

# THESE

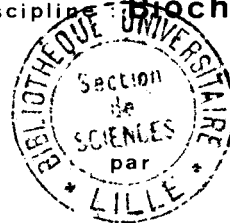
présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir le titre de

DOCTEUR DE TROISIEME CYCLE

Discipline : Biochimie



Carmen FERRAGUT

**CONTRIBUTION A L'ETUDE DU GC% CHEZ LES BACTERIES  
APPLICATION A UN GROUPE D'ENTEROBACTERIES  
APPARENTEES AU GENRE CITROBACTER**

Membres du Jury :	M. GUILLAUME	Président
	M. LECLERC	Rapporteur
	M. KREMBEL	Examineur
	M. ROMOND	Examineur
	M. DE LEY	Membre Invité

Soutenue le 28 Septembre 1976

*Ces pages représentent le bilan de trois années d'activité.*

*On n'y trouvera pas les conclusions définitives qui accompagnent souvent la fin d'un travail de recherche, car, qui peut affirmer que son sujet est épuisé ? Les résultats décrits ne sont que des approches partielles de certains aspects d'un grand problème. Ils sont proposés comme apport scientifique et tentent d'établir des axes de travail. Ils constituent aussi le fruit du développement de techniques rigoureuses. Ces moyens ne peuvent s'acquérir qu'avec de l'imagination et aussi grâce à la volonté d'aboutir et à la conviction que possède le chercheur de suivre une voie fructueuse.*

*Lorsqu'on a franchi une étape, au long de laquelle on est amené à mesurer, dans cet ample domaine de la science, la relativité des connaissances et les limites humaines, il faut parfois arrêter le temps, faire une halte et écrire.*

*Puisse cette thèse avoir pour but de dégager quelques points d'intérêt et d'élargir les chemins sur lesquels d'autres chercheurs essaient de s'avancer. Qu'elle soit le moyen et l'occasion d'exprimer simplement ma gratitude à tous ceux qui m'ont aidée, guidée et conseillée ainsi qu'à ceux qui vont être mes juges.*

TABLE DES MATIERES

INTRODUCTION	1
Première partie : ETUDE BIBLIOGRAPHIQUE	
I - ISOLEMENT DE L'ADN BACTERIEN	3
A - La lyse bactérienne	3
B - Extraction de l'ADN	3
C - Purification de l'ADN	5
D - Conservation	6
II - LE % GC ET LA DENATURATION THERMIQUE DE L'ADN	7
A - Rappels sur la configuration moléculaire de l'ADN	7
B - Différentes méthodologies	8
C - La dénaturation thermique	11
1 - Description du phénomène	11
2 - Facteurs qui influencent la dénaturation	13
a) - La force ionique	13
b) - Action des ions métalliques	14
c) - Le pH	14
d) - Action des composés organiques	15
e) - Action des rayonnements	16
f) - La composition en bases	16
g) - Le poids moléculaire	17
III - SIGNIFICATION DU % GC ET INCIDENCES TAXONOMIQUES	17

## Deuxième partie : ETUDE EXPERIMENTALE

I - MATERIEL ET METHODES	26
A - Souches bactériennes	26
1 - Origine	26
2 - Milieu de culture	26
3 - Conditions de culture	27
B - Extraction de l'ADN	27
1 - Procédé d'isolement	27
2 - Contrôle de pureté	27
C - Dénaturation thermique	28
D - Détermination du Tm	28
E - Tests statistiques	34
II - RESULTATS EXPERIMENTAUX	35
A - Etude comparative des méthodes de détermination du Tm de l'ADN bactérien	35
B - Les coefficients de CHARGAFF	38
III - ANALYSE DES RESULTATS	55
A - Méthodes de détermination du Tm	55
B - Les coefficients de CHARGAFF	57
1 - Corrélation entre taxonomie nu- mérique et analyse du % GC	57
2 - Evaluation et conséquences générales	63
CONCLUSIONS	67
BIBLIOGRAPHIE	

## INTRODUCTION

---

La classification des bactéries a été influencée durant de longues années par des considérations pratiques. Leur étude se résumait à une simple description de caractères morphologiques, physiologiques, biochimiques et sérologiques. La notion d'espèce, artificielle pour le monde bactérien, était tributaire de la complexité et de l'incertitude des résultats et des techniques.

Un grand pas a été franchi avec l'analyse numérique qui a permis une première approche objective de classification. Par l'étude d'un grand nombre de caractères phénétiques, elle établit des "phénons" ou "classes" où les bactéries sont groupées selon un coefficient de similitude. On évite ainsi le caractère subjectif et personnel des interprétations et l'on dispose de données infiniment plus riches et donc plus fécondes.

Pour donner à la classification une valeur phylogénétique comme dans le cas des organismes supérieurs, il faut avoir recours à la biochimie et à la génétique pour comparer les structures moléculaires qui possèdent une "originalité" et une "continuité" (44).

Dans cette perspective trois types d'études peuvent être abordées. La première concerne l'analyse du contenu en bases nucléiques de l'ADN ou % GC ; elle peut traduire une divergence évolutive globale entre deux organismes. Depuis les travaux de CHARGAFF (8) on sait, en effet, que la teneur en guanine et cytosine de l'ADN des bactéries est susceptible de larges variations d'une espèce à une autre.

Des valeurs proches sont l'indice d'une similitude entre ADN ; des valeurs éloignées ne peuvent appartenir à un même groupe phylogénique. Cependant des % GC similaires n'impliquent pas forcément une identité génomique.

La deuxième approche consiste à rechercher les séquences nucléotidiques communes entre 2 génomes : c'est l'hybridation.

La troisième possibilité est offerte par l'examen de la structure des protéines isofonctionnelles (44). On conçoit aisément que les différences existant au sein du génôme soient aussi perçues au niveau du produit de sa traduction.

Le but de notre présent travail concerne l'analyse du % GC chez les bactéries. Etape essentielle dans les méthodologies qui concourent à la classification, son intérêt et ses limites n'ont jamais été parfaitement précisés. Elle constitue donc le premier objectif de notre étude. Cette analyse a été appliquée systématiquement à un groupe de 111 souches d'entérobactéries apparentées au genre *Citrobacter*. Dans un contexte aussi précis, jamais poussé jusqu'à ce stade, il apparaît possible d'une part de définir les informations originales utiles à la classification de ces souches, d'autre part d'extrapoler ces données sur le plan général de la classification bactérienne.



Première Partie : ETUDE BIBLIOGRAPHIQUE

---

## I. - L'ISOLEMENT DE L'ADN BACTERIEN

### A. - LA LYSE BACTERIENNE

Les méthodes de lyse cellulaire doivent permettre d'une part la solubilisation de l'ADN dans le milieu choisi, d'autre part l'inhibition des DNases grâce à la présence d'ions bivalents ou d'agents chelateurs. Le citrate de sodium, souvent employé à concentration assez faible, joue ce double rôle (57). Parfois l'EDTA y est ajouté pour complexer les ions  $Mg^{++}$  et minimiser l'action de ces enzymes.

On opère par voie chimique ou par traitement mécanique. Lorsqu'on s'adresse à des bactéries Gram négatives la méthode aux détergents est la plus utilisée ; l'action pendant 5 à 10 minutes du lauryl sulfate de sodium (SLS) donne généralement des résultats satisfaisants ; ce composé possède aussi la propriété d'être un inhibiteur des DNases (80).

Les difficultés semblent accrues avec les bactéries Gram positives dont la paroi constitue une véritable barrière. Le lysozyme se révèle actif mais son action est parfois lente et son rendement faible.

Parmi les traitements mécaniques, on peut mentionner les ultrasons (80), l'agitation à billes (80), la cellule de FRENCH (41). Dans tous les cas il convient d'éviter la dégradation moléculaire du biopolymère par une méthodologie rigoureuse (66).

### B. - EXTRACTION DE L'ADN

On procède généralement selon la technique de MARMUR (80). Le lauryl sulfate de sodium utilisé permet de séparer les acides nucléiques de leurs composés associés : protéines ou lipo-

protéines. Son mécanisme d'action semble inconnu à l'heure actuelle mais selon NOLL et coll. (95) la chaîne hydrocarbonée du SLS entretrait en compétition avec les ponts hydrogène établis entre l'acide nucléique et les protéines. Cette propriété associée à la présence du groupement sulfate fortement hydrophile serait vraisemblablement la cause de formation de micelles.

Un autre procédé mis au point par SEVAG et coll. (116) emploie le mélange chloroforme-alcool isoamylique comme agent dénaturant des protéines et permet d'obtenir après émulsion et centrifugation trois phases parmi lesquelles la phase aqueuse est enrichie en ADN.

D'autres enfin comme KIRBY (62) utilisent le phénol qui présente le double avantage d'être inhibiteur des nucléases et fortement déprotéinisant ; lorsqu'il est saturé en sels lipophiles, ces caractéristiques se trouvent augmentées (62, 63, 64). Le mélange m-crésol-phénol a l'avantage de pouvoir être refroidi sans provoquer une cristallisation du phénol (65).

Selon SAITO et MIURA (109), le traitement au phénol est une des méthodes de choix pour la déprotéinisation dans le cas de *Bacillus subtilis*. Le pH de 9 est nécessaire pour éviter une contamination en ARN. On peut émettre à ce sujet quelques réserves : des résidus de phénol peuvent interférer dans l'analyse spectrale de l'ADN ; ils doivent être éliminés par voie chimique (40). Ils peuvent aussi causer dans certains cas une perte sélective d'ADN au cours du processus d'isolement (95) et sa présence dans la solution d'ADN extraite peut provoquer une dénaturation partielle de la macromolécule et abaisser ainsi sa température de transition ou  $T_m$  (86).

Actuellement, l'action d'un composé organique aux effets déprotéinisants est le plus souvent précédée par l'activité biologique de la pronase.

### C. - PURIFICATION DE L'ADN

---

L'élimination des résidus d'ARN, restant encore en solution avec l'ADN, est effectuée par l'action de la RNase pancréatique (80). D'après certains, l'utilisation complémentaire de deux RNases conduit à des rendements supérieurs (90, 109). Après la dégradation de l'ARN, il est possible de séparer les fractions protéiques cellulaires restantes et les polysaccharides à l'aide de l'alcool (57, 109).

En 1968, BELLAMY et coll. (5) développent une méthode de séparation des deux acides nucléiques par le bromure de cethyl triméthyl ammonium. DARBY et coll. (20) reprennent ce procédé pour travailler sur l'ADN de *Clostridium welchii* et aboutissent à des résultats satisfaisants. Plus tard, HILL et coll. (55) l'appliquent aux mycobactéries avec un succès identique.

Dans la pratique courante, la purification de l'ADN est obtenue par des étapes successives de dissolutions et de précipitations ; cette méthode est simple et reproductible (123). Pour obtenir une récupération quantitative de l'ADN, il est nécessaire de maintenir sa concentration dans la solution à une valeur supérieure à 5 unités de densité optique (DO) par millilitre à 260 nanomètres (129).

Les informations concernant la séparation des polysaccharides sont peu développées. Selon MARMUR (80), l'isopropanol permet leur élimination subtotale. KIRBY (62, 63) obtient une séparation spécifique avec le 2-méthoxy-éthanol ; en raison de son effet légèrement dénaturant, l'auteur choisit ultérieurement le crésol benzoate.

Pour une purification plus fine, il serait nécessaire de recourir soit à la chromatographie sur colonne de méthyl albumine kieselgur ou d'hydroxyapatite, soit à la centrifugation en gradient de chlorure de césium.

#### D. - CONSERVATION

Dans une étude publiée récemment, ANDO et FUKADA (2) énoncent les différentes méthodes de conservation de l'ADN. Trois alternatives sont proposées :

- les fibres d'ADN sont mises en solution dans un tampon convenable puis congelées à  $- 30^{\circ} \text{C}$  ;

- l'ADN est conservé sous sa forme précipitée dans l'alcool éthylique à 70 % à basse température ;

- l'ADN en fibres est séché et conservé à basse température.

Les observations récentes d'ANDO et de FUKADA (2) ont montré que le séchage de l'ADN, porté jusqu'à un vide de  $10^{-4}$  mm de mercure, provoque des altérations de ses propriétés physico-chimiques ; les valeurs de la viscosité, de la biréfringence d'écoulement, de la sédimentation en gradient de sucrose et du  $T_m$  sont alors abaissées. Ces modifications traduisent selon ces auteurs une perte du poids moléculaire, une réduction de la taille et une légère dénaturation de la double chaîne.

CROMBACH (18) montre que la congélation rapide de la solution d'ADN à  $- 70^{\circ} \text{C}$  et sa conservation à  $- 20^{\circ} \text{C}$  n'entraînent pas de modifications notables de ses propriétés. Pour DE LEY et coll. (29), l'ADN présente une bonne stabilité lorsqu'il est fortement concentré dans la solution saline citrate diluée, refroidi brutalement à  $- 50^{\circ} \text{C}$  puis conservé à  $- 90^{\circ} \text{C}$ .

Pour les études d'hybridation sur support, cet auteur propose la conservation de l'ADN dénaturé, fixé sur filtre et maintenu sous vide à  $+ 4^{\circ} \text{C}$ . Le délai de conservation ne pourrait pas cependant dépasser 20 mois (29).

## II - LE % GC ET LA DENATURATION THERMIQUE DE L'ADN

---

### A. - RAPPEL SUR LA CONFIGURATION MOLECULAIRE DE L'ADN

---

Selon le modèle actuellement admis de WATSON et CRICK (132, 133), l'ADN est formé de deux chaînes de nucléotides enroulées l'une autour de l'autre en hélice double ; la cohésion de cet édifice est assurée par des liaisons hydrogène entre bases complémentaires. Ce modèle a été confirmé par les études cristallographiques à l'aide des rayons X par WILKINS (136).

Les plans des cycles de désoxyribose sont parallèles à l'axe de l'hélice tandis que les bases puriques et pyrimidiques sont des molécules plates qui tendent à s'empiler les unes au-dessus des autres en se disposant perpendiculairement à l'axe de l'hélice (7, 52). L'adénine s'associe à la thymine et la guanine avec la cytosine. Ces arrangements sont les seuls possibles car l'appariement de deux purines occuperait un trop grand volume pour permettre une hélice régulière, alors que deux pyrimidines s'inscriraient dans un volume trop petit (134).

En solution, la conformation de l'ADN natif n'est pas notablement affectée par des changements de force ionique, même importants ; en cela elle diffère d'un polyélectrolyte typique. Ceci est sans doute dû à l'existence de la double hélice qui maintient les groupes phosphate négativement chargés à des intervalles réguliers rigidement fixés. En revanche, l'ADN dénaturé se comporte comme une chaîne polyélectrolytique flexible, enroulée de façon lâche (58).

Suivant le degré d'humidité relative, trois types de configurations, A, B, C, peuvent être rencontrées (38, 137). Dans les solutions d'ADN, la configuration B est la plus représentative et aussi la forme la plus stable lorsque cette molécule se trouve dans un excès d'eau. Il s'agit d'une structure paracristalline avec 90 % d'humidité relative. Il y a dix paires de bases par tour d'hélice et

le plan des nucléotides est perpendiculaire à l'axe de l'hélice. Elle dérive de la configuration A qui comporte 75 % d'humidité relative et se différencie de la structure C, observée avec des sels de lithium pour une humidité relative de 44 %.

#### B. - DIFFERENTES METHODOLOGIES

Les premières études chimiques détaillées de l'ADN sont dues à CHARGAFF et coll. (8). Ils ont montré que le rapport entre bases puriques et pyrimidiques est constant mais que le rapport  $\frac{A + T}{G + C}$  varie avec l'ADN. Ce rapport définit le coefficient de CHARGAFF, exprimé aussi par le % GC ou nombre de molécules de guanine et de cytosine pour 100 molécules de bases dans l'ADN.

Grâce à la technique chromatographique, ces auteurs ont pu accumuler dans un premier temps, un grand nombre de données sur la composition moyenne en bases de l'ADN d'une grande variété d'organismes. Elle comporte d'une part, la séparation sur papier des bases puriques et pyrimidiques libérées par hydrolyse de l'ADN et d'autre part, l'estimation quantitative des bases par l'emploi de la spectrophotométrie en lumière ultraviolette.

Plus tard, cette technique a été remplacée par des procédés physico-chimiques plus précis et reproductibles et il a été possible de déterminer le % GC par centrifugation en gradient densité (126, 128), par dénaturation thermique (81) et par dénaturation à pH acide (39).

La découverte de MESELSON, STAHL et VINOGRAD (87) sur la technique de formation de bandes d'ADN par ultracentrifugation dans un gradient de densité de chlorure de césium, a rendu praticable la détermination de la densité d'échantillons d'ADN de l'ordre du ug.

En 1959, SUEOKA et coll. (126, 128), ROLFE et MESELSON (104) cherchent à mettre en évidence une variation de densité avec des ADN à contenu en guanine et cytosine différent. L'examen systématique d'un grand nombre d'échantillons, leur permet de démontrer que la densité de la bande d'ADN est fonction de son % GC. L'équation établie par ces auteurs (104)

$$\rho_{25^\circ} = 1,658 + 0,100 (\text{GC}) \text{ g.cm}^{-3}$$

où  $\rho$  est la densité et GC la fraction molaire de guanine plus cytosine, traduit la dépendance qui existe entre la densité de l'ADN et le % GC. Une variation de 1 % en guanine et cytosine introduit un changement de densité de 0,00103 unités (104).

Plus tard, SCHILDKRAUT et coll. (112) confirment la relation précédente et déterminent par la méthode des moindres carrés, l'équation :

$$\rho = 0,098 (\text{GC}) + 1,660 \text{ g.cm}^{-3}$$

qui tient compte de la valeur trouvée pour l'ADN d'*Escherichia coli* ( $\rho = 1,710 \text{ g.cm}^{-3}$ ).

La mesure de la densité est facilitée par l'usage d'une relation linéaire entre l'indice de réfraction ( $\eta_D$ ) et la densité ( $\rho$ ) :

$$\rho_{25^\circ} = 10,8601 \eta_{D 25^\circ} - 13,4974$$

et celle-ci est calculée en utilisant la position de l'ADN standard de référence (37).

L'incertitude dans la détermination des densités est de l'ordre de  $0,001 \text{ g.cm}^{-3}$ , ce qui implique une précision de  $\pm 1$  % sur le % GC (112). La reproductibilité obtenue par SCHILDKRAUT et coll. (112) leur permet de considérer cette méthode comme une méthode standard pour évaluer la teneur en bases de l'ADN. Cependant, ces auteurs soulignent les risques d'erreur inhérents à la technique. En effet, le gradient de densité est affecté par le pH, par les impuretés de chlorure de rubidium et par la présence de bases mineures dans l'ADN.



Par rapport aux autres méthodes, celle-ci offre un certain nombre d'avantages. Un seul  $\mu\text{g}$  d'ADN est nécessaire et la purification de la macromolécule n'est pas requise ; elle fournit en même temps des informations sur l'hétérogénéité de composition en bases, sur la distribution uni ou bimodale et sur le poids moléculaire de l'ADN.

Quand on chauffe progressivement une solution d'ADN dans un spectrophotomètre en lumière ultraviolette, on constate un accroissement de la densité optique ou effet hyperchromique. Le  $T_m$  ou température moyenne de transition correspond à la moitié de l'augmentation de l'absorption relative et représente la température à laquelle 50 % de l'ADN est dénaturé.

Les recherches de MARMUR et DOTY (81) ont permis de relier le  $T_m$  de l'ADN à son contenu moyen en guanine et cytosine. L'équation donnée par ces auteurs (81) :  $T_m = 69,3 + 0,41 (GC)$ , pour une concentration en ion  $\text{Na}^+ = 0,195$ , est le résultat d'études comparatives menées entre la méthode chromatographique et la méthode de dénaturation thermique. Une correction est apportée ultérieurement par DE LEY (23) ; l'équation devient :  $GC = (T_m - 69,4) 2,44$ .

Une étude détaillée des paramètres qui interviennent dans la détermination du  $T_m$  sera décrite dans un prochain chapitre. Cette technique a été utilisée dans nos travaux comme moyen d'analyse.

Pour une évaluation quantitative et rapide des bases dans l'ADN, d'autres méthodes ont été mises au point. FREDERICK et coll. (39) montrent que le rapport des absorbances obtenues à 260 nm et à 280 nm atteint une constante à  $\text{pH} = 3$ , caractéristique des proportions molaires des bases. Elle apparaît valable pour des souches dont le % GC est compris entre 40 % et 70 % (23), mais son degré de précision peu élevé ne peut permettre de distinguer deux genres bactériens dont les valeurs en % GC sont proches et elle n'est pas conseillée pour une différenciation fine à l'intérieur d'un même genre.

Une variante décrite par ULITZUR (130), consiste à faire des mesures optiques de l'ADN natif à différentes longueurs d'onde. La méthode est basée sur une corrélation empiriquement démontrée, entre le spectre de l'ADN natif et son contenu en guanine et cytosine. Le rapport des absorbances à 245 nm et 270 nm peut aussi être suffisant pour une estimation du % GC (88).

La plupart des données recueillies à présent sur le % GC de l'ADN bactérien proviennent essentiellement des trois premières méthodes citées (45). Les résultats acquis avec des échantillons identiques d'ADN sont fréquemment comparés par les deux méthodes physiques mais le parallélisme avec le procédé chimique est plus rare.

