and ATP syntheses. Phosphates (and polyphosphates) certainly played an important role at the origin of life (Bernal, 1951 ; Cairns-Smith, 1982 ; Wächtershäuser, 1988 ; Danchin, 1989). Generation of a **chemiosmotic** function allowing synthesis of ATP from ADP requires steps that are **difficult** to see fulfilled in a single stage. It may therefore be considered that, originally, **cAMP** has been a precursor of ATP. However this requires a widely spread source for **cAMP** synthesis : the catalytic activities of self-splicing **introns** are involving processes of **transesterification**, and one can propose that **cAMP** or **cGMP**

Fig. 4. — Alignment of class III cyclases catalytic domains. A ; adenylate cyclases : *R. meliloti*, *S. cerevisiae*, *S. pombe*, bovine brain (domains a and b), *B. liquifaciens*, *S. aurantiaca* (isozymes 1 and 2), *T. brucei* (isozyme 2), *T. equiperdum*, rat olfactory cells ; B ; guanylate cyclases : rat atrial natriuretic peptide receptor (ANP_b), human ANP_b, human ANP_a, mouse ANP_a, *S. purpuratus*, rat cytoplasmic 1, bovine cytoplasmic 1, *A. punctulata*, rat cytoplasmic 2.

have been by-products of such reactions, providing an early source for these molecules. In this scheme enzymes predating class III **adenylyl cyclases** would be necessary for ATP synthesis ; a second, very important enzyme, **adenylate kinase**, would then scavenge AMP, generating ADP. A general source of ADP thus available would have triggered evolution towards generation of ATP **synthase**. It seems noteworthy that the class of toxic **adenylyl cyclases** can be **phylogenetically** related to **adenylate kinase**, at least when considering the first block of conserved amino-acid residues in *B. pertussis* and *B. anthracis* enzymes.

Another observation may substantiate the validity of this hypothetical relationship. Antibodies raised against *B. pertussis* **adenylyl cyclase**, or even against a peptide motif common to the *B. pertussis* and *B. anthracis* enzymes could recognize a human brain **adenylyl cyclase** (Goyard *et al.*, 1989 ; Monneron *et al.*, 1988 ; Orlando *et al.*, 1991). This suggested that higher eucaryotes enzymes might display similarities with the **calmodulin** activated toxic proteins. Identification of an **adenylyl cyclase** gene from bovine brain by the group of Gilman revealed however, that the enzyme was clearly similar to the yeast enzyme and not to **adenylyl cyclase** toxins (Krupinski *et al.*, 1989). It is therefore difficult at the moment to know whether there exists a true similarity between class II and class III enzymes, or whether the class of brain **cyclases** is heterogeneous.

A tentative **phylogenetic** tree could be proposed, where all **cyclases** derive from an ancestral nucleotide binding structure (Figure 5). Alternatively, we could face a specific case of convergent evolution, as for instance in other ATP binding enzymes, the different **adenylyl cyclases** having independently evolved, the oldest activity corresponding to class III enzymes, derived from

Ancestral **Cyclase**

GC ACIII ATPsyn ADK ATPase ACII ACI

Fig. 5 A tentative evolutionary tree linking the three classes of **cyclases** together. Similarities are so low that the structure of the tree is highly speculative.

an ancestral NTP synthesizing enzyme. Some similarity can be discovered with such enzymes (Takeyama *et al.*, 1990), as well as with enzymes using ATP for positive ions translocation (Tamanoi *et al.*, 1984). Finally a puzzling observation challenges the ubiquity of cAMP function : cyanobacteria synthesize cyclic nucleotides, but plants do not seem to have cAMP. This may be due to the fact that photosynthesis has produced a very efficient way for ATP synthesis, making the original use of cyclic nucleotides as precursors obsolete.