Il faut souligner que les formules utilisées pour le calcul du % GC à partir du  $T_m$  (77, 81) ou de la densité de l'ADN (126, 128) ont été établies empiriquement à partir de données chimiques sélectionnées : l'interconversion  $T_m - GC$ ,  $\rho - GC$  peut engendrer des incertitudes si l'application des formules est effectuée sans rigueur scientifique, surtout pour des valeurs extrêmes en GC.

C'est pourquoi dans toute étude, les résultats doivent être validés par l'introduction de souches de référence et par le contrôle précis de tous les paramètres. Seule la similitude des informations données par plusieurs méthodes sera finalement crédible.

## C. - LA DENATURATION THERMIQUE

### 1 - Description du phénomène

Les solutions d'ADN sont stables à la température du laboratoire. Lorsqu'elles sont exposées à l'action de températures critiques, de pH extrêmes ou de certains solvants organiques, on assiste à un changement structural de la macromolécule reflété par l'altération de certaines propriétés hydrodynamiques ou optiques. Ce phénomène est la conséquence de la séparation de la double hélice en deux

chaînes monocaténares. On l'appelle dénaturation par analogie avec le processus de dénaturation des protéines.

La structure double hélice est parfaitement ordonnée alors que la structure "random coil" ou chaînes séparées découle de la fusion complète de l'ADN où l'harmonie stéréochimique est totalement détruite. Le phénomène de transition (séparation des chaînes) est habituellement représenté graphiquement par une sigmoïde simple, définie par la densité optique et la température (figure n° 4, p 29).

L'ADN natif est caractérisé par son hypochromicité due à l'arrangement régulier des bases. Chaque plan de bases constitue un dipôle qui absorbe la lumière à 260 nm. Dans la double hélice, les bases sont parfaitement rangées parallèlement entre elles et c'est cette stéréorégularité qui est responsable de l'hypochromicité.

Le chauffage progressif d'une solution saline d'ADN natif de force ionique déterminée engendre la cassure des deux liaisons hydrogène entre l'adénine et la thymine puis celle des trois ponts hydrogène liant la guanine et la cytosine. La double chaîne polynucléotidique se sépare en deux chaînes monocaténares dans lesquelles les dipôles ne sont plus disposés parallèlement mais bout à bout. Ce passage provoque une augmentation de la densité optique de la solution de près de 40 % (figure n° 4, p 29).

D'après DE VOE et TINOCO (33), l'énergie des liaisons hydrogène est nécessaire pour maintenir l'appariement des bases. Elle n'est pourtant pas le facteur qui contribue le plus à la stabilité de la molécule dans son état natif. Les auteurs soulignent l'importance des interactions dipôle-dipôle, dipôle induit-dipôle et celle des forces de London qui relient les bases entre-elles. Il en résulte que l'énergie libre de l'hélice dépend de la composition en bases et de leurs séquences. La nature du solvant influera donc sur la force des interactions.

La transition définie expérimentalement peut être interprétée comme le résultat cumulatif de la fusion des différents fragments d'ADN caractérisés dans leurs séquences par une composition moyenne en bases (99 a). Chacun des fragments est dénaturé à une température critique après avoir acquis la même énergie d'activation (99 b). En termes thermodynamiques, la dénaturation constitue un phénomène enthalpique.

## 2 - Facteurs qui influencent la dénaturation

### a) - La force ionique

La stabilité de la macromolécule d'ADN à caractère polyelectrolytique nécessite qu'une grande proportion de groupements phosphate soient neutralisés ou protégés. Selon MANDEL, MARMUR et coll. (77, 84), le  $T_m$  est dépendant de la force ionique du solvant et de la nature des ions. Cette valeur est approximativement très proche d'une fonction linéaire du logarithme de la concentration en sel ou conductance spécifique (34).

La plupart des auteurs emploient comme solvant la solution saline citrate standard (SSC) (81) dont la concentration en ion  $Na^+$  est égale à 0,195 et dont la conductance spécifique est de  $16 \times 10^{-3} \text{ mho cm}^{-1}$  (34).

Lorsque la molécule d'ADN a une composition en guanine et cytosine élevée, l'ADN est très stable dans la SSC, on s'adresse alors à d'autres solvants ou à des concentrations plus faibles en sel qui déplacent le  $T_m$  ; une dilution au 1/10 de la solution saline citrate standard entraîne un abaissement du  $T_m$  de  $15,4^\circ \text{ C}$  (82).

Parallèlement, une augmentation de la force ionique de la solution accroît la stabilité thermique de l'ADN ; cet effet tend vers un maximum pour des forces ioniques très élevées (84).

Lorsque le perchlorate de sodium de molarité 7,2 ou le méthanol à 50 % sont utilisés comme solvant, on observe un abaissement du  $T_m$  de 30 et 50 degrés par rapport à la solution d'ADN dans la SSC (51).

b) - Action des ions métalliques

EICHORN (35), KATZ (61), MILLER (89), DOVES et coll. (34) se sont intéressés aux interactions entre ions métalliques et ADN. Deux catégories de phénomènes ont pu être observées. Les ions bivalents  $Mg^{2+}$ ,  $Co^{2+}$ ,  $Ca^{2+}$ ,  $Ba^{2+}$ ,  $Mn^{2+}$ ,  $Ni^{2+}$  et  $Zn^{2+}$  sont considérés comme des cations stabilisateurs contribuant par neutralisation des charges négatives des groupements phosphate, à maintenir la structure native de l'ADN.

Les ions  $Mg^{2+}$  et  $Co^{2+}$  agissent à faible force ionique et augmentent le  $T_m$  de 35 à 45° C (34). De même l'ion  $Ag^+$  (0,2 par base) accroît le  $T_m$  de 40° C et élargit le profil de transition. Ce cation peut être utilisé comme réactif pour une dénaturation sélective (34).

Au contraire la présence des ions  $Cu^{2+}$ ,  $Cd^{2+}$  et  $Pb^{2+}$  abaisse le  $T_m$  (35). Ils entraînent la rupture des ponts hydrogène et un déplacement d'électrons au niveau des bases. D'après KATZ (61), les ions  $Hg^{2+}$  provoquent aussi un abaissement du  $T_m$  en formant des complexes réversibles avec les bases de l'ADN et créant un pontage entre les deux brins de l'hélice.

c) - Le pH

Le pH, lorsqu'il est fortement acide ou basique agit directement sur la stabilité de l'ADN, provoquant sa dénaturation (84). Il n'influence pas le  $T_m$  pour des valeurs comprises entre 5,5 et 8,5 (42).

d) - Action des composés organiques

Plusieurs types de composés organiques peuvent modifier le  $T_m$  de l'ADN. Certains comme le formamide, dans des conditions données de force ionique et de température et à une concentration critique (60 % et 65 %) (85) dénaturent l'ADN par rupture des ponts hydrogène. L'urée, la guanidine, le salicylate de sodium ont le même effet (102).

D'après STOLLAR, GROSSMAN et coll. (53, 124) l'addition de formaldéhyde à une solution d'ADN diminue son  $T_m$  mais augmente son effet hyperchromique. Ces changements spectraux résulteraient de l'action de ce composé sur les groupements amine des bases puriques et pyrimidiques rendus accessibles pendant la dénaturation. Celui-ci pourrait en conséquence se révéler utile pour mesurer le degré de dénaturation de l'ADN soumis à des accroissements de température ou à des changements de pH.

D'autres composés de faible poids moléculaire ont une activité stabilisatrice vis-à-vis de l'ADN et augmentent significativement son  $T_m$ . Les diamines de formule générale  $NH_2 (CH_2)_n NH_2$  ont un effet maximal pour  $n = 5$  (73). Comme l'indiquent MAHLER et MEHROTA (73), la variation du  $T_m$  est fonction de la longueur de l'amine et de la composition en bases de l'ADN étudié.

La spermine, la spermidine et le sulfate de protamine causent aussi un accroissement de la température de demi-dénaturation (74). La spermine assure une protection à une molarité de  $10^{-6} - 10^{-5}$ , alors que des concentrations 10 fois plus élevées de spermidine sont nécessaires pour produire le même effet. On attribue ce phénomène aux différences des constantes d'association des amines avec le groupement phosphate de l'ADN (66).

e) - Action des rayonnements

L'exposition de l'ADN à la lumière ultraviolette, aux rayons X ou à l'action photodynamique des colorants (118) conduit à une altération chimique de la macromolécule qui correspond à un affaiblissement de la structure double hélice et se traduit par une baisse du  $T_m$  (84).

f) - La composition en bases

Le degré de stabilité de l'ADN natif aux divers agents dénaturants est lié à sa composition en bases (66, 77, 81, 82). La paire G - C semble être un meilleur donneur et accepteur d'électrons que la paire A - T, entraînant une plus grande stabilité.

Avec la SSC et à pH 7, le  $T_m$  est augmenté de  $0,41^\circ$  C par mole de guanine et cytosine. Les études d'IMMAN et BALDWIN (59) sur deux copolymères synthétiques : dA-T et dA- $\overline{BU}$  montrent qu'à des concentrations faibles en sels, il existe une différence de 9 degrés entre les deux températures de transition, dA- $\overline{BU}$  étant le copolymère le plus stable. Remarquons que cette différence est minimisée lorsque la force ionique est élevée.

En milieu alcalin, dA- $\overline{BU}$  fusionne à un pH inférieur de 1,5 unités à celui qui est nécessaire pour la dénaturation de dA-T. Cette diminution de la stabilité vis-à-vis des alcalis peut être expliquée par la différence de pK entre le bromo-uracile (pK = 8) et la thymine (pK = 9,8) (59). La substitution de la thymine par le bromo-uracile dans l'ADN naturel entraîne une température de fusion plus élevée qui pourrait être attribuée au renforcement des ponts hydrogène dû à l'effet inductif du brome dans le déplacement d'électrons du noyau pyrimidique.

Signalons que la présence de glucose ou de l'hydroxyméthylcytosine dans l'ADN ne déplace pas le  $T_m$  (77).

g) - Le poids moléculaire

La fragmentation d'un échantillon d'ADN ne semble pas avoir de répercussion sur la valeur du Tm. MANDEL et MARMUR (77) ont étudié les courbes de transition de deux échantillons d'ADN de thymus de veau, l'un étant soumis à la sonication. Ils n'ont pas observé de changement essentiel du Tm dans les deux cas.

III - SIGNIFICATION DU % GC ET INCIDENCES  
TAXONOMIQUES

---

Les compositions en bases caractéristiques de l'ADN de la plupart des groupes d'organismes ont été présentées par STANIER et coll. (120) (Fig. 1). Chez les plantes et les animaux cette valeur oscille entre 35 % et 40 % alors que l'on observe parmi les Protistes des variations beaucoup plus larges. Si l'on se rapporte au monde bactérien, ces fluctuations maximales s'étalent de 25 % à 75 % (72, 120) alors que le contenu GC moyen des souches de même espèce présente des valeurs similaires (Fig.2) (120).

On ignore la raison exacte de cette diversité mais il est facile à concevoir que, chez des organismes dont le temps de génération est court, les facteurs de mutation et de sélection ont joué un rôle primordial.

En général, les techniques courantes ne permettent de différencier l'ADN isolé des bactéries sauvages de celui de leurs mutants. Puisque la taille du génôme de ces microorganismes est de l'ordre de  $10^7$ - $10^8$  paires de bases (76), il semble difficile de mettre en évidence par la mesure du % GC, des mutations par délétion ou addition sans provoquer une action létale sur la bactérie.

WEED (135), SPIRIN (119) et GAUSE (47, 48) ont décrit des mutations qui affectaient considérablement le % GC de l'ADN. Ces observations se limitent à des cas où de fortes pressions sélectives



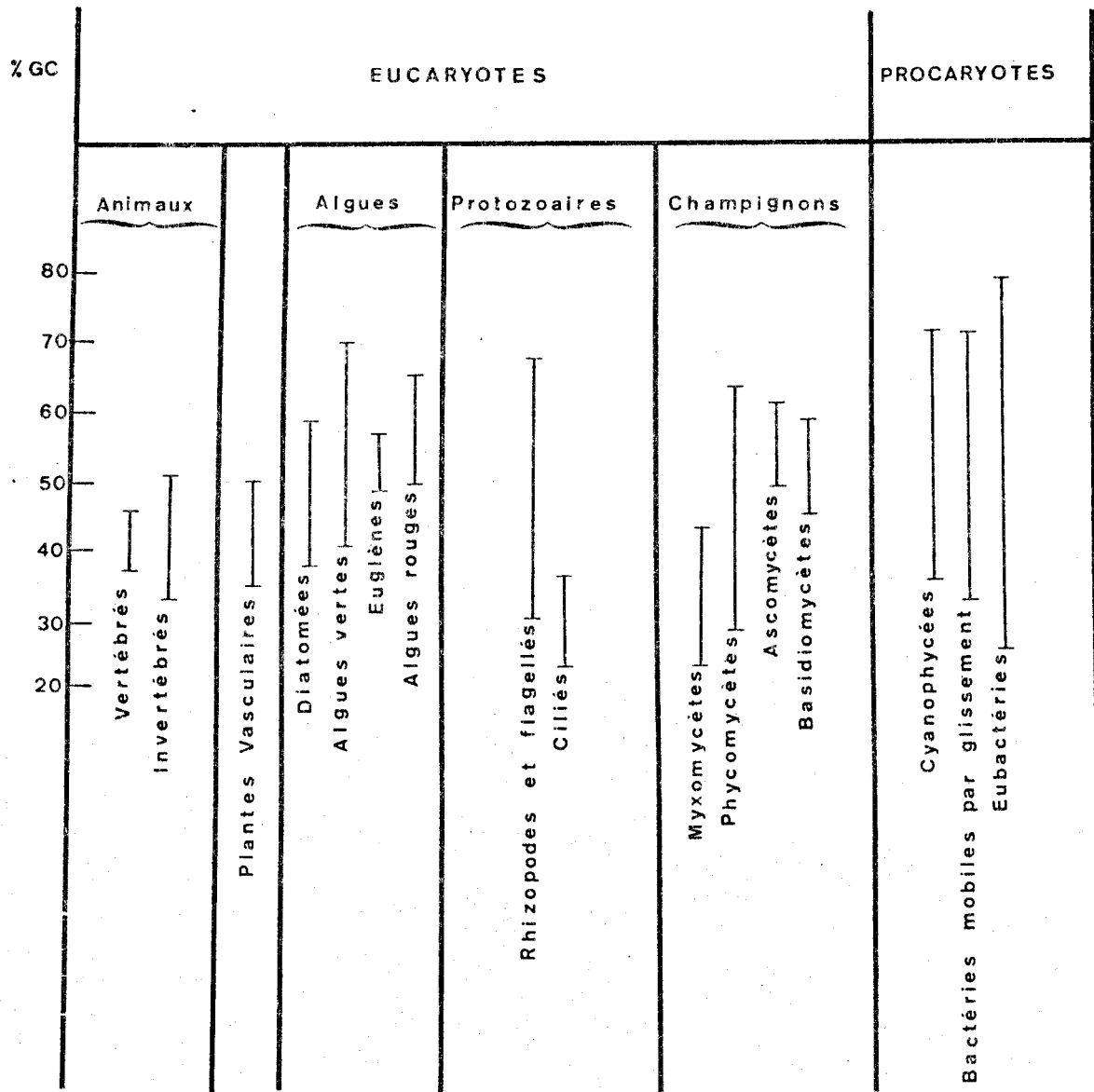


Fig 1 : Composition en bases de l'ADN

caractéristique des principaux groupes

biologiques.



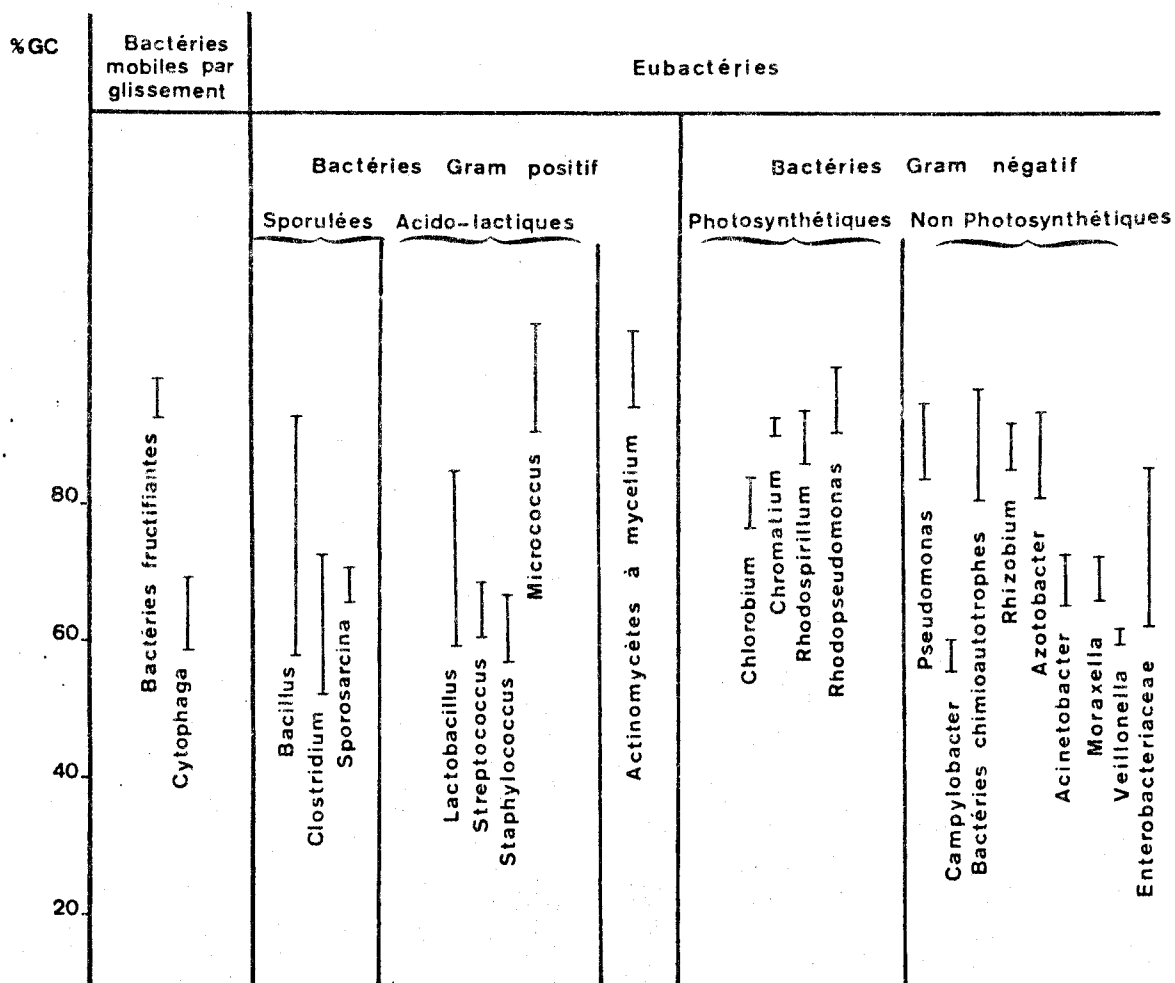


Fig 2 : Composition en bases de l'ADN

caractéristique de certains groupes

bactériens.



sont appliquées et où les mutants isolés ne sont guère semblables à la souche sauvage dont ils dérivent.

MANDEL (76), LEIFSON et coll. (71) mettent en doute la nature de ces mutants et n'éliminent pas l'hypothèse d'une contamination au cours de l'expérimentation.

La description de DE LEY (22) est compatible avec un changement de teneur GC important à la suite de mutations induites par les rayons U.V. Or, le contenu GC est déterminé dans ce cas selon la technique de dénaturation thermique : la différence de  $T_m$  observée peut être due à des altérations au niveau des bases nucléiques ou bien à des changements de structure de l'ADN affectant les forces de liaison interchaînes.

Les expériences rapportées par COX et YANOFSKY (15) montrent que chez une souche d'*E. coli* l'existence d'un gène mutateur (Treffers) provoque un accroissement de la fréquence de mutation de  $10^{-8}$  à  $10^{-5}$ . Celles-ci ont été identifiées comme une transversion de la paire AT vers GC et se traduisent par une variation de 1 % en guanine et cytosine sur les bases totales. Ces mutations ont pu être localisées en partie sur le cistron tryptophane synthétase A mais n'ont pas affecté le fonctionnement de l'enzyme.

D'une façon générale, on peut dire que le % GC est un reflet des acides aminés des protéines cellulaires : chez les bactéries à % GC élevé, les acides aminés seraient codés préférentiellement par des triplets riches en G + C ; dans le cas contraire par des triplets riches en A + T (17, 120).

Plus récemment, STREISINGER et coll. (125), OKADA et coll. (96) montrent que le mode de dégénérescence de la troisième base du codon affecte considérablement le rapport  $\frac{G + C}{A + T}$  et reflète la spéci-

ficité de l'espèce. Ces études indiquent que les bactéries riches en AT utilisent probablement des codons ayant A ou T en troisième position tandis que les espèces du type GC, ont des codons dont la troisième position est occupée par la guanine ou la cytosine. Par ailleurs les lois qui régissent la synthèse protéique ne contredisent pas l'existence d'organismes ayant un % GC différent et une composition semblable en amino acides.

Des recherches récentes (139) ont permis d'établir une relation entre le % GC et l'ARN de transfert : à des valeurs extrêmes en % GC correspond un nombre faible de molécules d'ARN de transfert.

DE LEY et coll. (30) signalent en outre que l'ADN des bactéries à forte activité biochimique présente en général une teneur en GC importante.