References

- ALEXANDRE S., PAINDAVOINE P., TEBABI P., PAYS A., HALLEUX S., STEINERT M. and PAYS E., 1990 — Differential expression of a family of putative adenylate/guanylate cyclase genes in *Trypanosoma brucei*. *Mol. Biochem. Parasitol.*, 43 : 279-288.
- BAKALYAR H.A. and REED R.R., 1990 — Identification of a specialized adenylyl cyclase that may mediate odorant detection. *Science*, 250 : 1403-1406.
- BEALL F.A., TAYLOR M.J. and THORNE C.B., 1962 — Rapid lethal effects in rats of a third component found upon fractionating the toxin of *Bacillus anthracis*. *J. Bacteriol.*, 83 : 1274-1280.
- BERNAL J.D., 1951 — *The Physical Basis of Life*, Routledge and Kegan Paul, London.
- BEUVE A., BOESTEN B., CRASNIER M., DANCHIN A. and O'GARA F., 1990 — *Rhizobium meliloti* adenylate cyclase is related to eukaryotic adenylate and guanylate cyclases. *J. Bacteriol.*, 172 : 2614-2621.
- CAIRNS-SMITH A.G., 1982 — *Genetic Takeover and the Mineral Origin of Life*, Cambridge University Press, Cambridge.
- COUSEN F., HAIECH J., D'ALAYER J. and MONNERON A., 1985 — Identification of the catalytic subunit of brain adenylate cyclase : a calmodulin binding protein of 135 kDa. *Proc. Nat. Acad. Sci. USA*, 82 : 6736-6740.
- DANCHIN A., 1989 — Homeotypic transformation and the origin of translation. *Prog. Biophys. Molec. Biol.*, 54 : 81-86.
- DANCHIN A. and LENZEN G., 1988 — Structure and evolution of bacterial adenylate cyclase: comparison between *E. coli* and *E. chrysanthemi*. *Second Messengers and Phosphoproteins*, 12 : 7-28.
- EHRMANN I.E., GRAY M.C., GORDON V.M., GRAY L.S. and HEWLETT E.L., 1991 — Hemolytic activity of adenylate cyclase toxin from *Bordetella pertussis*. *FEBS Lett.*, 278 : 79-83.
- ESCUYER V., DUFLOT E., SEZER O., DANCHIN A. and MOCK M., 1988 — Structural homology between virulence associated bacteria adenylate cyclases. *Gene*, 71 : 293-298.
- FANDL J.P., THORNER L.K. and ARTZ S.W., 1990 — Mutations that affect transcription and cyclic AMP-CRP regulation of the adenylate cyclase gene (*cya*) of *Salmonella typhimurium*. *Genetics*, 125 : 719-727.
- GARBERS D.L., 1990 — The guanylyl cyclase receptor family. *The New Biologist*, 2 : 499-504.
- GILLES A.-M., MUNIER H., ROSE T., GLASER P., KRIN E., DANCHIN A., PELLECUE C. and BARZU O., 1990 — Intrinsic fluorescence of a truncated *Bordetella pertussis* adenylate cyclase expressed in *Escherichia coli*. *Biochemistry*, 29 : 8126-8130.