L'intérêt du coefficient de CHARGAFF dans la classification bactérienne a été reconnu en premier lieu par LEE et coll. (70). En 1956, ces auteurs apportent les résultats d'analyse des ADN provenant de 60 souches. Leurs travaux font apparaître que le rapport  $\frac{A + T}{C + G}$  est un caractère lié à l'espèce et qu'il peut donc être utilisé comme un critère supplémentaire pour une classification en genre. D'autres travaux (11, 49, 54, 68, 105) ont confirmé son intérêt.

Les résultats expérimentaux accumulés sont en faveur de l'existence d'une relation entre la valeur du % GC de l'ADN d'une espèce bactérienne et sa position taxonomique. Une connaissance méthodique de ce coefficient est donc devenue importante pour comprendre l'évolution ou le maintien du matériel génétique et pour mettre en lumière la parenté entre organismes (21, 37).

Notons qu'un contenu GC identique entre deux souches dont les phénotypes sont semblables constitue un critère nécessaire mais non suffisant pour établir leur identité (21). A l'inverse, des souches dont les valeurs du % GC s'écartent sensiblement, ne sont pas

susceptibles d'homologie génétique étendue (21). Pour SUEOKA (127) deux bactéries dont le contenu GC diffère de 10 % ne possèdent pas de relations phylogénétiques ; dans ce cas il apparaît peu probable que les protéines de ces 2 souches aient la même structure primaire. Par un calcul théorique, DE LEY (25) définit les limites des variations des % GC entre deux organismes qui partagent un nombre appréciable de cistrons identiques ou presque identiques. Il estime qu'une divergence de 20 à 30 % dans le contenu GC indique une absence quasi-totale de séquences nucléotidiques communes aux deux bactéries.

En dépit de son caractère relatif, ce critère a fait l'objet d'une multitude d'études d'ordre taxonomique. Il a contribué à différencier des groupes bactériens dont les études physiologiques et biochimiques aboutissaient à des résultats confus ; les différences phénotypiques sont devenues significatives lorsqu'elles ont été comparées aux données du % GC (107, 113).

Dans certains cas, il a permis d'attirer l'attention sur l'hétérogénéité phylogénétique méconnue des genres et sur l'importance des caractères morphologiques comme facteur évolutif (117).

Les valeurs GC obtenues ont remis en cause certaines subdivisions existantes (9, 46, 75, 79, 114) et ont imposé la création de nouveaux genres (75, 114). Dans d'autres occasions ce critère a imposé le reclassement d'espèces dont la définition restait encore obscure (108) et montré la nécessité de révisions taxonomiques approfondies (111).

Les études parallèles de taxonomie numérique et du % GC paraissent étroitement complémentaires (12, 13, 14, 100, 117, 131). Les bactéries ayant un % GC différent s'insèrent dans des classes différentes ; inversement, les souches présentant le même % GC forment une même classe. Ces correspondances ne sont qu'un argument de plus en faveur de l'une et l'autre méthode.

Dans quelques groupes, la signification du % GC comme indice d'affinité évolutive a été démontrée par échange génétique (60, 83, 110). Les plus grandes fréquences de transformation (115, 120) ont été détectées entre souches de la même espèce et des phénomènes de conjugaison ont pu être observés entre bactéries à parenté étroite. La similitude de composition en bases entre l'ADN des deux partenaires s'est révélée une condition indispensable à ces échanges.

Par contre la transduction et les transferts plasmidiques semblent contredire cette relation entre le % GC et la parenté génétique (113). Ces transferts sont possibles entre bactéries de teneurs GC significativement différentes (113, 120). Cette incidence ne doit pas pour autant surprendre. Dans le premier cas, l'adsorption du phage est essentiellement dépendante de la structure de la paroi cellulaire quelle que soit la nature du matériel génétique, dans l'autre cas, le transfert n'est pas suivi d'intégration dans le génôme de la bactérie réceptrice et de ce fait ne constitue pas nécessairement une relation de parenté. Notons enfin que la composition en bases est fréquemment rapportée parallèlement aux expériences d'hybridation (10, 16, 43, 92, 94, 98, 101, 106, 122). Cette donnée est en effet capitale pour définir les conditions expérimentales de la méthode (31).

La teneur en guanine et cytosine des différents genres de la famille des *Enterobacteriaceae* recouvre une gamme de 37 % à 60 %. La valeur la plus basse est observée pour le genre *Proteus* et les genres *Enterobacter*, *Klebsiella* et *Serratia* atteignent les % GC les plus élevés. Les valeurs caractéristiques de chaque genre sont résumées sur la figure 3.

Les espèces des genres *Citrobacter*, *Escherichia*, *Proteus*, *Salmonella* et *Shigella* présentent des % GC homogènes. Par contre, les espèces identifiées dans les genres *Enterobacter*, *Erwinia*, *Klebsiella* et *Serratia* offrent une plus grande disparité.

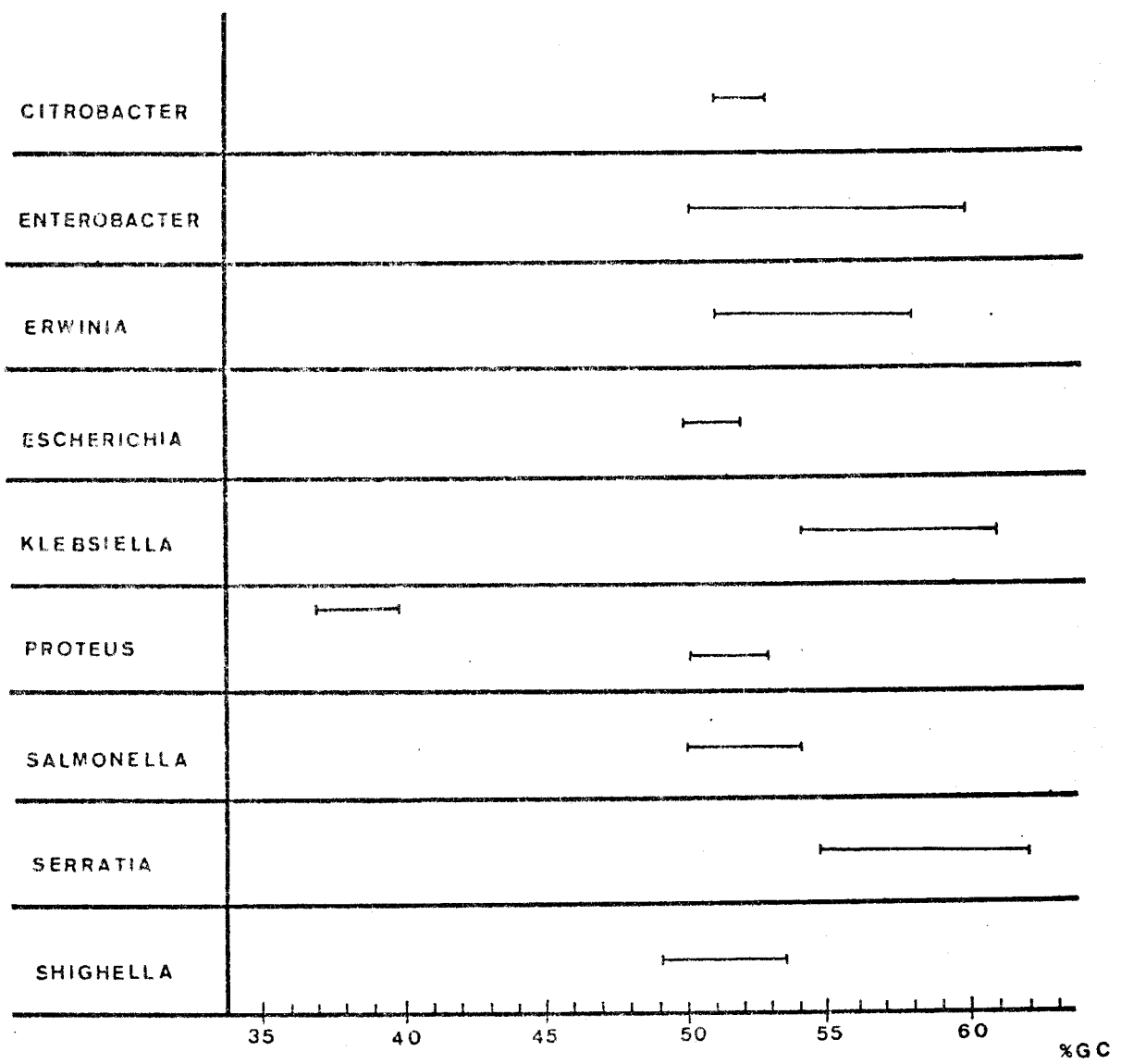


Fig 3 : Composition en bases de l'ADN

caractéristique des principaux genres  
de la famille des Enterobacteriaceae.



Les études effectuées sur le genre *Citrobacter* regroupent les espèces *C. freundii* ( $H_2S^+$  -  $H_2S^-$ ) et *C. diversus*. Actuellement YOUNG et coll. (141) ont créé le genre *Levinea* qui correspond à *C. diversus* et à *C. freundii*, souches  $H_2S^-$  indologènes.

Il existe peu de renseignements sur leur teneur en bases. Des données ont été fournies par MANDEL et ROWND (78), BAPTIST et coll. (4), STARR et MANDEL (121). Les premiers, analysant une souche d'*Escherichia intermedium*, E1-1, trouvent une teneur GC de 51 % ; les observations des seconds rapportent un % GC de 52 pour *C. freundii* ATCC 8455 ; les derniers citent pour *C. freundii* ATCC 8090 une valeur de 52,6 %. RICHARD et coll. (103) ont aussi analysé deux souches appartenant au genre *Levinea* ; les résultats de 56 % et 59 % respectivement pour les espèces *L. malonatica* et *L. amalonatica* leur permettent d'infirmier la validité de ce nouveau genre.

Ces informations sommaires ont motivé cette nouvelle étude des % GC des *Citrobacter* et bactéries apparentées.



Deuxième Partie : ETUDE EXPERIMENTALE

---

## I - MATERIEL ET METHODES

### A. - SOUCHES BACTERIENNES

#### 1 - Origine

Les 111 souches retenues pour notre étude sont isolées de matières fécales humaines, d'eaux usées, d'eaux de surface polluées, d'eaux d'alimentation et de terres vierges. Elles possèdent les caractères suivants :

- absence d'oxydase
- absence de production d'H<sub>2</sub>S et d'acétoïne à 30° C et 37° C
- fermentation de glucose
- présence de  $\beta$ -galactosidase
- utilisation de citrate comme seule source de carbone

Nous avons inséré dans nos travaux, deux souches d'*Escherichia coli* comme souches de référence :

- |                       |           |             |
|-----------------------|-----------|-------------|
| - <i>E. coli</i> K 12 | n° AB1157 | E. Adelberg |
| - <i>E. coli</i>      | n° 10536  | ATCC        |

#### 2 - Milieu de culture

milieu liquide : bouillon nutritif

peptone	10 g
extrait de viande	4 g
NaCl	5 g
eau distillée q.s.p.	1000 ml

pH : 7,2

stérilisation 20 minutes à 120° C

### 3 - Conditions de culture

Pour l'extraction de l'ADN, un litre et demi de bouillon de culture estensemencé avec 200 ml d'une préculture de 24 heures.

Pour se placer dans les conditions de croissance les plus favorables, les cultures sont faites dans chaque cas à la température optimale (37° C pour *E. coli*, 30° C pour les autres bactéries) et en agitation constante pendant 18 h - 24 h.

Chaque culture est suivie d'un contrôle de pureté avant l'extraction de l'ADN.

### B. - EXTRACTION DE L'ADN

#### 1 - Procédé d'isolement

L'ADN est extrait par la méthode de MARMUR (80) avec certaines modifications. Le perchlorate de sodium est remplacé par le chlorure de sodium. L'action de la RNase est poursuivie à 37°C pendant 1 heure. Les fibres d'ADN sont précipitées par un volume d'éthanol à 95 % à - 20° C et enroulées sur une tige de verre. Après légère dessiccation l'ADN est conservé sous sa forme précipitée, dans l'éthanol à 70 % à - 20° C. L'ADN est dissous dans la solution SSC standard pour l'analyse du % GC.

Lorsque la purification s'est révélée insuffisante, la solution d'ADN est soumise à l'action de la pronase ("self-digested" 2 h à 37° C) pendant 16 h à 37° C, à une concentration finale de 0,5 %.

#### 2 - Contrôle de la pureté

La pureté des préparations d'ADN a été vérifiée en calculant les rapports  $\frac{\text{absorption à } 260 \text{ nm}}{\text{absorption à } 280 \text{ nm}}$  et  $\frac{\text{absorption à } 260 \text{ nm}}{\text{absorption à } 230 \text{ nm}}$ .

Ils ont toujours été trouvés au moins égaux aux valeurs admises (1,8 et 2 respectivement) (80).

#### C. - DENATURATION THERMIQUE

Les courbes de dénaturation sont établies avec un spectrophotomètre Unicam SP 1800 ou Beckman Acta II, fonctionnant en double faisceau. L'ADN est placé dans une cellule que l'on chauffe progressivement au moyen d'une résistance électrique externe reliée à un programmeur de température (Accuron Series 2 ou Eurotherm). La densité optique (DO) et la température sont enregistrées simultanément sur un enregistreur "Philips" double voie, ou sur un enregistreur XY "Hwelet-Packard".

Le phénomène de dilution de la solution d'ADN consécutif à l'expansion thermique est corrigé dans tous les cas (77).

Le % GC des souches analysées a été calculé d'après la formule de DE LEY (26).

#### D. - DETERMINATION DU T<sub>m</sub>

Les valeurs de T<sub>m</sub> peuvent être recherchées selon quatre méthodes différentes :

. Soit graphiquement (28, 82, 106) à partir des courbes de dénaturation. Le T<sub>m</sub> est donné par la moyenne des températures correspondant aux points d'intersection des lignes de base et de la tangente de la courbe (figure 4).

. Soit par l'utilisation de papier probit. Cette méthode, préconisée par KNITTEL et coll. (67) analyse le pourcentage d'augmentation de DO en fonction de la température. Le pourcentage d'hyperchromicité est traduit sous forme logarithmique (figure 5).

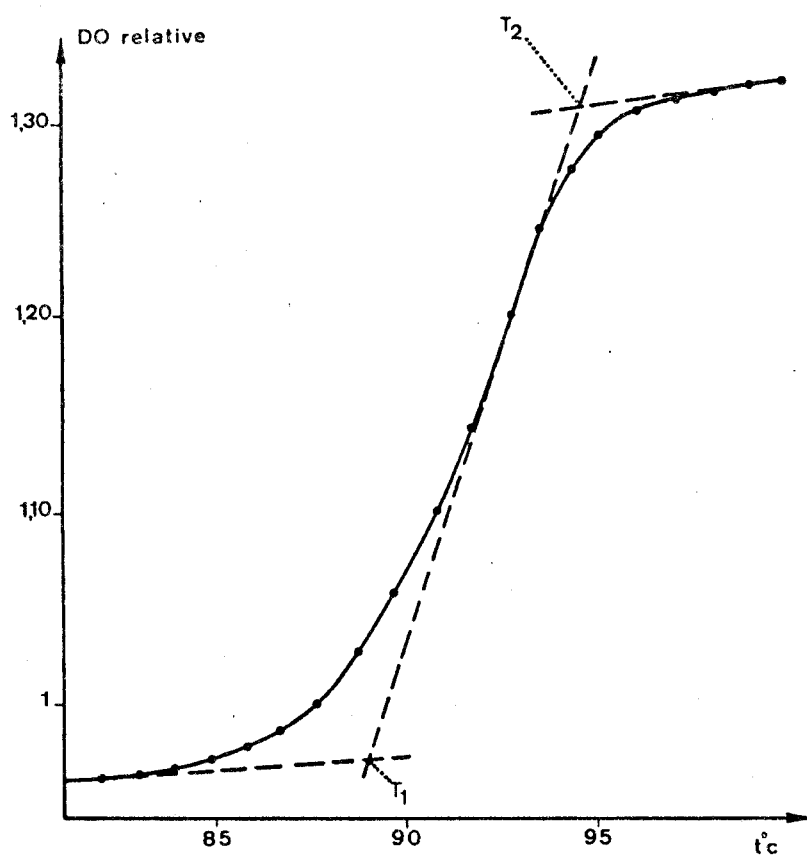


Fig 4 : Courbe de dénaturation de l'ADN d'une souche de *Citrobacter H<sub>2</sub>S<sup>-</sup>* (55)

Détermination graphique du  $T_m$ .  $T_m = \frac{T_1 + T_2}{2}$



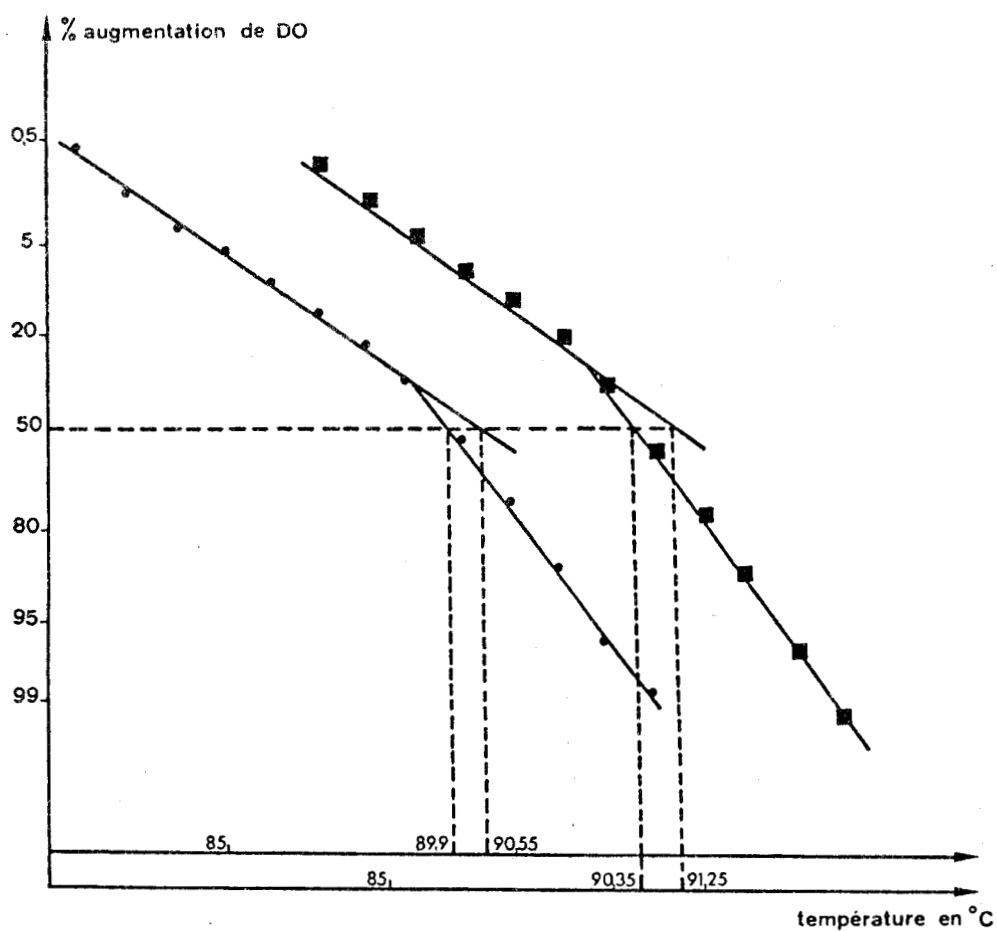


Fig 5 : Détermination du  $T_m$  sur papier probit.

●—● Citrobacter 32

■—■ Citrobacter 75



soit en traçant deux "droites de régression" après avoir étudié l'accroissement de la DO pour une augmentation de température d'un degré, en fonction de la DO moyenne dans l'intervalle de température considéré :

$$\frac{DO_{t + \Delta t} - DO_t}{\Delta t} = f \left( \frac{DO_t + DO_{t + \Delta t}}{2} \right)$$

Puisque le % GC recherché (valeur directement reliée au Tm) est une valeur moyenne globale, et si la distribution du % GC de l'ADN peut être assimilée à une courbe normale ou gaussienne (30), il peut être admis que le point d'intersection des deux droites, dessinées par extrapolation des 2 pentes - ascendante et descendante - de la courbe étudiée, correspond au sommet de celle-ci et permet de déterminer la valeur de la DO à partir de laquelle on calcule la valeur de Tm. (figure 6).

Soit par dérivation des mesures de DO en fonction de la température suivant l'équation mathématique  $\frac{\Delta DO}{\Delta t} = f(t)$  et interpolation linéaire de la médiane (91, 93) (figure 7). Nous considérons des intervalles de température de 1 degré (classe) et nous calculons la somme des variations de DO (fréquence cumulée). Ceci nous permet de connaître la valeur de la fréquence cumulée médiane qui est située dans la classe dont les limites réelles sont tx et tx + 1. Nous effectuons l'interpolation linéaire en appliquant la formule suivante :

$$\frac{\text{limite supérieure} - \text{limite inférieure}}{\text{limite supérieure} - \text{médiane}} = \frac{\text{fréquence cumulée de la limite supérieure} - \text{fréquence cumulée de la limite inférieure}}{\text{fréquence cumulée de la limite supérieure} - \text{médiane}}$$

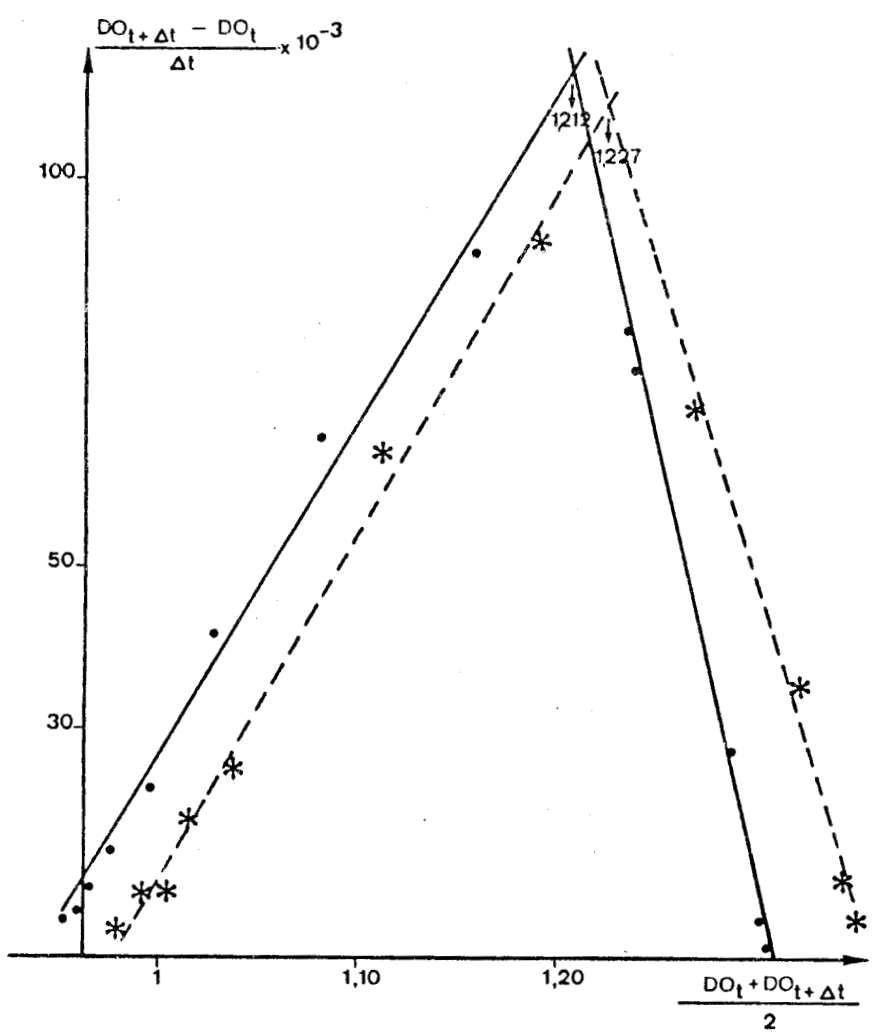


Fig 6 : Etude de la fonction  $\frac{DO_{t+\Delta t} - DO_t}{\Delta t} = f\left(\frac{DO_t + DO_{t+\Delta t}}{2}\right)$  à partir des courbes de dénaturation de l'ADN de deux souches de *Citrobacter H<sub>2</sub>S<sup>-</sup>*. Représentation graphique par «droites de régression»

- *Citrobacter 81*
- \*---\* *Citrobacter 32*



$\Delta t$ : variation de température pour l'intervalle considéré (1°C)



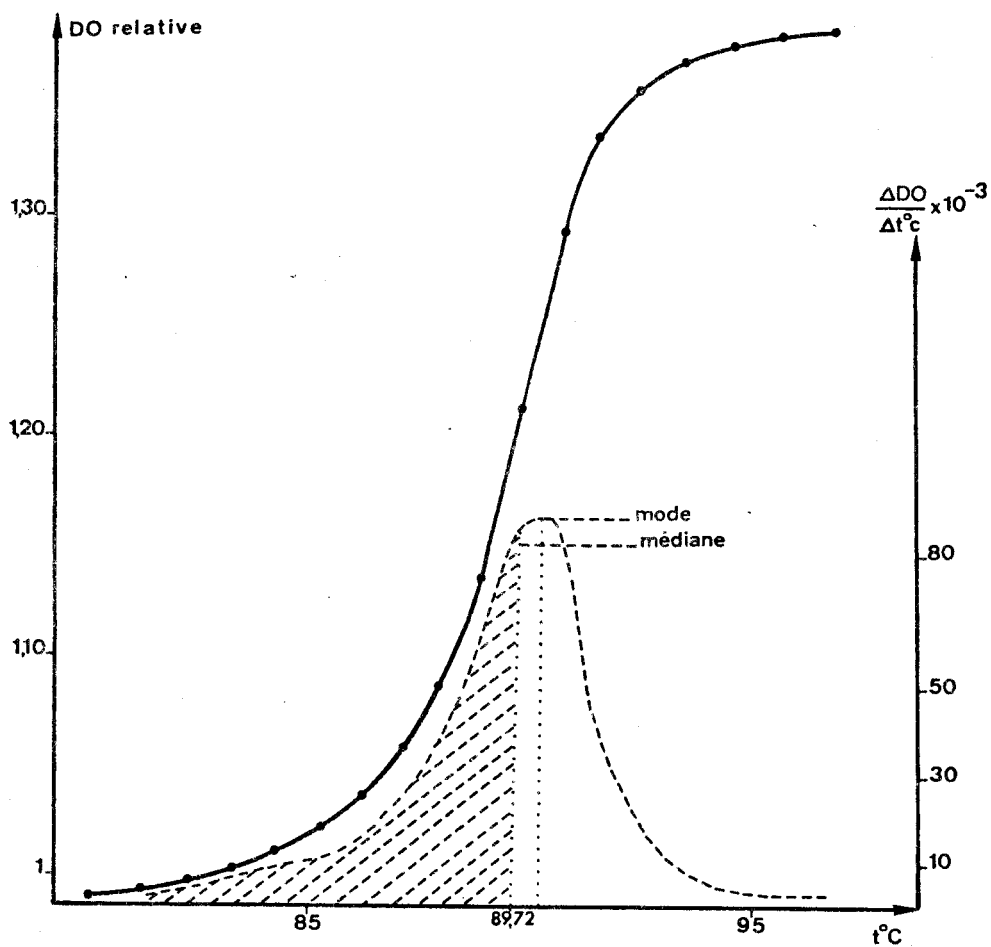


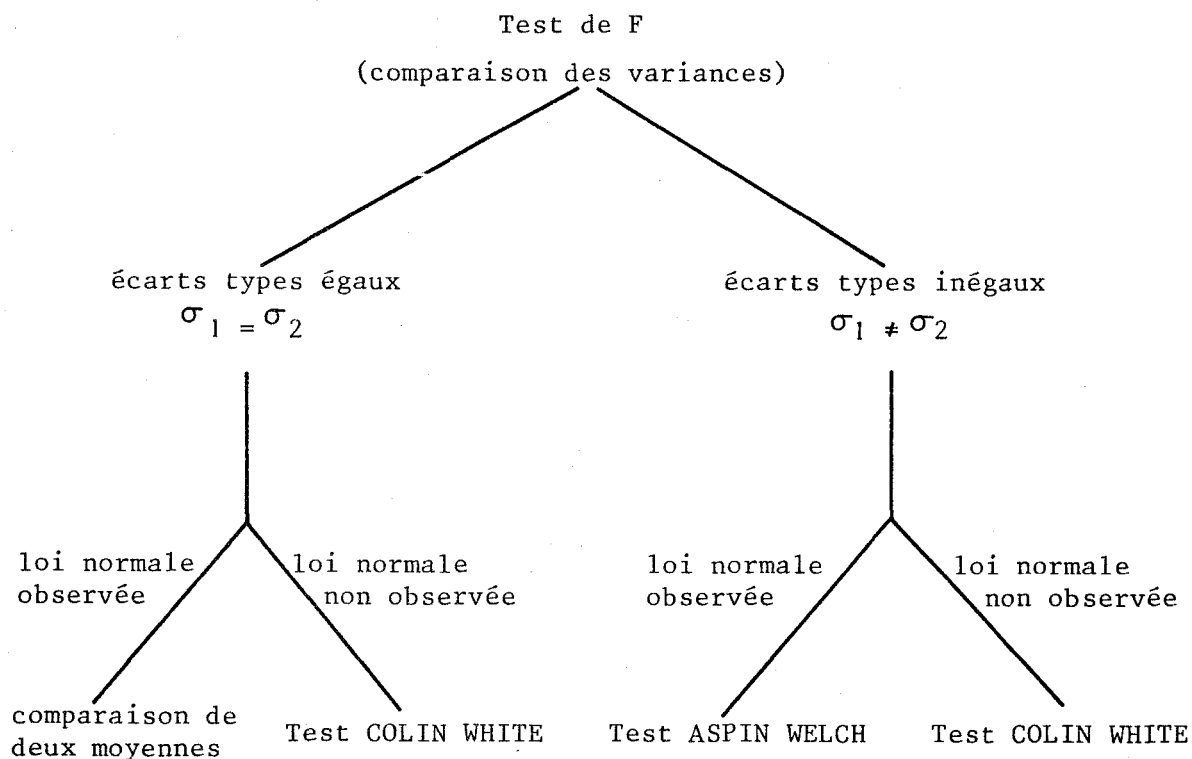
Fig 7: Courbe de dénaturation (•—•—•) et dérivée de la courbe de dénaturation (-----) de l'ADN natif de *E. coli* ATCC n°10536. Interprétation graphique de la médiane.  $\Delta t$  correspond à un intervalle de 1°. La médiane sépare la surface située sous la courbe de fusion différentielle en deux moitiés égales. Elle indique la valeur du  $T_m$ .



E. - TESTS STATISTIQUES (1a, 1b)

Les tests utilisés portent sur la comparaison de deux séries d'observations.

Nous avons calculé en premier lieu l'égalité ou non des écarts types entre chaque classe ; ensuite, nous avons appliqué suivant le schéma ci-dessous, le test statistique adéquat aux populations à comparer.



## II - RESULTATS EXPERIMENTAUX

Les compositions en bases de l'ADN de 111 souches initialement apparentées au genre *Citrobacter* ont été calculées à partir des données du  $T_m$  après dénaturation thermique. Le choix de cette technique a été déterminé par le grand nombre de souches à analyser.

Compte tenu des variations en % GC entraînées par un faible décalage du  $T_m$ , il a semblé opportun dans une première étape d'approfondir les diverses méthodes qui nous permettent de mieux définir ce paramètre ; ensuite nous avons regroupé et interprété les résultats à la lumière de l'analyse statistique et des travaux réalisés précédemment en taxonomie numérique (50).

### A. - ETUDE COMPARATIVE DES METHODES DE DETERMINATION DU $T_m$ DE L'ADN BACTERIEN

Quatre méthodes de détermination du  $T_m$  ont été étudiées et comparées à propos des souches regroupées dans le tableau I. Le nombre de mesures effectuées y est indiqué.

Nous avons négligé pour cette étude comparative le volume d'expansion thermique de la solution d'ADN. Le facteur de correction, qui intervient de la même façon pour toutes les méthodes, apparaît dans ce cas négligeable.

La figure 4 illustre la détermination graphique du  $T_m$  (p 29). L'évaluation du maximum d'hyperchromicité est parfois rendue difficile par une certaine dérive qui accompagne le palier final de la courbe de dénaturation, et qui fait apparaître une absence de parallélisme à ses deux extrémités. Ce phénomène, qui a été observé avec l'ADN

SOUCHES	METHODES	Tm moyen	ECART TYPE	I.C. pour Pe = 0,05	% GC moyen	I.C. pour Pe = 0,05
<i>E. coli</i> K 12 (16 répliques)	A	89,8	0,2	+ 0,11	49,8	+ 0,26
	B	90	0,31	+ 0,16	50,2	+ 0,40
	C	90,3	0,3	+ 0,16	51,1	+ 0,39
	D	89,8	0,28	+ 0,15	49,7	+ 0,27
<i>E. coli</i> TCG n° 10536 (14 répliques)	A	89,9	0,47	+ 0,27	49,9	+ 0,66
	B	90	0,26	+ 0,15	50,2	+ 0,37
	C	90,4	0,30	+ 0,17	51,3	+ 0,42
	D	89,7	0,21	+ 0,12	49,5	+ 0,3
<i>Citrobacter</i> H <sub>2</sub> S <sup>-</sup> (81) (4 répliques)	A	90,4	0,18	+ 0,29	51,3	+ 0,70
	B	90,6	0,23	+ 0,37	51,7	+ 0,9
	C	91	0,26	+ 0,41	52,7	+ 1
	D	90,3	0,13	+ 0,21	51	+ 0,52
<i>Citrobacter</i> H <sub>2</sub> S <sup>-</sup> (75) (5 répliques)	A	90,5	0,24	+ 0,30	51,5	+ 0,74
	B	90,9	0,29	+ 0,36	52,5	+ 0,88
	C	91	0,30	+ 0,37	52,7	+ 0,89
	D	90,4	0,20	+ 0,23	51,2	+ 0,57
<i>Citrobacter</i> H <sub>2</sub> S <sup>-</sup> (55) (4 répliques)	A	91,5	0,52	+ 0,82	54	+ 2,01
	B	91,5	0,61	+ 0,98	54	+ 2,39
	C	-	-	-	-	-
	D	91,3	0,48	+ 0,77	53,5	+ 1,87
<i>Citrobacter</i> H <sub>2</sub> S <sup>-</sup> (82) (4 répliques)	A	89,6	0,34	+ 0,55	49,2	+ 1,33
	B	89,7	0,41	+ 0,66	49,5	+ 1,61
	C	89,9	0,41	+ 0,65	50,1	+ 1,58
	D	89,5	0,35	+ 0,56	49,1	+ 1,37

## ETUDE STATISTIQUE DES DETERMINATIONS DU Tm

CHEZ DEUX SOUCHES DE "*E. coli*" ET QUATRE SOUCHES DE *Citrobacter*

I.C. = intervalle de confiance pour probabilité d'erreur (Pe)

Méthodes d'études : A = détermination graphique  
 B = utilisation du papier probit  
 C = "droites de régression"  
 D = interpolation linéaire de la médiane



de certaines souches, est reproductible pour tous les essais répétés ; il pourrait être la cause des fluctuations apparues dans certains résultats (tableau I, *E. coli* ATCC). La reproductibilité de cette méthode est généralement satisfaisante (tableau I).

Dans la figure 5 (p 30) les courbes de dénaturation tracées sur papier probit ne se traduisent pas en une seule droite ; elles mettent en évidence une distribution non normale des bases le long de la molécule d'ADN. Dans le cas d'une répartition normale, la valeur de  $T_m$  serait située sur l'abscisse en un point correspondant sur l'axe des ordonnées à une augmentation de 50 % de  $D_0$ . Dans le cas présent, la valeur de  $T_m$ , correspondant au point 50 % d'hyperchromicité, est définie par rapport aux deux segments de droite. Elle doit être calculée en faisant la moyenne des deux valeurs obtenues sur l'axe des abscisses par la projection du point 50 % d'hyperchromicité sur les deux segments de droite. L'erreur sur la valeur ainsi calculée du  $T_m$  sera d'autant plus faible que les deux segments de droite tendront vers une droite unique. Les écarts types importants qui apparaissent dans les valeurs moyennes des expériences répétées (tableau I) proviennent donc du phénomène ci-dessus décrit.

Lorsque les mesures de  $D_0$  sont représentées par des "droites de régression" respectivement ascendante et descendante, nous observons que les valeurs de ces pentes sont différentes (figure 6 - p 32) . Il est alors faux de définir le  $T_m$  par la mesure de la  $D_0$  correspondant au point d'intersection des deux droites considérées. L'erreur commise serait d'autant plus importante que l'asymétrie des deux droites par rapport à l'axe central est plus prononcée.

Les valeurs de  $T_m$  obtenues par cette méthode sont différentes de celles des autres méthodes par des écarts pouvant aller de  $0,2^\circ\text{C}$  à  $0,7^\circ\text{C}$  (tableau I).

La mesure du  $T_m$  par l'étude de la fonction  $\frac{\Delta DO}{\Delta t} = f(t)$

suivie de l'interpolation linéaire de la médiane est illustrée sur la figure 7 (p 33). Elle fait apparaître un pic unique, mais l'absence de coïncidence du mode et de la médiane. La valeur médiane devient la seule mesure rigoureuse de tendance centrale divisant l'aire de dispersion délimitée par la courbe en deux moitiés égales. Elle indique directement la valeur du  $T_m$ . La comparaison des résultats rapportés dans le tableau I montre que cette méthode est la plus reproductible. Le traitement statistique (tableau I) des données obtenues avec les souches témoins et les souches de *Citrobacter* 75, 81, montre que l'estimation de l'erreur sur le % GC est inférieure à  $\pm 1$  %. La précision obtenue est pourtant inférieure avec les souches de *Citrobacter* 32 et 55 (p 36).

Pour les souches de référence, nos résultats sont en accord avec ceux présentés dans la littérature (3, 6, 82, 106).

#### B. - LE COEFFICIENT DE CHARGAFF

Sur le tableau II nous présentons l'ensemble des résultats obtenus. Chacun de ceux-ci est la moyenne d'au moins trois analyses. La reproductibilité des valeurs des % GC, pour une même préparation d'ADN est comprise entre  $\pm 0,2$  % et  $\pm 1,8$  %, pour une probabilité d'erreur de 0,05.

L'étude de 206 caractères des 111 souches analysées en taxonomie numérique (50) et la comparaison de leurs similitudes, a permis d'établir le dendrogramme de la figure 9 b (p 46). Sur ce dendrogramme on peut distinguer six classes :

N° Souche	Nombre de mesures	Tm moyen	Ecart type	I.C. pour Pe = 0,05	% GC moyen	I.C. pour Pe = 0,05
1	6	89,45	0,71	0,75	49,9	1,82
2	4	89,47	0,28	0,45	49	1,09
3	6	90,66	0,6	0,64	51,9	1,55
4	4	91,1	0,16	0,25	52,9	0,61
5	5	90,75	0,58	0,72	52,1	1,77
6	5	90,04	0,24	0,3	50,3	0,72
7	4	90,56	0,23	0,37	51,6	0,89
8	5	90,05	0,25	0,31	50,4	0,76
9	4	90,83	0,27	0,43	52,3	1,06
10 *	11	90,25	0,79	0,56	50,9	1,37
11	3	90,78	0,06	0,15	52,2	0,36
12	4	90,56	0,32	0,50	51,6	1,23
13	4	90,48	0,34	0,54	51,4	1,32
14	5	90,38	0,51	0,64	51,2	1,56
15	5	90,65	0,52	0,65	51,8	1,58
16	5	90,28	0,61	0,76	50,9	1,86
17	4	91,22	0,27	0,42	53,3	1,03
18	5	91,01	0,42	0,52	52,7	1,26
19	4	90,65	0,32	0,51	51,9	1,24
20	4	90,3	0,28	0,44	51	1,08
21	4	90,81	0,31	0,49	52,3	1,21
22	3	91,33	0,03	0,08	53,5	0,20
23	3	90,97	0,20	0,5	52,6	1,23
24	4	91,04	0,38	0,6	52,8	1,46
25	5	89,67	0,3	0,35	49,4	0,85
26	3	89,91	0,19	0,46	50	1,13

BIS  
LILLE

27	4	91,17	0,28	0,44	53,1	1,08
28	3	90,7	0,08	0,21	52	0,5
29	4	90,8	0,35	0,55	52,2	1,35
30	4	92,3	0,18	0,29	55,9	0,72
31	4	89,56	0,23	0,36	49,2	0,88
32	4	89,51	0,33	0,53	49,1	1,31
33	3	90,52	0,06	0,15	51,5	0,37
34 *	3	91,29	0,08	0,20	53,4	0,5
35	7	88,41	0,40	0,37	46,4	0,9
36	4	90,49	0,36	0,57	51,4	1,39
37	3	90,77	0,1	0,24	52,2	0,59
38	5	89,32	0,5	0,62	48,6	1,51
39	6	89,11	0,45	0,48	48,1	1,17
40	5	90,67	0,48	0,6	51,9	1,46
41	4	89,73	0,15	0,25	49,6	0,6
42	3	89,57	0,07	0,17	49,2	0,42
43	4	89,75	0,29	0,47	49,7	1,14
44	6	89,34	0,41	0,43	48,7	1,06
45	4	90,83	0,27	0,43	52,3	1,06
46	4	89,47	0,14	0,22	49	0,54
47	6	90,62	0,59	0,62	51,8	1,51
48 *	3	90,77	0,1	0,24	52,1	0,59
49	4	90,46	0,08	0,13	51,4	0,32
50	4	90,4	0,21	0,33	51,3	0,81
51	3	90,15	0,01	0,03	50,6	0,07
52	4	89,39	0,38	0,60	48,8	1,46
53	4	92,21	0,24	0,38	55,6	0,93
54	5	90,16	0,43	0,53	50,7	1,3
55	6	91,6	0,58	0,61	54,2	1,49
56	4	90,08	0,2	0,31	50,4	0,76