- GLASER P., LADANT D., SEZER O., PICHOT F., ULLMANN A. and DANCHIN A., 1988a — The calmodulin-sensitive adenylate cyclase of *Bordetella pertussis* : cloning and expression in *Escherichia coli*. *Mol. Microbiol.*, **2** : 19-30.
- GLASER P., SAKAMOTO H., BELLALOU J., ULLMANN A. and DANCHIN A., 1988b — Secretion of cyclolysin, the calmodulin-sensitive adenylate cyclase-haemolysin bifunctional protein of *B. pertussis*. *EMBO J.*, **7** : 3997-4004.
- GLASER P., ELMAOGLU-LAZARIDOU A., KRIN E., LADANT D., BARZU O. and DANCHIN A., 1989 — Identification of residues essential for catalysis and binding of calmodulin in *B. pertussis* adenylate cyclase by site-directed mutagenesis. *EMBO J.*, **8** : 967-972.
- GLASER P., MUNIER H., GILLES A-M., KRIN E., PORUMB T., BARZU O., SARFATI R., PELLECUER C. and DANCHIN A., 1991 — Functional consequences of single amino-acid substitutions in calmodulin-activated adenylate cyclase of *Bordetella pertussis*. *EMBO J.*, **10** : 1683-1688.
- GOEBEL W. and HEDGPETH J., 1982 — Cloning and characterization of the plasmid-encoded hemolysin determinant of *Escherichia coli*. *J. Bacteriol.*, **151** : 1290-1298.
- GOLDHAMMER A. and WOLFF A., 1982 — Assay of calmodulin with *Bordetella pertussis* adenylate cyclase. *Anal. Biochem.*, **124** : 45-52.
- GORDON V.M., LEPPA S.H. and HEWLETT E.L., 1988 — Inhibitors of receptor-mediated endocytosis block the entry of *Bacillus anthracis* adenylate cyclase toxin but not that of *Bordetella pertussis* adenylate cyclase toxin. *Infect. Immun.*, **56** : 1066-1069.
- GOYARD S., ORLANDO C., SABATIER J.-M., LABRUYÈRE E., d'ALAYER J., FONTAN G., VAN RIETSCHOTEN J., MOCK M., DANCHIN A., ULLMANN A. and MONNERON A., 1989 — Identification of a common domain in calmodulin-activated eukaryotic and bacterial adenylate cyclases. *Biochemistry*, **28** : 1964-1967.
- HEDEGAARD L. and DANCHIN A., 1985 — The *cya* gene region of *Erwinia chrysanthemi* B374: organisation and gene products. *Mol. Gen. Genet.*, **201** : 38-42.
- HEWLETT E.L. and WOLFF J., 1976 — Soluble adenylate cyclase from the culture medium of *Bordetella pertussis* : purification and characterization. *J. Bacteriol.*, **127** : 890-898.
- HOLLAND M.M., LEIB T.K. and GERLT J.A., 1988 — Isolation and characterization of a small catalytic domain released from the adenylate cyclase of *Escherichia coli* by digestion with trypsin. *J. Biol. Chem.*, **263** : 14661-14668.
- KATAOKA T., BROEK D. and WIGLER M., 1985 — DNA sequence and characterization of the *S. cerevisiae* gene encoding adenylate cyclase. *Cell*, **43** : 493-505.
- KAUL R. and WENMAN W.M., 1986 — Cyclic AMP inhibits developmental regulation of *Chlamydia trachomatis*. *J. Bacteriol.*, **168** : 722-727.
- KHANDELWAL R.L. and HAMILTON I.R., 1972 — Effectors of purified adenylate cyclase from *Streptococcus salivarius*. *Arch. Biochem. biophys.*, **151** : 75-84.
- KIELY B. and O'GARA F., 1983 — Cyclic 3'5'-adenosine monophosphate synthesis in *Rhizobium*: identification of a cloned sequence from *Rhizobium meliloti* coding for adenylate cyclase. *Mol. Gen. Genet.*, **192** : 230-234.
- KRUPINSKI J., COUSSEN F., BAKALYAR H.A., TANG W-J., FEINSTEIN P.G., ORTH K., SLAUGHTER C., REED R.R. and GILMAN A.G., 1989 — Adenyl cyclase amino acid sequence : possible channel — or transporter-like structure. *Science*, **244** : 1558-1564.

- LEIB T.K. and GERLT J.A., 1983 — Evidence for a small catalytic domain in the adenylate cyclase from *Salmonella typhimurium*. *J. Biol. Chem.*, **258**: 12982-12987.
- LEPPLA S.H., 1982 — Anthrax toxin edema factor : a bacterial adenylate cyclase that increases cyclic AMP concentrations in eukaryotic cells. *Proc. Nat. Acad. Sci. USA*, **79** : 3162-3166.
- LEPPLA S.H., 1984 — *Bacillus anthracis* calmodulin-dependant adenylate cyclase: chemical and enzymatic properties and interactions with eukaryocytic cells. *Adv. Cyclic Nucleotide Protein Phosphorylation Res.*, **17**: 189-198.
- LOWE D.G., CHANG M-S., HELLMISS R., CHEN E., SINGH S., GARBERS D.L. and GOEDDEL D.V., 1989 — Human atrial natriuretic peptide receptor defines a new paradigm for second messenger signal transduction. *EMBO J.*, **8** : 1377-1384.
- MASSON P., JACQUEMIN J.M. and CULOT M., 1984 — Molecular cloning of the *tsm0185* gene responsible for adenylate cyclase activity in *Saccharomyces cerevisiae*. *Ann. Microbiol.*, **135** : 343-351.
- MASSON P., LENZEN G., JACQUEMIN J.M. and DANCHIN A., 1986 — Yeast adenylate cyclase catalytic domain is carboxyterminal. *Curr. Genet.*, **10**: 343-352.
- MOCK M., CRASNIER M., DUFLOT E., DUMAY V. and DANCHIN A., 1991 — Structural and functional relationships between *Pasteurella multilocidea* and enterobacterial adenylate cyclases. *J. Bacterial.*, **173**: 6265-6269.
- MONNERON A., LADANT D., d'ALAYER J., BELLALOU J., BARZU O. and ULLMANN A., 1988 — Immunological relatedness between *Bordetella pertussis* and rat brain adenylate cyclases. *Biochemistry*, **27** : 536-539.
- ORLANDO C., d'ALAYER J., BAILLAT G., CASTETS F., JEANNEQUIN O., MAZIE J-C. and MONNERON A., 1992 — A monoclonal antibody directed against the catalytic site of *Bacillus anthracis* adenylyl cyclase identifies a 175 kDa counterpart as a novel mammalian brain catalytic subunit. *Biochemistry* (in press).
- PARMA J., STENGEL D., GANNAGE M-H., POYARD M., BAROUKI R. and HANOUNE J., 1991 — Sequence of a human brain adenylyl cyclase partial cDNA. Evidence for a consensus cyclase specific domain. *Biochem. Biophys. Res. Commun.*, in press.
- PAVETO C., EGIDY G., GALVAGNO M.A. and PASSERON S., 1990 — A guanine nucleotide-sensitive, glucagon-stimulated adenylyl cyclase in *Candida albicans* : effect of glucagon on cell morphology. *Biochem. Biophys. Res. Commun.*, **167**: 1177-1183.
- PETERS E.P., WILDERSPIN A.F., WOOD S.P., ZWELEBIL M.J.J.M., SEZER O. and DANCHIN A., 1991 — A pyruvate-stimulated adenylate cyclase has a sequence related to the *fes/fps* oncogenes and to eukaryocytic cyclases. *Mol. Microbiol.*, **5** : 1175-1181.
- PFEUFFER T. and METZGER H., 1982 — 7-O-hemisuccinyl-deacetyl forskolin-sepharose: a novel affinity support for purification of adenylate cyclase. *FEBS Lett.*, **146** : 369-375.
- PFEUFFER E., DREHER R-M., METZGER H. and PFEUFFER T., 1985a — Catalytic unit of adenylate cyclase: purification and identification by affinity crosslinking. *Proc. Nat. Acad. Sci. USA*, **82**: 3086-3090.
- PFEUFFER E., MOLLNER S. and PFEUFFER T., 1985b — Adenylate cyclase from bovine brain cortex : purification and characterization of the catalytic unit. *EMBO J.*, **4** : 3675-3679.
- PFEUFFER E., MOLLNER S., LANCET D. and PFEUFFER T., 1989 — Olfactory adenylate cyclase. Identification and purification of a novel enzyme form. *J. Biol. Chem.*, **264**: 18803-18807.