57	4	89,71	0,1	0,16	49,5	0,40
58	3	90,84	0,08	0,2	52,3	0,48
59	6	90,55	0,57	0,60	51,6	1,46
60	4	90,97	0,21	0,34	52,6	0,82
61	4	89,79	0,2	0,32	49,7	0,78
62	5	90,6	0,35	0,43	51,7	1,06
63	6	90,86	0,48	0,5	52,4	1,23
64	4	91,4	0,4	0,59	53,7	1,44
65	7	89,77	0,8	0,75	49,7	1,83
66	4	91,1	0,42	0,66	52,9	1,62
67	4	91,54	0,23	0,37	54	0,9
68	4	90,97	0,44	0,7	52,6	1,7
69	4	89,33	0,4	0,64	48,6	1,56
70	7	90,65	0,58	0,54	51,9	1,32
71	5	90,58	0,57	0,71	51,7	1,74
72	5	91,75	0,46	0,58	54,5	1,4
73	4	90,13	0,03	0,05	50,6	0,13
74	4	90,73	0,13	0,20	52,1	0,49
75	9	90,36	0,17	0,13	51,1	0,43
76 *	5	91,71	0,28	0,35	54,4	0,84
77	5	89,03	0,47	0,58	47,9	1,42
78	4	89,62	0,05	0,09	49,3	0,21
79 *	5	91,34	0,13	0,16	53,5	0,4
80 *	7	90,24	0,67	0,62	50,8	1,51
81	4	90,33	0,15	0,23	51,1	0,56
82	4	89,8	0,2	0,33	49,8	0,80
83	4	89,48	0,07	0,11	49	0,28
84	5	88,99	0,4	0,51	47,8	1,24
85	4	89,06	0,32	0,51	48	1,24
86	4	88,7	0,15	0,24	47,1	0,58



87	4	88,48	0,25	0,39	46,6	0,96
88 *	7	89,17	0,57	0,52	48,2	1,28
89	4	90,33	0,11	0,17	51,1	0,42
90	4	89,07	0,35	0,56	48	1,37
91	5	91,45	0,24	0,3	53,8	0,72
92	4	90,04	0,19	0,3	50,4	0,74
93	5	90,52	0,33	0,42	51,5	1,02
94	5	89,08	0,3	0,38	48	0,92
95	4	90,5	0,38	0,61	51,5	1,49
96 *	4	89,73	0,26	0,42	49,6	1,02
97	6	88,55	0,57	0,6	46,7	1,47
98	6	90,27	0,12	0,14	50,9	0,29
99	6	90	0,58	0,61	50,3	1,49
100	4	90,25	0,34	0,54	50,9	1,32
101	4	90,18	0,38	0,6	50,7	1,46
102	4	90,35	0,24	0,38	51,1	0,94
103	4	90,4	0,36	0,58	51,2	1,4
104	4	90,65	0,33	0,53	51,9	1,29
105	5	89,47	0,3	0,37	49	0,9
106	4	90,91	0,28	0,44	52,5	1,08
107	4	90,08	0,25	0,39	50,5	0,95
108 (E.A.)	5	90,52	0,54	0,67	51,5	1,64
109 (E.A.)	3	91,19	0,14	0,35	53,2	0,85
110 (E.A.)	4	89,99	0,18	0,28	50,2	0,68
111 **	3	90,06	0,07	0,17	50,4	0,41

Tm et % GC DE L'ADN DE 111 SOUCHES APPARENTEES AU GENRE *Citrobacter*



\* Résultat déduit de deux ou trois extractions

\*\* *Citrobacter intermedium* de WERKMAN et GILLEN ATCC 6750

E.A. *Enterobacter agglomerans*

Classe A	:	<i>Levinea</i>
Sous-classe A <sub>1</sub>	:	<i>Levinea malonatica</i>
Sous-classe A <sub>2</sub>	:	<i>Levinea amalonatica</i>
Classe B	:	<i>Citrobacter</i> H <sub>2</sub> S -
Classe C	:	Entérobactéries inclassables
Classe D	:	Entérobactéries inclassables
Classe E	:	Entérobactéries acétoïne - proches de <i>Enterobacter cloacae</i>
Classe F	:	Entérobactéries inclassables

Afin de rendre plus apparentes les valeurs du coefficient de CHARGAFF obtenues, nous transcrivons celles-ci sous forme de diagrammes (nombre de souches en fonction de leur % GC). Ils regroupent d'une part le % GC de l'ensemble des souches étudiées (figure 8), d'autre part, le % GC des six classes différenciées en taxonomie numérique (figures 9a, 9b).

La figure 8 nous montre l'allure générale de la population examinée vis-à-vis de la teneur en guanine et cytosine de l'ADN. Le pourcentage GC s'étend de 46 à 56 entre souches extrêmes et 59 % des souches étudiées ont un contenu en guanine et cytosine compris entre 50 % et 53 %.

Les différents diagrammes qui regroupent les souches des six classes, fournissent une première évidence de similitude à l'intérieur de chaque groupe (figure 9a). Entre ces classes, 5 souches restent isolées ; nous signalons dans le tableau II leur % GC. Nous ne tenterons pas de donner des précisions sur ces souches solitaires ; toute interprétation nous paraît illusoire. Nous nous limiterons à l'étude et à la comparaison des classes.

Les premières déductions peuvent être faites en examinant les résultats reportés sur le tableau III et en considérant les

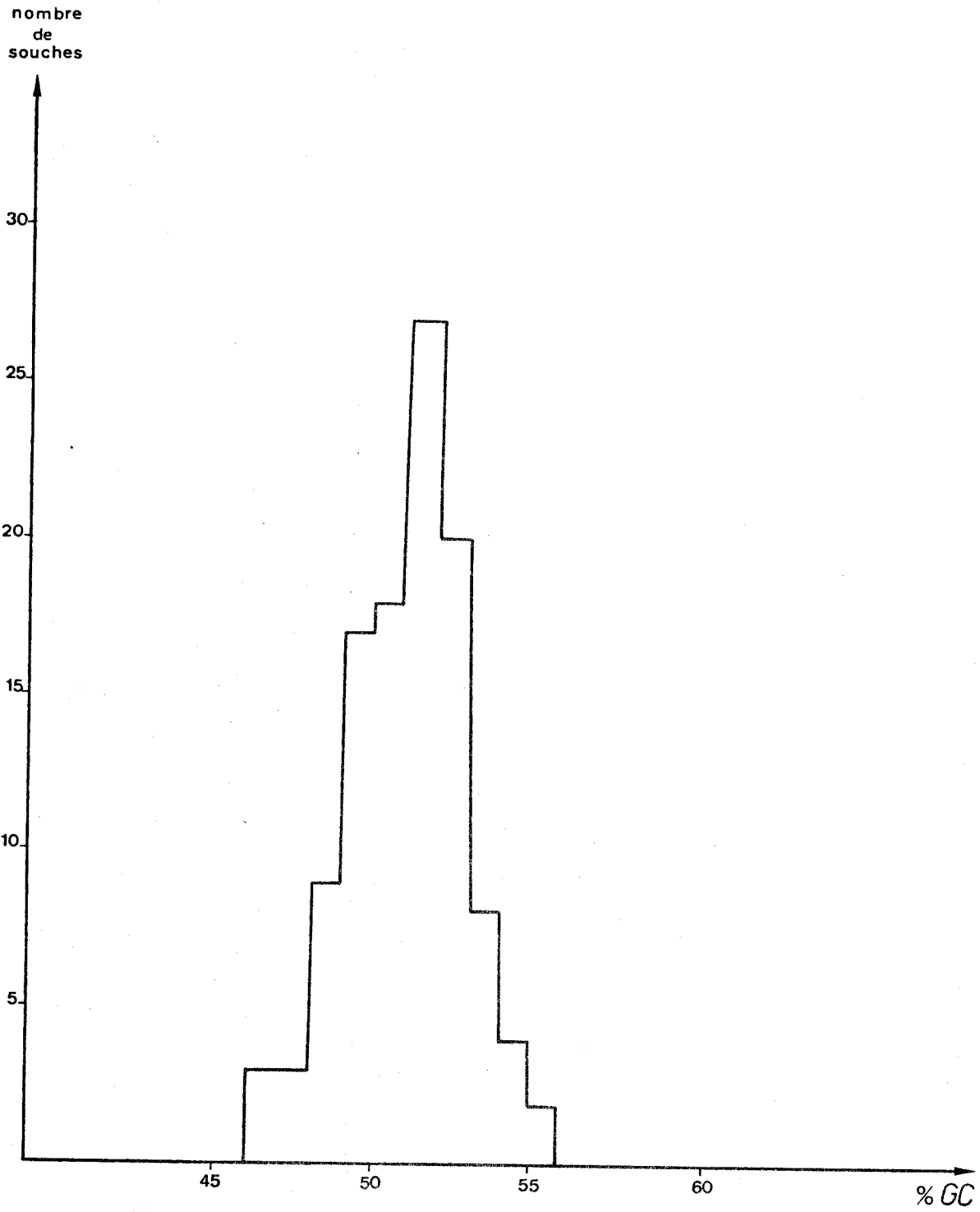
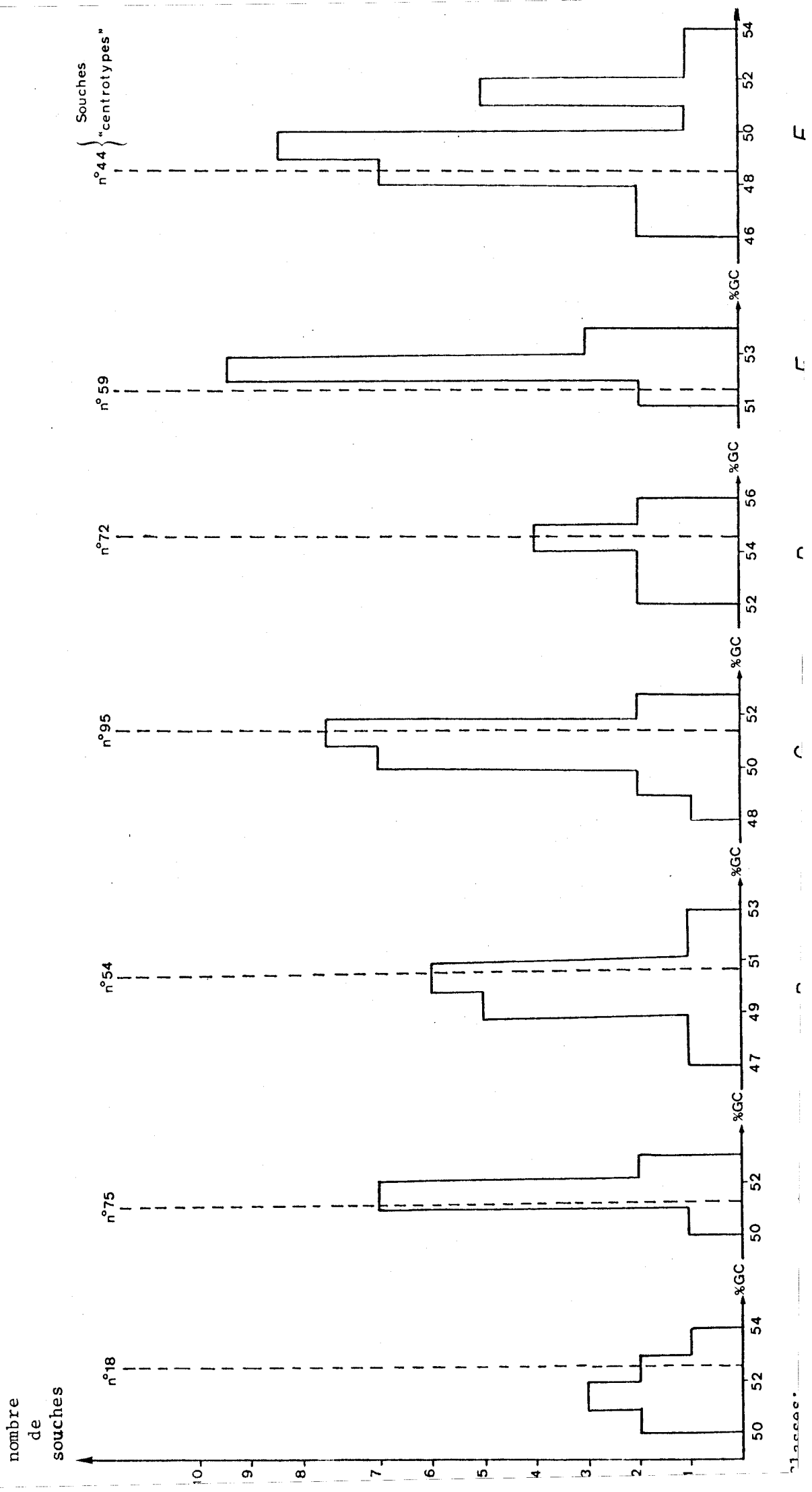


Fig 8 : Histogramme de la distribution des bactéries en fonction du contenu GC de leur ADN





Fig 9a : Histogramme de la distribution des souches en fonction du %GC de leur ADN, suivant les classes définies en taxonomie numérique.



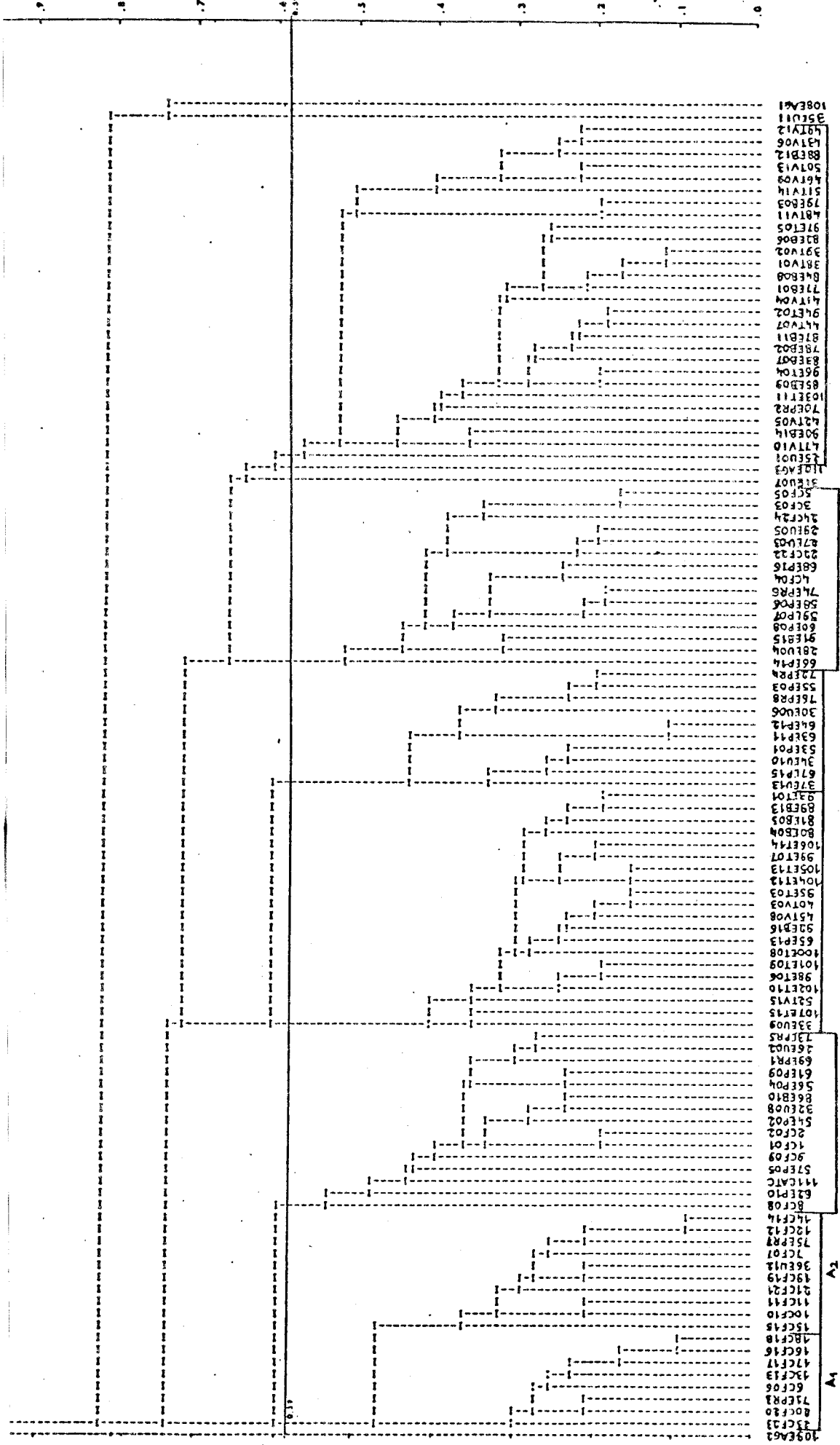


Fig 9b : Dendrogramme de l'agrégation selon la distance moyenne

- CF : fécale
- EU : eaux usées
- Origin
  - EPR et EP : eaux de surface propres et polluées
  - EB : eaux de distribution brutes
- ET : eaux de distribution traitées
- TV : terre vierge
- CATC : collection ATCC
- EAG : Enterobacter agglomerans



Tableau III

CLASSE	EFFECTIF	% GC MOYEN	ECART TYPE
A <sub>1</sub>	8	51,7	1,02
A <sub>2</sub>	10	51,6	0,44
A <sub>1</sub> + A <sub>2</sub>	18	51,65	0,73
B	15	50	1,25
C	20	50,9	0,98
D	10	54	1,22
E	15	52,6	0,62
F	28	49,5	1,71

VALEURS REPRESENTATIVES DE CHAQUE CLASSE



valeurs moyennes de chaque classe. On constate une différence plus ou moins prononcée entre les groupes, exception faite des deux sous-classes  $A_1$  et  $A_2$ .

L'écart type des moyennes de chaque classe rend compte de la dispersion globale des % GC autour de leur moyenne. Dans les classes  $A_2$  et E la valeur de l'écart type ( $\sigma$ ) est l'indice d'une grande homogénéité ; au contraire, le phénon F est très hétérogène, la valeur de  $\sigma$  étant la plus élevée.

L'ensemble des conclusions des tests statistiques sont récapitulées sur les tableaux IV et V. Les données concernant le % GC entre les classes  $A_1$  et  $A_2$ ,  $A_1$  et C, B et F, démontrent clairement qu'il n'existe pas de différence significative dans le contenu global en GC de ces populations. L'écart entre leurs moyennes est inférieur à une unité en % GC.

Par contre et pour les mêmes raisons les autres différences entre les classes sont significatives. La classe D est celle qui se distingue le plus de toutes les autres classes.

L'examen des résultats atteints avec un effectif plus restreint, nous conduit à considérer :

1. les souches "médianes" ou "centrotypes" définies à partir des coefficients de similitude phénétique ;
2. un échantillonnage de 5 souches comprenant la souche centrotipe, les deux souches les plus proches et les 2 souches les plus éloignées de la centrotipe ;
3. un échantillonnage de 10 souches rassemblant la souche centrotipe, les 4 souches les plus proches et les 5 souches les plus éloignées.

Le choix de celles-ci repose sur leurs distances à l'égard de la souche médiane de chaque classe ; sur les diagrammes (figure 9a) les



Tableau IV

A <sub>2</sub>	5,41*							
B	1,49	8,05*						
C	1,09	4,97*	1,62					
D	1,42	7,66*	1,05	1,54				
E	2,68	2,02	3,99*	2,46	3,79*			
F	2,81	15,19*	1,88	3,05*	1,98	7,52*		
A <sub>1</sub>								
		A <sub>2</sub>	B	C	D	E		

COMPARAISON DES VARIANCES

TEST DE F

\* Rejet de l'hypothèse  $\sigma_1 = \sigma_2$

Valeurs trouvées pour un risque d'erreur de 5 %



A <sub>2</sub>	0,36 (9)						
B	3,45 (21) <sup>**</sup>	4,67 (19) <sup>**</sup>					
C	2,00 (26)	2,66 (28) <sup>*</sup>	2,52 (33) <sup>*</sup>				
D	4,23 (16) <sup>**</sup>	5,92 (11) <sup>**</sup>	8,05 (23) <sup>**</sup>	7,56 (28) <sup>**</sup>			
E	2,40 (21) <sup>*</sup>	4,20 (23) <sup>*</sup>	7,22 (21) <sup>**</sup>	5,70 (33) <sup>**</sup>	3,5 (12) <sup>**</sup>		
F		4,096 (-) <sup>**</sup>	1,35 (-)	3,024 (-) <sup>**</sup>	58 (-) <sup>**</sup>	4,82 (-) <sup>**</sup>	
	A <sub>1</sub>	A <sub>2</sub>	B	C	D	E	F

DIFFERENCES SIGNIFICATIVES DU CONTENU GC MOYEN DE L'ADN DES SOUCHES DE CHAQUE CLASSE

Le nombre de degrés de liberté est donné entre parenthèses pour les valeurs de t calculées.

\* Paires significativement différentes pour une  $P_e = 0,05$

\*\* Paires significativement différentes pour une  $P_e = 0,01$

$P_e$  = Probabilité d'erreur



souches "centrotypes" se situent à proximité des valeurs centrales (% GC moyen  $\pm$  1 %), exception faite des classes A<sub>1</sub> et E. Si dans ces deux cas le % GC des "centrotypes" semble s'éloigner sensiblement de celui attribué à la classe, le calcul de leur différence n'excède pas l'unité (1 % et 0,8 % respectivement).

Les valeurs assignées à chaque classe (moyenne  $\pm$  1 écart type) sont représentées sur la figure 10, d'une part pour l'ensemble de la population, d'autre part pour des échantillonnages réduits. Sans se superposer totalement on peut dire qu'elles sont très voisines.