- READ L.K. and MIKKELSEN R.B.**, 1991 — *Plasmodium falciparum*-infected erythrocytes contain an adenylate cyclase with properties which differ from those of the host enzyme. *Mol. Biochem. Parasitol.*, 45 : 109-120.
- ROGEL A., MELLER R. and HANSKI E.**, 1991 — Adenylate cyclase toxin from *Bordetella pertussis*. The relationship between induction of cAMP and hemolysis. *J. Biol. Chem.*, 266 : 3154-3161.
- SARFATI R.S., KANSAL V.K., MUNIER H., GLASER P., GILLES A-M., LABRUYÈRE E., MOCK M., DANCHIN A. and BARZU O.**, 1990 Binding of 3'-anthraniloyl-2'-deoxy ATP to calmodulin-activated adenylate cyclase from *Bordetella pertussis* and *Bacillus anthracis*. *J. Biol. Chem.*, 265 : 18902-18906.
- SCHULZ S., CHINKERS M. and GARBERS D.L.**, 1989 — The guanylate cyclase/receptor family of proteins. *FASEB J.*, 3 : 2026.
- SINGH S., LOWE D.G., THORPE D.S., RODRIGUEZ H., KUANG W-J., DANGOTT L.J., CHINKERS M., GOEDDEL D.V. and GARBERS D.L.**, 1988 — Membrane adenylate cyclase is a cell-surface receptor with homology to protein kinases. *Nature*, 334 : 708-712.
- TAKAI K., KURASHIMA C., SUZUKI-HOVI, OKAMOTO H. and HAYAISHI O.**, 1974 Adenylate cyclase from *Brevibacterium liquefaciens*. I. Purification, crystallization and some properties. *J. Biol. Chem.*, 265 : 21279-21284.
- TAKEYAMA M., IHARA K., MORIYAMA Y., NOUMI T., IDA K., TOMIOKA N., ITAI A., MAEDA M. and FUTAI M.**, 1990 — The glycine-rich sequence of the B subunit of *Escherichia coli* Mg^{2+} -ATPase is important for activity. *J. Biol. Chem.*, 265 : 21279-21284.
- TAMANOU F., WALSH M., KATAOKA T. and WIGLER M.**, 1984 — A product of yeast *RAS2* gene is a guanine nucleotide binding protein. *Proc. Nat. Acad. Sci. USA*, 81 : 6924-6928.
- THORNE C.B., MOLNAR D.M. and STRANGE R.E.**, 1960 — Production of toxin *in vitro* by *Bacillus anthracis* and its separation into two components. *J. Bacteriol.*, 79 : 450-455.
- THORNER L.K., FANDL J.P. and ARTZ S.W.**, 1990 Analysis of sequence elements important for expression and regulation of the adenylate cyclase gene of *Salmonella typhimurium*. *Genetics*, 125 : 709-717.
- THORPE D.S. and GARBERS D.L.**, 1989 — The membrane form of guanylate cyclase : homology with a subunit of the cytoplasmic form of the enzyme. *J. Biol. Chem.*, 264 : 6545-6549.
- VERNI F. and ROSATI G.**, 1987 — Adenylate cyclase and cyclic AMP in the ciliate *Euploea crassus* : involvement in both cell cycle and sexual reproduction. *J. Exp. Zoology*, 244 : 289-298.
- WÄCHTERSHÄUSER G.**, 1988 — Before enzymes and templates : theory of surface metabolism. *Microbiol. Rev.*, 52 : 452-480.
- WOLFF J., COOK G.H., GOLDHAMMER A.R. and BERKOWITZ S.A.**, 1980 — Calmodulin activates prokaryotic adenylate cyclase. *Proc. Nat. Acad. Sci. USA*, 77 : 3840-3844.
- YOUNG D., RIGGS M., FIELD J., VOJTEK A., BROEK D. and WIGLER M.**, 1989 — The adenylyl cyclase gene from *Schizosaccharomyces pombe*. *Proc. Nat. Acad. Sci. USA*, 86 : 7989-7993.