La figure 11 représente quelques exemples (souches centrotypes) de courbes de dénaturation représentées en fonction de l'accroissement d'absorbance relative (intervalles de 1° C) et de la température. La distribution du contenu moyen en GC s'est montrée dans tous les cas asymétrique. On remarque la nette individualité des souches 72 (classe D) et 44 (classe F). Les courbes dérivées de dénaturation des autres centrotypes connaissent un chevauchement plus ou moins important. Elles sont donc concernées par une même population statistique de molécules d'ADN. L'étendue des profils de fusion de l'ADN de la souche 72 (classe D) est significative d'une plus grande hétérogénéité de composition en bases intramoléculaires.

Un essai de corrélation entre les coefficients de similitude fournis par la taxonomie numérique et le % GC est esquissé sur la figure 12. Chaque classe est représentée par un triangle inversé dont la base est constituée par les variations des teneurs en GC (moyenne  $\pm$  1 $\sigma$ ) et le sommet, par la valeur de la distance phénotypique qui caractérise chaque population bactérienne.

La forme que prend le triangle témoigne de la dispersion à l'intérieur d'un groupe. Taxonomie numérique et étude du % GC se complètent pour dénoncer l'hétérogénéité particulière du groupe F : l'étendue des teneurs en GC est maximale et le coefficient de similitude est supérieur à celui des autres classes.

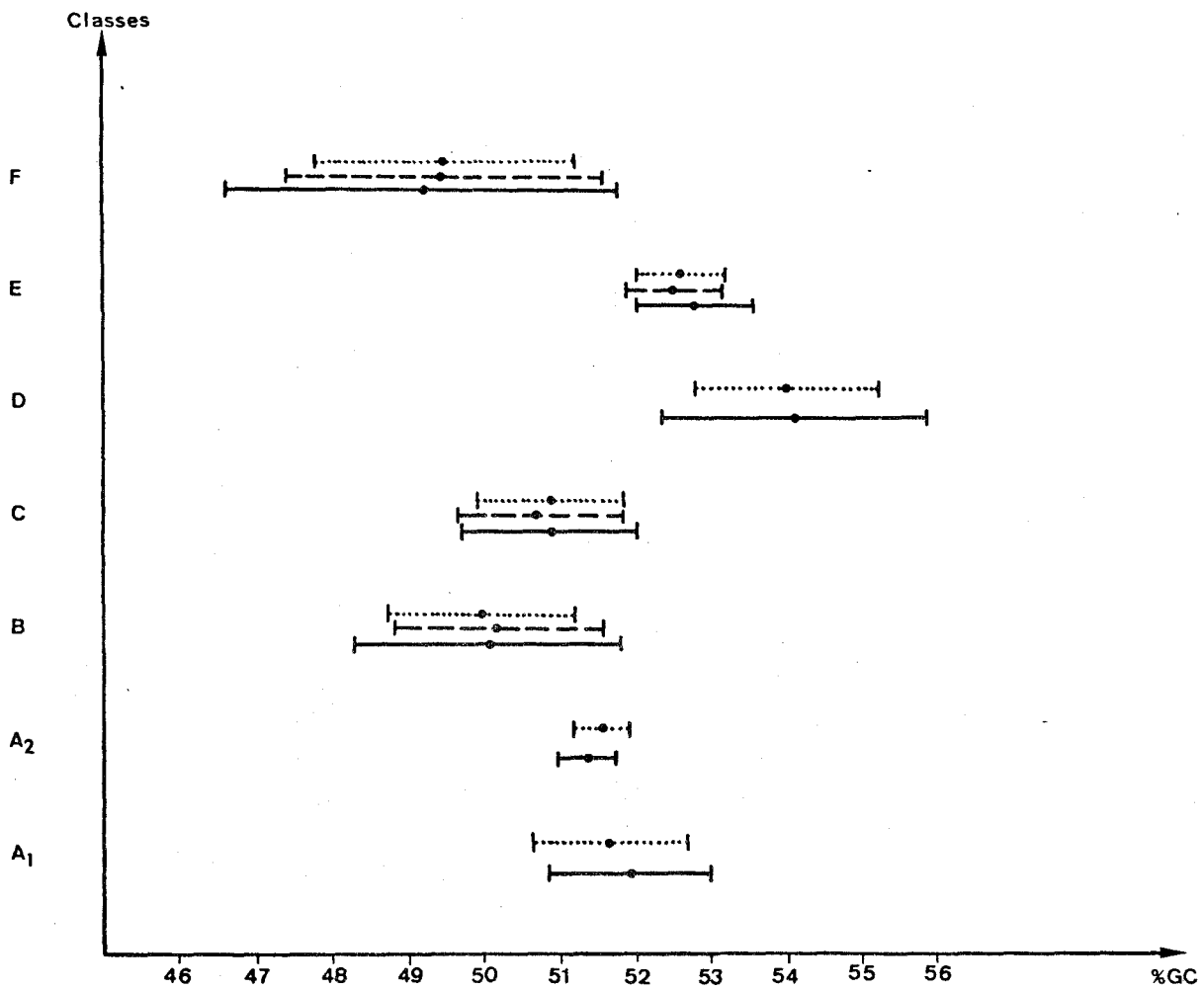


Fig 10 %GC de l'ADN des souches de chaque classe

..... ensemble de la population

----- échantillonnage réduit (10 souches)

———— échantillonnage réduit (5 souches)

( les valeurs représentent la moyenne  $\pm$  un écart type )



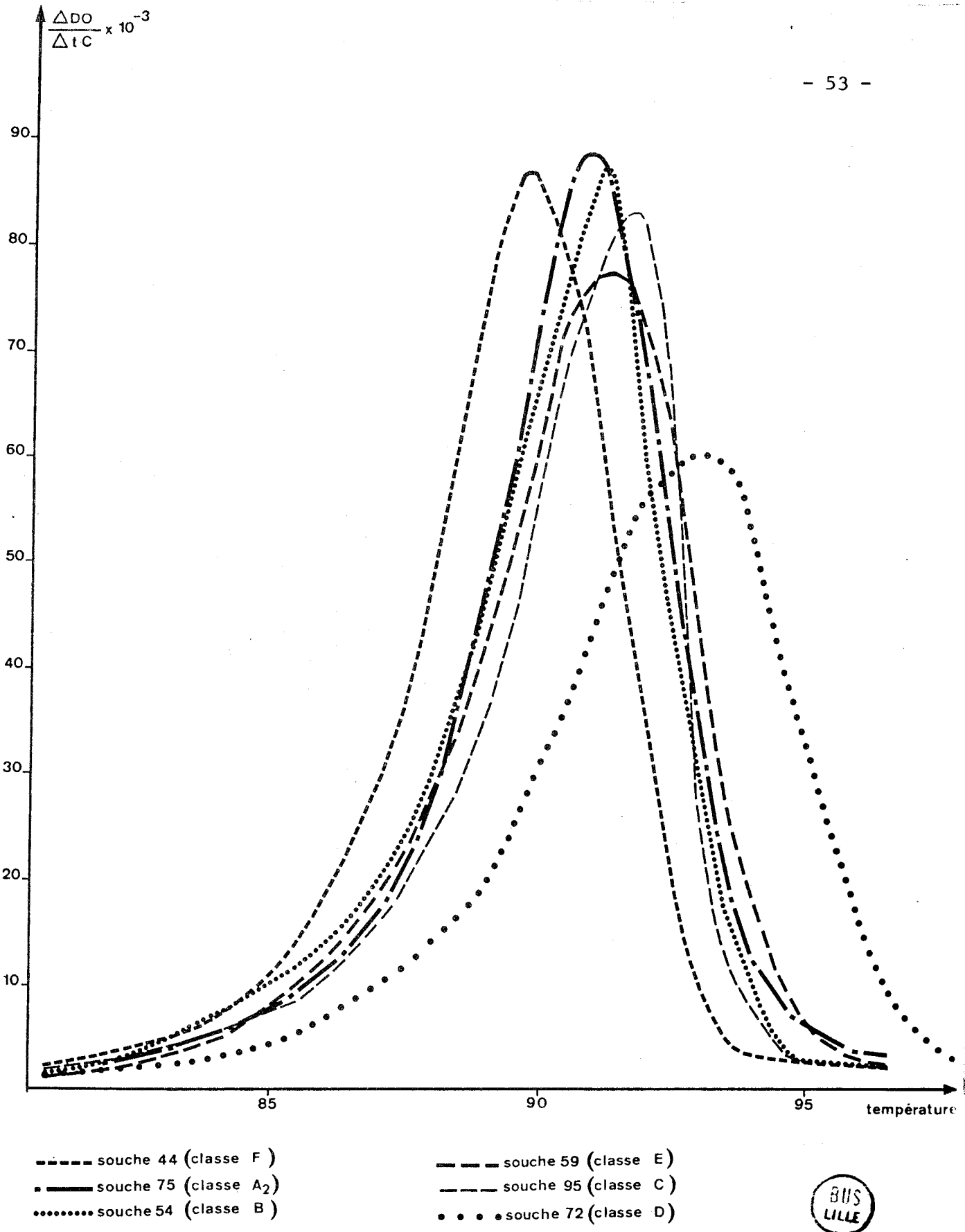


Fig 11 : Dérivées des courbes de dénaturation de l'ADN des différentes souches représentatives de chaque classe

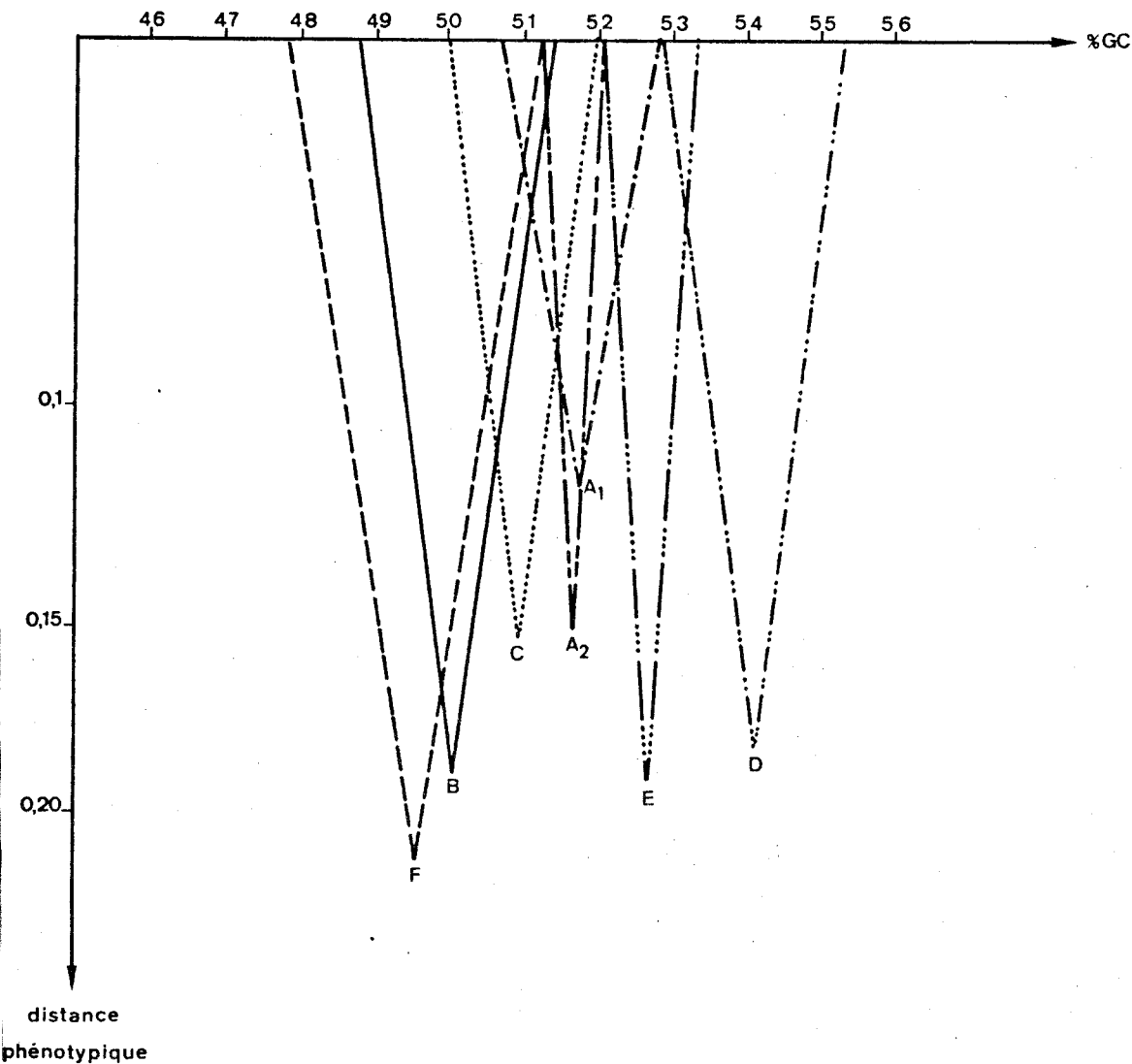


Fig 12 : Relations entre les différentes classes basées sur un paramètre phénotypique (distance) et un paramètre génomique (%GC)

- classe F
- classe A<sub>1</sub>
- classe B
- ..... classe E
- ..... classe C
- ..... classe D
- classe A<sub>2</sub>



### III - ANALYSE DES RESULTATS

#### A. - METHODES DE DETERMINATION DU $T_m$

Les méthodes étudiées, exception faite de la méthode graphique, nous conduisent à confirmer l'asymétrie des courbes de dénaturation thermique de l'ADN bactérien.

Ce phénomène avait déjà été constaté par DE LEY et coll. (32) pour les genres *Pseudomonas*, *Xanthomonas* et quelques espèces de *Chromobacterium*. Quoique de signification encore douteuse, plusieurs auteurs ont cherché à l'expliquer. MANDEL et MARMUR (38) supposent qu'il traduit une certaine polydispersion de la molécule dans sa longueur. DE LEY et coll. (32) pensent à une hétérogénéité intramoléculaire mais n'excluent pas la possibilité d'un artefact dû au procédé d'extraction. KNITTEL et coll. (67) attribuent cette asymétrie à la présence de plasmides. Les travaux récents de YAMAGISHI et coll. (138, 139, 140) ont permis de déceler la présence de régions riches en adénine et thymine dans l'ADN de plusieurs espèces bactériennes. On peut logiquement penser que l'utilisation d'intervalles de chauffage réduits, traduit, sur les courbes de dénaturation, dans les zones de température les plus basses, la présence de ces régions riches en adénine et thymine. Le choix d'une méthode de détermination du  $T_m$  doit tenir compte de ce phénomène qui semble être assez général avec l'ADN bactérien (25).

La méthode qui a recours au papier probit et celle des "droites de régression" apparaissent, sur la base des constatations précédentes, d'un usage limité. Les avantages énoncés par KNITTEL (67) lorsqu'il emploie le papier probit, perdent leur signification lorsque l'analyse de l'ADN bactérien met en évidence deux populations

gaussiennes. L'extrapolation des résultats à une population unique entraîne des imprécisions importantes. De même, l'étude de la fonction

$$\frac{DO_{t+\Delta t} - DO_t}{\Delta t} = f\left(\frac{DO_t + DO_{t+\Delta t}}{2}\right)$$

que nous avons assimilée à deux "droites de régression", ne conduira à une mesure rigoureuse du  $T_m$  que si la distribution des bases à l'intérieur de la molécule d'ADN, suit la loi de LAPLACE-GAUSS. Son application fera donc intervenir dans la plupart des cas un facteur de correction variable suivant l'ADN de la souche analysée. Ces deux techniques sont difficilement applicables, étant donné leur manque d'universalité.

Dans le cadre des analyses de routine, la détermination graphique du  $T_m$  apparaît comme une méthode recommandable. Elle est simple et rapide. Les résultats montrent son excellente reproductibilité. Cependant, l'étude de la fonction  $\frac{\Delta DO}{\Delta t} = f(t)$  suivie de

l'interpolation linéaire de la médiane, nous semble la méthode de choix. Elle n'est pas contingente de certaines évaluations subjectives dues à l'absence de parallélisme des extrémités de la courbe de dénaturation ou au tracé de la tangente. Elle nous apporte une meilleure connaissance du génôme bactérien. En plus du % GC moyen, elle fait apparaître, en effet, l'hétérogénéité de l'ADN analysé : les préparations d'ADN qui montreront un profil de fusion large seront l'indice d'une grande hétérogénéité dans la composition en bases intramoléculaires ; le coefficient de dispersion autour de la médiane, ainsi que la distance qui sépare le mode et la médiane, indiqueront, lorsqu'ils varieront pour des souches de même % GC moyen, l'existence de différences dans la séquence des bases de l'ADN ; la proportion de molécules à contenu GC commun dans deux échantillons d'ADN pourront être estimées à partir des aires de distribution communes (25) ; enfin, le rapport des surfaces délimitées par les courbes dérivées de dénaturation de deux souches différentes nous conduira à évaluer la taille relative d'un échantillon d'ADN.



La méthode est plus élaborée, donc plus complexe. Il est facile de remédier à cet inconvénient par l'usage d'une machine à calculer.

Le traitement statistique des données confirme ces conclusions. Il nous indique également la nécessité d'effectuer un nombre répété de mesures, parfois supérieur à trois, pour en augmenter la précision et les situer dans les valeurs limites de l'espèce étudiée, valeurs qui varient généralement de  $\pm 2,5 \%$  à  $\pm 3,5 \%$  autour du % GC moyen.

## B. - LES COEFFICIENTS DE CHARGAFF

### 1 - Corrélation entre taxonomie numérique et analyse du % GC

Les caractères analysés en taxonomie numérique sont l'expression phénotypique de l'information conservée dans l'ADN, alors que l'étude du % GC relève d'une connaissance directe du génôme. Les deux méthodes font appel à des options de classification différentes. Il semble donc opportun d'évaluer leur corrélation et de rechercher si elles se complètent ou se contredisent.

Notre discussion doit prendre en considération le travail de taxonomie numérique pour effectuer cette comparaison (50).

#### Classe A : *Levinea*

La classe A est composée de 18 souches dont les % GC sont compris entre 50 et 54. Elle comprend deux sous-classes  $A_1$  et  $A_2$  qui correspondent respectivement aux espèces *L. malonatica* et *L. amalonatica*. Ces deux sous-classes ont des teneurs en GC moyennes de 51,7 % et 51,6 % ; la différence entre ces 2 moyennes n'est absolument pas significative (tableaux III et V - p 47 et 50).

Il n'est pas étonnant de constater que ces deux sous-classes ont des valeurs des % GC identiques puisqu'elles ont en commun un grand nombre de propriétés phénotypiques ; signalons aussi qu'elles sont voisines par leur origine.

L'étude génomique met donc en évidence un groupe relativement homogène, confirmant les relations établies par les études phénétiques. Les valeurs obtenues dans nos expériences ( $51,65 \pm 0,73$ ) ne correspondent pas à celles de RICHARD et coll. (103) si l'on se réfère aux analyses de l'ADN des deux souches qu'ils ont analysées : le % GC est de 56 avec *L. malonatica* et de 58 avec *L. amalonatica*. Dans notre étude, cette divergence ne peut tenir compte ni des erreurs expérimentales ni des écarts types observés. Il apparaît donc à l'évidence que la détermination de la valeur GC moyenne d'une espèce ne puisse reposer sur des analyses isolées, mais sur un ensemble d'échantillons représentatifs d'une population. La proposition d'un statut de genre pour les *Levinea* confirmée par la valeur du % GC n'apparaît pas valable à la lumière de nos résultats personnels. Par contre, l'étude taxonomique qui a été faite, met en évidence son individualité en tant que classe. La moyenne des valeurs de cette classe ne présente pas de différence significative par rapport à celle de la classe C.

Classe B : *Citrobacter* H<sub>2</sub>S<sup>-</sup>

Le contenu moyen en guanine et cytosine de 73 % des souches réunies dans ce groupe se situe entre 49 % et 51 %. Quatre souches ont des pourcentages plus éloignés (47 % - 53 %). La teneur moyenne en bases peut être évaluée à 50 %.

Les études phénétiques identifient ces souches au genre *Citrobacter*. La souche ATCC 6750 incluse dans cette classe a été considérée comme représentative de l'espèce *Citrobacter intermedium*

jusqu'à ce que EWING et DAVIS (36) aient mis en évidence son aptitude à produire de l' $H_2S$  sur TSI agar. Il s'agit donc d'une souche typique de *C. freundii*. Les valeurs des % GC sont approximativement celles déterminées pour cette espèce (51 - 52,6 %) (4, 78, 121). On peut donc aussi bien par leurs caractères d'identification que par leur teneur en GC, rapprocher ces souches de l'espèce type de ce genre.

Les différences phénotypiques entre les classes A et B se trouvent affermies sur le plan génomique. Les résultats des coefficients de CHARGAFF entre celles-ci présentent une différence significative pour une probabilité d'erreur de 1 %. Compte tenu des études antérieures (50) ces deux classes sont identifiables à deux genres différents : les *Levinea* et les *Citrobacter*. Or, les résultats d'hybridation de CROSA et coll. conduisent à considérer *L. malonatica* comme une espèce séparée de *C. freundii* tandis que *L. amalonatica* pourrait être incluse dans le genre *Citrobacter*.

Dans l'état actuel des connaissances et en nous référant à notre double étude phénétique et génomique la création d'une classe correspondant aux *Levinea* apparaît valide. La définition d'un statut de genre ou d'espèce apparaît, pour l'instant, prématurée. Nos travaux ultérieurs en particulier à propos de *C. freundii* permettront sans doute de confirmer ces premières données car, rappelons-le, la classe B est par hypothèse assimilée à l'espèce *C. freundii* dont elle constituerait l'un des biotypes ( $H_2S^-$ ).

Ce critère apparaît également significatif dans les comparaisons des classes B - F, B - D et B - E.

#### Classe C : Entérobactéries inclassables (classe nouvelle)

Cette classe réunit 15 souches dont les teneurs en GC s'étendent de 50 % à 52 %, 3 souches dont le % GC est inférieur

à 50 (48 % - 50 %), 2 dont le % GC est supérieur à 52.

L'étude phénétique montre ici une classe particulièrement homogène et peu comparable aux autres classes, exceptée la D. Or, dans ce cas, les teneurs en % GC respectives les séparent sans ambiguïté.

Ce critère ne peut opérer une discrimination avec la classe A<sub>1</sub> ; les différences signalées par traitement statistique entre les classes A<sub>2</sub> et C, B et C sont significatives à un risque d'erreur de 5 %. Si l'on tient compte de l'individualité phénotypique des souches de cette classe, exprimée aussi par leur biotope, il est légitime de regarder toute similitude de composition en bases avec d'autres classes comme une coïncidence.

Le caractère lysine + de ces souches a été rapporté antérieurement par LECLERC et BUTTIAUX (69). Analysées par CROSA et coll. (19) au cours d'études d'hybridation, elles ont montré une similitude maximale de 47 % avec le genre *Serratia*. Pour poursuivre ce rapprochement on peut noter que l'espèce *Serratia liquefaciens* (51 - 52 %) a un % GC proche de celui de la classe C. Il apparaît pourtant que cette parenté soit illusoire compte tenu de l'homogénéité très nettement marquée de cette classe C sur le plan phénotypique et de son individualité par rapport au genre *Serratia* dans son ensemble.

#### Classe D : Entérobactéries inclassables (classe nouvelle)

10 souches viennent se grouper dans cette classe ; elle est caractérisée par un % GC moyen de 54. L'étude de ce coefficient objective des différences importantes entre cette classe et les classes restantes : parmi les six classes étudiées, celle-ci est la seule à présenter une différence hautement significative ( $P_e = 0,01$ ) avec toutes les autres classes.

Bien que différenciée de la classe C par l'étude de taxonomie numérique, cette distinction n'apparaissait pas particulièrement tranchée sur la base de caractères hautement significatifs. Elle se trouve ici confirmée par les valeurs des GC : l'écart moyen entre le % GC des deux classes atteint trois unités.

Si les bactéries de ces deux groupes ont des génômes dont la taille correspond approximativement à  $10^7$  paires de nucléotides, cette différence de 3 % représente approximativement  $2,9 \times 10^4$  GC substitutions de la paire AT dans l'ADN des bactéries de cette classe (75). Ces substitutions survenant au hasard, il est logique de prévoir que les souches de ces 2 classes auront un nombre faible de séquences nucléotidiques communes. Les expériences d'hybridation qui sont en cours permettront de statuer sur cette hypothèse. Déjà les courbes de distribution du contenu GC entre les souches "centrotypes" de ces deux classes ne laisse présumer qu'un faible pourcentage d'homologie (figure 11 - p 53).

Le rapprochement éventuel des souches de cette classe avec d'autres genres de la famille des *Enterobacteriaceae* sur la base du % GC (*Serratia* 56 % - 58 % - *Klebsiella* 54 % - 59 %) (14, 68) n'est pas satisfaisant à la lumière des informations phénétiques.

Classe E : Entérobactéries acétoïne - proches des *Enterobacter cloacae*

Cette classe est composée de 15 souches dont le % GC moyen est caractérisé par une valeur de 52,6. Elle représente une entité sur le plan phénétique qui s'exprime aussi génétiquement. En effet, les souches rassemblées dans cette classe forment un groupe très homogène dont la valeur moyenne du contenu GC est peu dispersée ( $52,6 \pm 0,62$ ).

Les travaux de taxonomie numérique ont conduit à comparer cette classe à l'espèce *Enterobacter cloacae*.

Seule l'analyse de l'ADN de deux souches de cette espèce a été rapportée dans la littérature ; elle a révélé des teneurs en GC de 54 et 57 % ce qui est en totale contradiction avec nos résultats propres. Cette contradiction pourrait conduire à différencier la classe E de l'espèce *Enterobacter cloacae* malgré les similitudes phénétiques observées. Cette hypothèse ne peut être retenue car la valeur du % GC de l'une de ces souches de collection (ATCC 13047) est de 54,6 % - 55,4 % dans la littérature (97) et de 51 % dans notre étude (2 extractions consécutives). Ces divergences tendent à toute interprétation aléatoire. Il apparaît sage d'attendre les informations complémentaires que nous fourniront les études phénétiques et génômiques en cours.

#### Classe F : Entérobactéries inclassables

Les souches rassemblées dans ce groupe sont au nombre de 28 dont 71 % ont une teneur en GC inférieure à 50 %. Pour les souches restantes la moyenne des % GC est de 51-52.

L'hétérogénéité des valeurs des % GC dans ce groupe est frappante ; 7 souches sont nettement séparées des autres ; elles ont un % GC de 51-53,5. La souche n° 79 dont le contenu GC est extrême (53,5 %) correspond à une des souches de cette classe à distance de similitude phénétique très élevée.

On pourrait évoquer l'existence de deux populations dont les % GC moyens se situeraient à 49 % et 51,5 % respectivement. Une distribution bimodale identique a déjà été observée par COLWELL et MANDEL (14) pour l'espèce *Serratia marcescens*. Ceci n'avait pas retenu l'attention des auteurs. Il n'est pas impossible d'imaginer dans cette classe la présence de deux sous-groupes bactériens différents ayant un grand nombre de caractères communs.

L'analyse étendue de ces souches par hybridation doit permettre de mieux interpréter ces données. Si celle-ci permet d'envisager ultérieurement deux sous-classes, un examen phénotypique plus approfondi deviendra indispensable.

L'apparement des souches de cette classe au genre *Citrobacter* ne paraît pas indiqué par référence aux caractères phénétiques ; il apparaît au contraire valide sur la base des % GC.

Le rapprochement avec le genre *Levinea* n'est pas souhaitable en analyse numérique ; elle n'est pas davantage confirmée par l'analyse des % GC : la variation de 2 % en GC entre cette classe et les classes A<sub>1</sub> et A<sub>2</sub> est significative à 95 %.

La différence observée entre la moyenne en GC de cette classe et celle des classes C, D et E est statistiquement très significative.

Les études phénétiques ont permis d'envisager une parenté avec l'espèce *E. agglomerans*. Si la diversité des caractères de cette espèce rend difficile toute comparaison objective, le % GC exclut totalement ce rapprochement (55 % à 56 % pour *Erwinia herbicola*) (94).

## 2 - Evaluation et conséquences générales

Les informations apportées par l'analyse du % GC apparaissent étroitement complémentaires des données fournies par la taxonomie numérique. Pourtant, des réserves peuvent être faites en ce qui concerne la classe F. Le contenu en bases de l'ADN se montre assez homogène à l'intérieur de chaque classe et indique que les groupements fondés sur le degré de similitude phénétique traduisent une relation génétique entre bactéries.

Cette approche parallèle a l'avantage de permettre une meilleure description des unités taxonomiques de base avec un niveau d'information plus élevé qualitativement et quantitativement. Les deux méthodes se complètent parfaitement mais isolément elles sont assujetties à des limites de confiance moindres. La figure 12 nous en apporte l'illustration par l'intrication de certaines classes (p. 54).

La teneur en bases de l'ADN d'un certain nombre de souches doit être précisée car il s'agit d'un critère fondamental préliminaire à toute étude d'homologie de séquences. De plus, les informations d'ordre génétique pourront être confrontées avec profit à partir de cette donnée. Le besoin d'user de ce critère pour l'élaboration d'une taxonomie moléculaire apparaît dès lors impératif.

Il serait donc rationnel de concevoir les étapes qui conduisent à la construction de la nouvelle classification des bactéries dans l'ordre suivant :

taxonomie numérique  $\longrightarrow$  % GC  $\longrightarrow$  hybridation

On peut à juste titre rechercher le nombre le plus raisonnable de souches à analyser pour que chaque unité taxonomique soit effectivement représentée.

Eriger une relation directe et absolue entre les valeurs GC d'une souche considérée comme représentative (centrotype) et celles de la classe paraît difficile. Les erreurs pourraient être excessives. Par contre, on peut admettre qu'un échantillonnage minimum donnera une meilleure image de la population. C'est pour répondre à ce type d'interrogation que nous avons opéré un choix sur 5 ou 10 souches tirées à l'intérieur d'un groupe par les critères d'indice de similitude proche ou éloigné du centrotype ;



ce choix peut être considéré fait au hasard quant aux valeurs du % GC. En effet, des distances phénétiques faibles entre deux souches n'impliquent pas obligatoirement une teneur en GC identique et vice-versa. Les résultats sont ainsi d'interprétation statistique valable.

L'examen de la figure 10 où l'on passe de l'analyse de toutes les souches à l'analyse d'un effectif plus réduit, fait apparaître des divergences faibles. Le % GC moyen d'un échantillonnage de 10 souches s'éloigne au maximum de 0,2 % par rapport à la moyenne de la population totale. La déviation observée pour un nombre de 5 souches ne dépasse pas 0,3 %. Seuls les écarts types considérés diffèrent notablement dans l'un et l'autre cas. Ceci est totalement concevable et répond aux règles statistiques des échantillons issus d'une même population (p. 52).

Il ne semble donc pas essentiel d'étudier tous les individus appartenant à une classe pour que le % GC serve d'appui à la classification.

On peut raisonnablement avancer que l'homogénéité observée dans les valeurs du % GC ne dispense pas de sélectionner dans le lot des souches à teneur extrême en bases nucléiques. Lorsque des classes nouvelles sont créées, l'examen d'un nombre plus élevé de souches apparaît souhaitable, afin de reporter leur % GC moyen de façon plus rigoureuse.

Ces études détaillées nous conduisent à considérer par ailleurs la dispersion en % GC de chaque classe (tableau III - p. 47) Celle-ci est variable mais ne correspond jamais à une grande hétérogénéité (sauf classe F).

Quelle est la variation de composition en bases intrinsèque à une classe que l'on peut admettre pour définir une véritable entité taxonomique ?

Si l'on examine les valeurs rapportées dans la littérature, la plupart des teneurs des % GC attribuées aux souches appartenant à un même genre proviennent de l'expérience de différents chercheurs. De plus, ceux-ci s'adressent souvent à des cultures référencées qui ne sont pas toujours représentatives d'un phénon ou d'une classe. Le faible nombre des souches analysées restreint la valeur et la portée de l'information. Ainsi, les limites d'hétérogénéité acceptées pour un même genre varient dans une moyenne de 5 % à 8 %.

Plus rares sont les études où les auteurs prêtent attention à la fois au nombre de souches et à l'espèce. Citons en exemple certains travaux de DE LEY et coll. (24, 27), MANDEL (75), HOGAN et COLWELL (56). Le premier examinant un grand nombre de souches recueille des écarts extrêmes qui s'élèvent à 2,6 % en GC avec *Agrobacterium tumefaciens* et à 5,5 % avec *Moraxella lwoffii*. Les observations du second sur *Pseudomonas fluorescens* rapportent des écarts de l'ordre de 4 %. Les derniers citent pour *Pseudomonas bathycetes* une variation de 5 %. D'autres exemples pourraient confirmer cette diversité des résultats et les fluctuations admises pour une même espèce.

Notre travail se distingue surtout des précédents par le large éventail d'échantillonnage. Il est probable que des études identiques feraient apparaître des dispersions équivalentes pour les classes délimitées en taxonomie numérique, ce qui permettrait d'assurer d'une façon générale la compatibilité de telles variations (5 %) avec le niveau taxonomique le plus élémentaire. A présent, la décision d'une division, lorsqu'on utilise la composition en bases comme une partie de la définition d'un groupe, doit plutôt être fondée sur l'uniformité fournie à partir des souches analysées.

## CONCLUSIONS

---

Au cours de ces travaux, l'analyse de l'ADN de 111 souches, apparentées au genre *Citrobacter*, nous a permis d'aboutir à un certain nombre de conclusions originales.

I - Sur le plan taxonomique cette étude fait apparaître soit des divergences très significatives des % GC entre les différentes classes, divergences qui confirment l'analyse numérique, soit des résultats en contradiction apparente avec cette dernière:

- L'individualité de la classe D qui comprend des souches lysine +, mise en lumière par l'analyse numérique est parfaitement confirmée par l'étude du % GC.

- De même la classe C, qui comprend également des souches lysine +, apparaît indépendante de la classe précédente par les caractères phénétiques et par les teneurs en % GC.

- La classe B, assimilée à l'espèce *Citrobacter freundii* par l'étude phénétique, se voit confirmée sur la base du contenu GC de l'ADN de ses souches.

- La création d'une classe correspondant aux *Levinea* apparaît valide dans la double approche phénétique et génomique.

- Les souches rassemblées dans la classe E ont un % GC qui les individualise des autres classes, mais qui exige d'autres approches taxonomiques avant d'en préciser la situation.

- Les classes A et C, B et F ne peuvent être différenciées à la lumière de ce seul critère. Ce fait met en relief le besoin de l'utiliser en corrélation avec les données phénétiques.

- La classe F est caractérisée par une grande hétérogénéité du % GC des souches constitutives, hétérogénéité non mise en évidence par la taxonomie numérique. L'analyse de ces souches par hybridation s'impose.

II - Sur le plan général il est possible de faire quelques remarques :

- La sensibilité de la détermination du  $T_m$  conduit à opérer un nombre répété de mesures ; un nombre minimum de 3 répliques paraît indispensable pour aboutir à des résultats cohérents et statistiquement significatifs.

- Les valeurs isolées des % GC retrouvées dans la littérature comme représentatives d'une espèce ne paraissent pas significatives. Il importe en effet de les définir à partir d'une population globale déjà analysée en taxonomie numérique.

- L'élaboration d'une classification rationnelle tenant compte des différentes méthodes d'approche doit suivre un cheminement ordonné qui pourrait être schématisé par les étapes suivantes :

1. analyse numérique
2. mesure du % GC
3. hybridation

- Il ne semble pas indispensable d'étudier toutes les souches appartenant à une classe pour que le % GC soit représentatif de cette classe. Un nombre de 5 souches apparaît généralement suffisant.

- Lorsque l'analyse ainsi programmée fait apparaître une hétérogénéité inattendue à l'intérieur d'une classe, l'examen d'un plus grand nombre de souches apparaît indispensable.

- La définition d'une véritable entité taxonomique lorsqu'on utilise la composition en bases devrait être fondée sur l'uniformité fournie à l'intérieur d'une classe à la fois par les caractères phénétiques et génômiques.

BIBLIOGRAPHIE

1. AFNOR (Norme Française homologuée)
  - a) NF X 06-062 : "Estimation du rapport des variances (des écarts types) de deux lots"  
Mars 1971
  - b) NF X 06-065 :
    - . "Comparaison de deux moyennes"(écarts types inconnus mais égaux)
    - . "Comparaison de deux moyennes. Test d'ASPIN WELCH"  
(écarts types inconnus et inégaux)
    - . "Comparaison de deux distributions. Test des rangs de COLIN WHITE"  
Décembre 1971
2. ANDO Y., FUKADA E. : "Effect of drying on deoxyribonucleic acid"  
J. Rad. Res., 1974, 15, n° 4, 212-218
3. BAILIE W.E., COLES E.H., WEIDE K.D. : "Deoxyribonucleic acid characterization of a microorganism isolated from infectious thromboembolic meningo-encephalomyelitis of Cattle"  
Int. J. Syst. Bact., 1973, 23, 231-237
4. BAPTIST J.N., SHAW C.R., MANDEL M. : "Zone electrophoresis of enzymes in bacterial taxonomy"  
J. Bact., 1969, 99, 180-188
5. BELLAMY A.R., RALPH R.K. : "Recovery and purification of nucleic acids by means of cethyl trimethylammonium bromide"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE), 12 part. B (p. 156-160), Acad. Press, New York & London, 1968
6. BELOZERSKY A.N., SPIRIN A.S. :  
In "The Nucleic Acids" (E. CHARGAFF & J.N. DAVIDSON)3, (p.147)  
Acad. Press, New York, 1960  
Op. cité par MARMUR J. et DOTY P. (voir réf. 82)

7. CHAMPETIER G., MONNERIE L. : "L'acide désoxyribonucléique"  
In "Introduction à la Chimie macromoléculaire", (p. 706),  
Ed. Masson et Cie, Paris, 1969
8. CHARGAFF E., CRAMPTON Ch., LIPSHITZ R. : "Separation of calf  
thymus deoxyribonucleic acid into fractions of different  
composition"  
Nature, 1953, 172, 289-292
9. CHARLIER G., DEKEYSER P., FLORENT A., STROBBE R., DE LEY J. :  
"DNA base composition and biochemical characters of  
*Campylobacter* strains"  
Antonie V. Leeuwenhoek, 1974, 40, 145-151
10. CITARELLA R.V., COLWELL R.R. : "Polyphasic taxonomy of the germs  
*Vibrio* : polynucleotide sequence relationships among selec-  
ted *Vibrio* species"  
J. Bact., 1970, 104, 434-442
11. COLWELL R.R. : "Genetic and phenetic classification of bacteria"  
In "Advances in applied Microbiology" (D. PERLMAN), 16,  
(p. 137-175), Acad. Press, New York & London, 1973
12. COLWELL R.R., CITARELLA R.V., RYMAN I. : "Deoxyribonucleic acid  
base composition and adansonian analysis of heterotrophic,  
aerobic *Pseudomonas*"  
J. Bact., 1965, 90, 1148-1149
13. COLWELL R.R., MANDEL M. : "Adansonian analysis and deoxyribonucleic  
acid base composition of some gram negative bacteria"  
J. Bact., 1964, 87, 1412-1422
14. COLWELL R.R., MANDEL M. : "Adansonian analysis and deoxyribonucleic  
acid base composition of *Serratia marcescens*"  
J. Bact., 1965, 89, 454-461



15. COX E.C., YANOFSKY C. : "Altered base ratios in the DNA of an *Escherichia coli* mutator strain"  
Proc. Nat. Acad. Sci., 1967, 58, 1895-1902
16. COYKENDALL A.L., SPECHT P.A. : "DNA base sequence homologies among strains of *Streptococcus sanguis*"  
J. Gen. Microbiol., 1975, 91, 91-98
17. CRAVERI R., HILL L.R., MANACHINI P.L., SILVESTRI L.G. : "Deoxyribonucleic acid base compositions among thermophilic actinomycetes : the occurrence of two strains with low GC content"  
J. Gen. Microbiol., 1965, 41, 335-339
18. CROMBACH W.H.J. : "Deep-freezing of bacterial DNA for thermal denaturation and hybridization experiments"  
Antonie V. Leeuwenhoek, 1973, 39, 249-255
19. CROSA J.H., STEIGERWALT A.G., FANNING G.R., BRENNER D.J. : "Polynucleotide sequence divergence in the genus *Citrobacter*"  
J. Gen. Microbiol., 1974, 83, 271-282
20. DARBY G.K., JONES A.S., KENNEDY J.F., WALKER R.T. : "Isolation and analysis of the nucleic acids and polysaccharides from *Clostridium welchii*"  
J. Bact., 1970, 103, 159-165
21. DE LEY J. : "*Pseudomonas* and related genera"  
Ann. Rev. Microbiol., 1964, 18, 17-46
22. DE LEY J. : "Effect of mutation on DNA-composition of some bacteria"  
Antonie V. Leeuwenhoek, 1964, 30, 282-288
23. DE LEY J. : "The quick approximation of DNA base composition from absorbancy ratios"  
Antonie V. Leeuwenhoek, 1967, 33, 203-208

24. DE LEY J. : "DNA base composition and taxonomy of some *Acinetobacter* strains"  
Antonie V. Leeuwenhoek, 1968, 34, 109-114
25. DE LEY J. : "Compositional nucleotide distribution and the theoretical prediction of homology in bacterial DNA"  
J. Theoret. Biol., 1969, 22, 89-116
26. DE LEY J. : "Reexamination of the association between melting point, buoyant density and chemical base composition of deoxyribonucleic acid"  
J. Bact., 1970, 101, 738-754
27. DE LEY J., BERNAERTS M., RASSEL A., GUILMOT J. : "Approach to an improved taxonomy of the genus *Agrobacterium*"  
J. Gen. Microbiol., 1966, 43, 7-17
28. DE LEY J., CATTOIR H., REYNAERTS A. : "The quantitative measurement of DNA hybridization from renaturation rates"  
Europ. J. Biochem., 1970, 12, 133-142
29. DE LEY J., DE SMEDT J. : "Improvements of the membrane filter method for DNA : rRNA hybridization"  
Antonie V. Leeuwenhoek, 1975, 41, 287-307
30. DE LEY J., SCHELL J. : "Deoxyribonucleic acid base composition of acetic acid bacteria"  
J. Gen. Microbiol., 1963, 33, 243-253
31. DE LEY J., TITGAT R. : "Evaluation of membrane filter methods for DNA-DNA hybridization"  
Antonie V. Leeuwenhoek, 1970, 36, 461-474
32. DE LEY J., VAN MUYLEM J. : "Some application of deoxyribonucleic acid base composition in bacterial taxonomy"  
Antonie V. Leeuwenhoek, 1963, 29, 344-358