Cartographie génomique chez les végétaux supérieurs

Fernand VEDEL *

Résumé : Au cours des dix dernières années, le domaine du marquage et de la cartographie génétique s'est considérablement développé chez les végétaux supérieurs, grâce à l'utilisation des marqueurs ADN représentés par les polymorphismes de longueur des fragments de restriction (**RFLP**) auxquels sont venus s'ajouter récemment les polymorphismes de longueur des produits d'amplification par **PCR** à l'aide d'amorces aléatoires (**RAPD**). Actuellement, des cartes génétiques saturées ont été construites pour une douzaine d'espèces. Ces cartes dont la densité des marqueurs s'accroît rapidement présentent un intérêt considérable pour de nombreuses applications en amélioration des plantes, l'étude de l'évolution des génomes, l'isolement de gènes, et la construction de cartes physiques. Les intervalles de résolution des cartes saturées nécessitent la manipulation de grands fragments d'ADN pour le clonage de gènes. Cette dernière est rendue possible par la construction de banques de clones recombinants à l'aide des chromosomes artificiels de levure (**YAC**). Ces banques qui existent maintenant pour plusieurs espèces végétales vont permettre aussi d'établir des cartes physiques des génomes correspondants, comme c'est le cas chez la levure, *Coenorhabditis elegans*, la drosophile, l'homme. Cartes génétiques et cartes physiques d'une espèce donnée pourront ensuite être alignées pour constituer la carte **génomique**.

Mots-clés: cartes génétiques, **RFLP**, **RAPD**, cartes physiques, génétique réverse.

Abstract: Genetic mapping in higher plants has considerably **developped** over the past ten years with the use of DNA markers such as restriction fragment length **polymorphisms** (**RFLP**) and more recently random **PCR** amplified polymorphic DNA (**RAPD**). Saturated genetic maps are available today for a dozen species. These maps the marker density of which is rapidly increasing are extremely useful for : numerous applications in plant breeding, evolutionary studies of **genomes**, gene isolation, construction of physical maps. Resolution intervals of saturated maps require manipulation of large DNA fragments for gene cloning. This is now possible through the construction of recombinant clone libraries in yeast artificial chromosomes (**YAC**). Such libraries which are now available in several plant species will help establish physical maps from corresponding **genomes**, as it is the case in yeast, *Coenorhabditis elegans*, *Drosophila* and man. Genetic maps and physical maps from a given species could then be aligned to build the **genomic** map.

Key words : genetic maps, **RFLP**, **RAPD**, physical maps, reverse genetics.

* Laboratoire de Génétique Moléculaire des Plantes, URA 115, Bâtiment 360, Université de Paris-Sud, 91405 Orsay cedex, France.

Introduction

La carte **génomique** représente une référence unique pour une espèce donnée, permettant d'intégrer les différentes informations fournies par les cartes génétiques et les cartes physiques préalablement établies pour cette espèce. La carte **génomique** est graduée en distance physique (nombre de nucléotides) avec comme repères les localisations des gènes et des différents marqueurs. Les cartes génétiques sont constituées des groupes de liaison dont chacun représente l'ordre linéaire de différents marqueurs génétiques le long d'un chromosome. Ces cartes sont graduées en **centimorgans** (cM), unités proportionnelles à la fréquence de recombinaison entre marqueurs voisins. Les cartes physiques représentent l'ordre linéaire de repères moléculaires ; elles sont construites à partir de banques de clones recombinants. Sur ces cartes, la distance entre deux repères est proportionnelle au nombre de nucléotides qui les séparent.

Actuellement, chez les végétaux supérieurs, les cartes **génomiques** sont représentées presque exclusivement par les cartes génétiques. Cependant, des cartes physiques sont en construction chez plusieurs espèces.

Construction d'une carte génétique

Le principe consiste généralement à croiser deux lignées qui se distinguent par un ensemble de marqueurs et à analyser la ségrégation de ces marqueurs dans les descendances. Au moment de la méiose, les chromosomes non homologues **ségrègent** de façon indépendante ainsi que les gènes localisés sur les différents chromosomes non homologues. Au contraire, les gènes portés par un même chromosome ne **ségrègent** pas indépendamment. Dans ce dernier cas, les fréquences de recombinaison permettent de calculer les distances entre ces marqueurs et de les ordonner dans un groupe de liaison. La construction d'une carte génétique nécessite de disposer d'un matériel végétal génétiquement bien défini, de marqueurs appropriés, d'une méthode d'analyse des résultats des ségrégations (**Helentjaris et al., 1986a** ; **Bernatzky et Tanksley, 1986**).