33. DE VOE H., TINOCO I. jr. : "The stability of helical polynucleotides : base contributions"  
J. Mol. Biol., 1962, 4, 500-517
34. DOVE W.F., DAVIDSON N. : "Cation effects on the denaturation of DNA"  
J. Mol. Biol., 1962, 5, 467-478
35. EICHHORN G.L. : "Metal ions as stabilizers or destabilizers of the deoxyribonucleic acid structure"  
Nature, 1962, 194, 474-475
36. EWING W.H., DAVIS B.R. : "Biochemical characterization of *Citrobacter diversus* (Burkey) WERKMAN and GILLEN and designation of the neotype strain"  
Int. J. Syst. Bact., 1972, 22, 12-18
37. FALKOW S. : "Nucleic acids, genetic exchange and bacterial speciation"  
Am. J. Med., 1965, 39, 753-765
38. FRANKLIN R.E., GOSLING R.G. : "Molecular configuration in sodium thymonucleate"  
Nature, 1953, 171, 740-741
39. FREDERICQ E., OTH A., FONTAINE F. : "The ultraviolet spectrum of deoxyribonucleic acids and their constituents"  
J. Mol. Biol., 1961, 3, 11-17
40. FREIFELDER D. : "The use of NaClO<sub>4</sub> to isolate bacteriophage nucleic acids"  
In "Methods in Enzymologie" (L. GROSSMAN & K. MOLDAVE)  
12, part. A, (p. 550-554)  
Acad. Press, New York & London, 1967

41. FRENCH C.S., MILNER H.W. : "Disintegration of bacteria and small particles by high-pressure extrusion"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE),  
1 (p. 64-67), Acad. Press, New York & London, 1955
42. FRONTALI C., HILL L.R., SILVESTRI L.G. : "The base composition of deoxyribonucleic acids of *Streptomyces*"  
J. Gen. Microbiol., 1965, 38, 243-250
43. GARDNER J.M., KADO C.J. : "Comparative base sequence homologies of the deoxyribonucleic acids of *Erwinia* species and other *Enterobacteriaceae*"  
Int. J. Syst. Bact., 1972, 22, 201-209
44. GASSER F. : "Introduction à la taxonomie des bactéries. Taxonomie et phylogénie"  
In "Microbiologie générale : Bactéries, Bactériophages" 1  
(p. 15-23) Ediscience, Paris, 1975
45. GASSER F., MANDEL M. : "Deoxyribonucleic acid base composition of the genus *Lactobacillus*"  
J. Bact., 1968, 96, 580-588
46. GASSER F., SEBALD M. : "Composition en bases nucléiques des bactéries du genre *Lactobacillus*"  
Ann. Inst. Pasteur, 1966, 110, 261-275
47. GAUSE G.F. : "Alterations of DNA base composition in bacteria"  
In "Progress in Nucleic Acid Research and Molecular Biology" (J.N. DAVIDSON, W.E. COHN), 8 (p. 49-71)  
Acad. Press, New York & London, 1968
- 48 GAUSE G.F., DUDNIK Y.V., LAIKO A.V., NETYKSA E.M. : "Induction of mutants with altered DNA composition : effect of ultra-violet on *Bacterium paracoli* 5099"  
Science, 1967, 157, 1196-1197

49. GAVINI F. : "Contribution à l'étude taxinomique des groupes  
*Cytophaga - Flavobacterium - Xanthomonas*"  
Thèse d'Université en Pharmacie, Lille II, 1972-1973
50. GAVINI F., LEFEBVRE B., LECLERC H. : "Positions taxonomiques  
d'Entérobactéries  $H_2S^-$  par rapport au genre *Citrobacter*"  
Ann. Microbiol., 1976, 127 A, 275-295
51. GEIDUSCHEK E.P. : "On the factors controlling the reversibility  
of DNA denaturation"  
J. Mol. Biol., 1962, 4, 467-487
52. GOUDOT A. : "Les acides nucléiques"  
In "La Chimie quantique appliquée à l'étude du DNA et du  
RNA", (p. 13-17), Ed. Gauthier-Villars, Paris, 1966
53. GROSSMAN L., LEVINE S.S., ALLISON W.S. : "The reaction of formal-  
dehyde with nucleotides and  $T_2$  bacteriophage DNA"  
J. Mol. Biol., 1961, 3, 47-60
54. HILL L.R. : "An index to deoxyribonucleic acid base compositions  
of bacterial species"  
J. Gen. Microbiol., 1966, 44, 419-437
55. HILL E.B., WAYNE L.G., GROSS W.M. : "Purification of mycobacterial  
deoxyribonucleic acid"  
J. Bact., 1972, 112, 1033-1039
56. HOGAN M.A., COLWELL R.R. : "DNA base composition and esterase  
patterus of bacteria isolated from deep sea sediments"  
J. Appl. Bact., 1969, 32, 103-111
57. HOTCHKISS R.D. : "Isolation of sodium deoxyribonucleate in bio-  
logically active form from bacteria"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE), 3  
(p. 692-696), Acad. Press, New York & London, 1957

58. INGRAM V.M., GRUNBERG-MANAGO M. : "ADN"  
In "Biosynthèse des macromolécules", (p. 20-28)  
Ediscience, Paris, 1970
59. INMAN R.B., BALDWIN R.L. : "Helix-random coil transitions in  
synthetic DNAs of alternating sequence"  
J. Mol. Biol., 1962, 5, 172-184
60. JONES D., SNEATH P.H.A. : "Genetic transfer and bacterial  
taxonomy"  
Bact. Rev., 1970, 34, 40-81
61. KATZ S. : "Mechanism of the reaction of polynucleotides and Hg<sup>2+</sup>"  
Nature, 1962, 194, 569
62. KIRBY K.S. : "A new method for the isolation of deoxyribonucleic  
acids : evidence on the nature of bonds between deoxy-  
ribonucleic acid and protein"  
Biochem. J., 1957, 66, 495-504
63. KIRBY K.S. : "The preparation of deoxyribonucleic acids by the  
p-aminosalicylate-phenol method"  
Biochem. Biophys. Acta, 1959, 36, 117-124
64. KIRBY K.S. : "Separation of deoxyribonucleic acid from ribonucleic  
acid without the use of enzymes"  
Biochem. Biophys. Acta, 1961, 47, 18-26
65. KIRBY K.S. : "Isolation of nucleic acids with phenolic solvents"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE), 12  
part. B (p. 87-99), Acad. Press, New York & London, 1968
66. KIT S. : "Deoxyribonucleic acids"  
Ann. Rev. Biochem., 1963, 32, 43-82

67. KNITTEL M.D., BLACK C.H., SANDINE W.E., FRASER D.K. : "Use of normal probability paper in determining thermal melting values of deoxyribonucleic acid"  
Canad. J. Microbiol., 1968, 14, 239-245
68. LASKIN A.I., LECHEVALIER H.A. : "Microbial composition"  
In "Handbook of Microbiology" 2, (p. 585-691), CRC Press, U.S.A., 1973
69. LECLERC H., BUTTIAUX R. : "Les *Citrobacter*"  
Ann. Inst. Pasteur, Lille, 1965, 16, 67-74
70. LEE K.Y., WAHL R., BARBU E. : "Contenu en bases puriques et pyrimidiques des acides désoxyribonucléiques des bactéries"  
Ann. Inst. Pasteur, 1956, 91, 212-224
71. LEIFSON E., MANDEL M. : "The DNA base composition of a flagellar mutant of *Comamonas terrigena* ATCC 8461"  
Antonie V. Leeuwenhoek, 1966, 32, 57-59
72. MAHLER H.R., CORDES E.H. : "Nucleic acids"  
In "Biological Chemistry", 186-266, HARPER & ROW, New York, 1971
73. MAHLER H.R., MEHROTA B.D. : "Dependence of deoxyribonucleic amine interactions on deoxyribonucleic acid composition"  
Biochem. Biophys. Acta, 1962, 55, 252-254
74. MANDEL M. : "The interaction of spermine and native deoxyribonucleic acid"  
J. Mol. Biol., 1962, 5, 435-441
75. MANDEL M. : "Deoxyribonucleic acid base composition in the genus *Pseudomonas*"  
J. Gen. Microbiol., 1966, 43, 273-292

76. MANDEL M. : "New approaches to bacterial taxonomy : perspective and prospects"  
Ann. Rev. Microbiol., 1969, 23, 239-271
77. MANDEL M., MARMUR J. : "Use of ultraviolet absorbance temperature profile for determining the guanine + cytosine content of DNA"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE), 12, part. B, (p. 195-206), Acad. Press, New York & London, 1968
78. MANDEL M., ROWND R. : "DNA base composition in the *Enterobacteriaceae* : an evolutionary sequence"  
In "Taxonomic Biochemistry and Serology" (C.A. LEONE), 585, Ronald Press, New York, 1964
79. MANDEL M., WEEKS O.B., COLWELL R.R. : "Deoxyribonucleic acid base composition of *Pseudomonas piscicida*"  
J. Bact., 1965, 90, 1492-1493
80. MARMUR J. : "A procedure for the isolation of deoxyribonucleic acid from micro-organisms"  
J. Mol. Biol., 1961, 3, 208-218
81. MARMUR J., DOTY P. : "Heterogeneity in deoxyribonucleic acids. 1. Dependence on composition of the configurational stability of deoxyribonucleic acids"  
Nature, 1959, 183, 1427-1429
82. MARMUR J., DOTY P. : "Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature"  
J. Mol. Biol., 1962, 5, 109-118
83. MARMUR J., FALKOW S., MANDEL M. : "New approaches to bacterial taxonomy"  
Ann. Rev. Microbiol., 1963, 17, 329-372



84. MARMUR J., ROWND R., SCHILDKRAUT C.L. : "Denaturation and renaturation of deoxyribonucleic acid"  
In "Progress in Nucleic Acid Research" (J.N. DAVIDSON, W.E. COHN), 1, (p. 231-300), Acad. Press, New York & London, 1963
85. MARMUR J., TS' O P.O.P. : "Denaturation of deoxyribonucleic acid by formamide"  
Biochem. Biophys. Acta, 1961, 51, 32-36
86. MASSIE H.R., ZIMM B.H. : "The use of hot phenol in preparing DNA"  
Proc. Nat. Acad. Sci., 1965, 54, 1641-1643
87. MESELSON M., STAHL F.W., VINOGRAD J. : "Equilibrium sedimentation of macromolecules in density gradients"  
Proc. Nat. Acad. Sci., 1957, 43, 581-588
88. MEYER S.A., SCHLEIFER K.H. : "Rapid procedure for the approximate determination of the deoxyribonucleic acid base composition of *micrococci*, *Staphylococci* and other bacteria"  
Int. J. Syst. Bact., 1975, 25, 383-385
89. MILLER I.R. : "The structure of DNA and RNA in the water mercury interface"  
J. Mol. Biol., 1961, 3, 229-240
90. MIURA K. : "Preparation of bacterial DNA by the phenol pH 9. RNases method"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE), 12 part A (p. 543-545), Acad. Press, New York & London, 1967
91. MONJALLON A. : "Analyse d'une série statistique"  
In "Introduction à la méthode statistique" (A. MONJALLON), (p. 57-62), Vuibert, Paris, 1969

92. MOORE R.L., BRUBAKER R.R. : "Hybridization of deoxyribonucleotide sequences of *Yersinia enterocolitica* and other selected members of *Enterobacteriaceae*"  
Int. J. Syst. Bact., 1975, 25, 336-339
93. MORONEY M.J. : "Interpolation linéaire de la médiane"  
In "Comprendre la statistique", 76-81,  
GERARD & Co, Verviers (Belgique)
94. MURATA N., STARR M.P. : "Intragenetic clustering and divergence of *Erwinia* strains from plants and man in the light of deoxyribonucleic acid segmental homology"  
Canad. J. Microbiol., 1974, 20, 1545-1565
95. NOLL H., STUTZ E. : "The use of sodium and lithium dodecyl sulfate in nucleic acid isolation"  
In "Methods in Enzymology" (L. GROSSMAN & K. MOLDAVE), 12 part. B, (p. 129-155), Acad. Press, New York & London, 1968
96. OCADA Y., AMAGASE S., TSUGITA A. : "Frameshift mutation in the lysozyme gene of bacteriophage T<sub>4</sub> : demonstration of the insertion of five bases and a summary of in vivo codons and lysozyme activity"  
J. Mol. Biol., 1970, 54, 219-246
97. OUELLETTE C.A., BURRIS R.H., WILSON P.W. : "Deoxyribonucleic acid base composition of species of *Klebsiella*, *Azotobacter* and *Bacillus*"  
Antonie V. Leeuwenhoek, 1969, 35, 275-286
98. PARIZA M.W., JANDOLO J.J. : "Base ratio and deoxyribonucleic acid homology studies of six *Staphylococcus aureus* typing bacteriophages"  
Appl. Microbiol., 1974, 27, n° 2, 317-323

99. PEACOCKE A.R., WALKER I.O. :
- a) "The thermal denaturation of sodium deoxyribonucleate.  
I. The limits"  
J. Mol. Biol., 1962, 5, 550-559
  - b) "The thermal denaturation of sodium deoxyribonucleate.  
II. Kinetics"  
J. Mol. Biol., 1962, 5, 560-563
100. POPOFF M., VERON M. : "A taxonomic study of the *Aeromonas hydrophila*  
*Aeromonas punctata* group"  
J. Gen. Microbiol., 1976, 94, 11-22
101. PRIEST F.G., SOMERVILLE H.J., COLE J.A., HOUGH J.S. : "The taxo-  
nomic position of *Obesumbacterium Proteus*, a common  
brewery contaminant"  
J. Gen. Microbiol., 1973, 75, 295-307
102. RICE S.A., DOTY P. :
- J. Am. Chem. Soc., 1957, 79, 3937
  - Op. cité par MARMUR J. et DOTY P. (voir réf. 81)
103. RICHARD C., BRISOU B., LIOULT J. : "Etude taxonomique de "*Levinea*",  
nouveau genre de la famille des Entérobactéries"  
Ann. Inst. Pasteur, 1972, 122, 1137-1146
104. ROLFE R., MESELSON M. : "The relative homogeneity of microbial DNA"  
Proc. Nat. Acad. Sci., 1959, 45, 1039-1043
105. ROMOND Ch. : "Contribution à l'étude de la systématique des  
streptocoques anaérobies. Introduction d'un nouveau  
critère : la composition du DNA bactérien"  
Thèse d'Etat en Pharmacie n° 22, Strasbourg, 1965

106. ROOP D.R., ORVIN MUNDT J., STUART RIGGSEY W. : "Deoxyribonucleic acid hybridization studies among some strains of group D and group N *Streptococci*"  
Int. J. Syst. Bact., 1974, 24, 330-337
107. ROSYPALOVA R., BOHACEK J., ROSYPAL S. : "Deoxyribonucleic acid base composition and taxonomy of violet pigmented cocci"  
Antonie V. Leeuwenhoek, 1966, 32, 105-112
108. ROSYPALOVA A., BOHACEK J., ROSYPAL S. : "Deoxyribonucleic acid base composition of some *Micrococci* and *Sarcinae*"  
Antonie V. Leeuwenhoek, 1966, 32, 192-196
109. SAITO H., MIURA K. : "Preparation of transforming deoxyribonucleic acid by phenol treatment"  
Biochem. Biophys. Acta, 1963, 72, 619-629
110. SANDERSON K.E. : "Genetic homology in the *Enterobacteriaceae*"  
In "Advances in Genetics" (E.W. CASPARI), vol. 16, (p. 35-51)  
Acad. Press, New York & London, 1971
111. SAUNDERS G.F., CAMPBELL L.L., POSTGATE J.R. : "Base composition of deoxyribonucleic acid of sulfate reducing bacteria"  
J. Bact., 1964, 87, 1073-1078
112. SCHILDKRAUT C.L., MARMUR J., DOTY P. : "Determination of the base composition of deoxyribonucleic acid from its buoyant density in CsCl"  
J. Mol. Biol., 1962, 4, 430-443
113. SEBALD M., GASSER F., WERNER H. : "Teneur GC % et classification. Application au groupe des Bifidobactéries et à quelques genres voisins"  
Ann. Inst. Pasteur, 1965, 109, 251-269

114. SEBALD M., VERON M. : "Teneur en bases de l'ADN et classification des vibrions"  
Ann. Inst. Pasteur, 1963, 105, 897-910
115. SEKI T., OSHIMA T., OSHIMA Y. : "Taxonomic study of *Bacillus* by deoxyribonucleic acid-deoxyribonucleic acid hybridization and interspecific transformation"  
Int. J. Syst. Bact., 1975, 25, 258-270
116. SEVAG M.G., LACKMAN D.B., SMOLENS J. : "The isolation of the components of streptococcal nucleoproteins in serologically active form"  
J. Biol. Chem., 1938, 124, 425-436
117. SILVESTRI L.G., HILL L.R. : "Agreement between deoxyribonucleic acid base composition and taxometric classification of Gram positive cocci"  
J. Bact., 1965, 90, 136-140
118. SIMON M.I., VAN VUNAKIS K. : "The photodynamic reaction of methylene blue with deoxyribonucleic acid"  
J. Mol. Biol., 1962, 4, 488-499
119. SPIRIN A.S., BELOZERSKY A.N., KUDLAI D.G., SKAVRONKAYA A.G., MITEREVA V.G. : "Changes in the nucleic acid composition as a result of experimental hereditary transformations of intestinal bacteria"  
Biochem., 1958, 23, 154-163
120. STANIER R.Y., DOUDOROFF M., ADELBERG E.A. : "The classification of bacteria"  
In "General Microbiology", (p. 523-544), The Macmillan Press Ltd, London & Basingstoke, 1971

121. STARR M.P., MANDEL M. : "DNA base composition and taxonomy of phytopathogenic and other Enterobacteria"  
J. Gen. Microbiol., 1969, 56, 113-123
122. STARR M.P., MANDEL M., MURATA N. : "The phytopathogenic Coryneform bacteria in the light of DNA base composition and DNA-DNA segmental homology"  
J. Gen. Appl. Microbiol., 1975, 21, 13-26
123. STOLERU R. : "Hybridations des acides nucléiques bactériens : revue bibliographique et application à la taxonomie du sous-genre IV des *Salmonella*"  
Thèse d'état en Médecine, Paris, 1974
124. STOLLAR D., GROSSMAN L. : "The reaction of formaldehyde with denatured DNA : spectrophotometric, immunologic and enzymic studies"  
J. Mol. Biol., 1962, 4, 31-38
125. STREISINGER G., OKADA Y., EMRICH J., NEWTON J., TSUGITA A., TERZAGHI E., INOUE M. : "Frameshift mutations and the genetic code"  
Cold Spr. Harb. Symp. Quant. Biol., 1966, 31, 77-84
126. SUEOKA N. : "A statistical analysis of deoxyribonucleic acid distribution in density gradient centrifugation"  
Proc. Nat. Acad. Sci., 1959, 45, 1480-1490
127. SUEOKA N. : "Variation and heterogeneity of base composition of deoxyribonucleic acids : a compilation of old and new data"  
J. Mol. Biol. 1961, 3, 31-40

128. SUEOKA N., MARMUR J., DOTY P. : "Heterogeneity in deoxyribonucleic acids. II - Dependence of the density of deoxyribonucleic acids on guanine-cytosine content"  
Nature, 1959, 183, 1429-1431
129. TRAVAGLINI E.C. : "Methods for the extraction and purification of deoxyribonucleic acids from eukaryote cells"  
Methods in cell Biology, 1973, 7, 105-127
130. ULITZUR S. : "Rapid determination of DNA base composition by ultraviolet spectroscopy"  
Biochem. Biophys. Acta, 1972, 272, 1-11
131. VERON M. : "Taxonomie numérique des vibrions et de certaines bactéries comparables. II - Corrélation entre les similitudes phénétiques et la composition en base de l'ADN"  
Ann. Inst. Pasteur, 1966, 111, 671-709
132. WATSON J.D., CRICK F.H.C. : "Molecular structure of nucleic acids. A structure for deoxyribose nucleic acid"  
Nature, 1953, 171, 737-738
133. WATSON J.D., CRICK F.H.C. : "Genetical implications of the structure of deoxyribonucleic acid"  
Nature, 1953, 171, 964-967
134. WATSON J.D., GROS F. : "La répllication et organisation génétique de l'ADN"  
In "Biologie Moléculaire du Gène" (p. 33-70), Ediscience, Paris, 1968
135. WEED L.L. : "Effects of copper on *Bacillus subtilis*"  
J. Bact., 1963, 85, 1003-1010

136. WILKINS M.H.F. : "Molecular configuration of nucleic acids.  
From extensive diffraction data and molecular model  
building a more detailed picture is emerging"  
Science, 1963, 140, 941-950
137. WILKINS M.H.F., STOKES A.R., WILSON H.R. : "Molecular structure  
of deoxypentose nucleic acids"  
Nature, 1953, 171, 738-740
138. YAMAGISHI H. : "Nucleotide distribution in the DNA of *Escherichia  
coli*"  
J. Mol. Biol., 1970, 49, 603-608
139. YAMAGISHI H. : "Nucleotide distribution in bacterial DNA's differing  
in G + C content"  
J. Mol. Evol., 1974, 3, 239-242
140. YAMAGISHI H., TAKAHASHI J. : "Heterogeneity in nucleotide compo-  
sition of *Bacillus subtilis* DNA"  
J. Mol. Biol., 1971, 57, 369-371
141. YOUNG V.M., KENTON D.M., HOBBS B.J., MOODY M.R. : "*Levinea*, a  
new genus of the family *Enterobacteriaceae*"  
Int. J. Syst. Bact., 1971, 21, 58-63