Le matériel végétal

Quatre types de descendances en ségrégation ont été utilisés pour cartographier les génomes des plantes :

- ensemble de plantes F2 issu de l'autofécondation d'un hybride F1,
- ensemble de plantes provenant d'un **rétrocroisement** (BC) mettant en jeu un hybride F1 et un gamète parental,
- population de lignées recombinantes (RI) obtenues après plusieurs cycles d'autofécondation,
- ensemble de lignées issues **d'haplodiploïdisation** (HD) à partir de la F1.

Si les deux premiers types de descendances en ségrégation ont été les plus utilisés pour la cartographie génétique, les 2 derniers sont plus intéressants car ils permettent entre autre d'obtenir un grand nombre d'individus par génotype. Les lignées RI présentent 2 avantages supplémentaires. D'une part, lorsque l'*homozygotie* est atteinte, elles peuvent être propagées indéfiniment sans ségrégation, ainsi la même population peut être utilisée par différents laboratoires pour la cartographie. D'autre part, les gènes liés ont une plus forte probabilité de recombinaison à cause des méiose multiples nécessaires pour atteindre l'*homozygotie*. Le temps nécessaire pour obtenir les lignées RI constitue leur principal inconvénient. Les lignées RI ont été très utilisées pour la cartographie génétique chez la souris ; chez les végétaux supérieurs elles ont surtout été utilisées pour la cartographie génétique du maïs (Burr et Burr, 1991).

Les lignées HD offrent de moindres possibilités de recombinaison entre marqueurs liés que les lignées RI et de plus, la culture d'anthère induit une importante variabilité.

La localisation de marqueurs sur les chromosomes et la corrélation entre groupes de liaison et chromosomes ont été menées chez plusieurs espèces grâce à l'utilisation de lignées monosomiques, trisomiques et également de lignées d'addition et de lignées de substitution (Helentjaris *et al.*, 1986b ; Young *et al.*, 1987).

Les marqueurs génétiques

Différents types de marqueurs ont servi à établir les cartes génétiques végétales : marqueurs morphologiques (depuis les années 1920), marqueurs moléculaires répartis en marqueurs protéiques (depuis les années 1960) et marqueurs à ADN (depuis 1980). Les marqueurs morphologiques et marqueurs protéiques représentés surtout par les *isoenzymes* possèdent deux inconvénients majeurs, leur nombre réduit et leur sensibilité aux conditions de culture. Les marqueurs à ADN peuvent être distribués en trois catégories : polymorphismes de longueur des fragments de restriction (**RFLP**), polymorphismes de longueur de produits d'amplification par **PCR** à l'aide d'amorces aléatoires (**RAPD**), polymorphismes des séquences répétées (micro et **minisatellites**, séquences **télomériques**, **rDNA**...). Les progrès importants réalisés depuis près de 10 ans dans la cartographie génétique des plantes sont dus à l'utilisation des **RFLP**. Ces progrès vont en s'accélérant avec la **mise** en évidence récente des marqueurs **RAPD** (Williams *et al.*, 1990). Les séquences répétées du type micro et **minisatellites** restent peu utilisées.

Marqueurs **RFLP**

Ils ont pour origine des différences de profils de restriction entre plantes pour une enzyme donnée ; ces différences sont liées soit à la perte (ou au gain) d'un site de restriction soit à des mutations de type insertion-délétion. Leur mise en évidence repose sur la comparaison des profils d'hybridation entre l'ADN cible transféré sur membrane après restriction et un ADN sonde préalablement marqué ; après hybridation, la visualisation de la position des fragments cibles homologues de la sonde est obtenue soit par autoradiographie (si la sonde est radioactive) soit par des techniques bio-