

CARACTÈRES CHETOTAXIQUES DE DISTINCTION
ENTRE LARVES (STADE IV) DE L'ESPÈCE A ET DE L'ESPÈCE B
DU COMPLEXE *Anopheles gambiae*

A MADAGASCAR

par

G. CHAUVET * et J. DÉJARDIN **

RÉSUMÉ.

Des études biométriques ont été effectuées à Madagascar sur douze soies larvaires dont le nombre de branches a été relevé sur quatre échantillons expérimentaux de chacune des espèces A et B du complexe A. gambiae. Deux soies ont été retenues : la soie suturale interne et la soie mésothoracique n° 1. Ces soies permettent de distinguer dans les deux régions étudiées une espèce de l'autre, en travaillant sur des générations F1 obtenues de pontes de femelles. Ces deux variables semblent suffisamment puissantes pour permettre de distinguer les deux espèces à l'échelle de Madagascar.

SUMMARY.

This paper deals with the results of biometrical studies on the larval setae of species A and B of the Anopheles gambiae complex in Madagascar.

The number of branches of the setae have been compared in the two species for twelve of these setae. The material used was composed of larvae obtained from females captured in two different areas. It was already known that in each of these regions only one of the two species existed. The numbers of branches observed in the inner sutural seta and the mesothoracic seta 1 allow to discriminate, with a low enough number of larvae, species A from species B in Madagascar. These two variables appear to be powerful enough to distinguish the two species at the scale of Madagascar.

* Maître de Recherches, Entomologiste médical au Centre O.R.S.T.O.M. de Tananarive (Madagascar).

** Maître de Recherches, Chef du Service de Biométrie aux S.S.C. de l'O.R.S.T.O.M., 93 - Bondy (France).

PLAN

- 1 -- INTRODUCTION.
- 2 -- LE COMPLEXE *Anopheles gambiae* A MADAGASCAR.
- 3 -- ETUDE COMPARÉE DE LA CHETOTAXIE DES LARVES (STADE IV) DE L'ESPÈCE A ET DE L'ESPÈCE B.
 - 3.1. — Protocole de travail.
 - 3.2. — Etude de l'ensemble des soies examinées.
 - 3.3. — Utilisation de la soie suturale interne.
 - 3.3.1. — Nombre de branches de la soie.
 - 3.3.2. — Proportion de larves porteuses de soie à deux branches.
 - 3.3.3. — Autres caractéristiques.
 - 3.3.4. — Résumé et conclusions relatives à l'utilisation de la soie suturale interne.
 - 3.4. — Utilisation de la soie mésothoracique n° 1.
 - 3.4.1. — Nombre de branches. Modèle de distribution et homogénéité.
 - 3.4.2. — Limites de tolérance pour des moyennes d'échantillons issus de A ou de B.
 - 3.4.3. — Résumé et conclusions relatives à l'utilisation de la soie mésothoracique n° 1.
 - 3.4.4. — Indépendance des critères utilisés.
- 4 -- ETUDE COMPARÉE DE LA CHETOTAXIE DES NYMPHES DE L'ESPÈCE A ET DE L'ESPÈCE B.
- 5 -- CONCLUSIONS.
- 6 -- REMERCIEMENTS.

1. - INTRODUCTION

Les travaux de MUIRHEAD-THOMSON (1948), DAVIDSON (1956), DAVIDSON et JACKSON (1962), BURGESS (1962), PATERSON (1962), puis de DAVIDSON (1963) et de PATERSON et coll. (1963) ont amené progressivement la preuve de l'existence de cinq « types » ou « formes » différentes chez l'espèce *Anopheles gambiae* Giles. Peu de temps après, ces cinq « types » reçurent le statut d'espèces distinctes à la suite des travaux de PATERSON (1963) et de MASON (1964). En effet, le premier signala la coexistence de trois des « formes » en un même lieu sans aucune perte de leurs caractéristiques propres et le second démontra une incompatibilité chromosomique au moins chez les deux « formes » qu'il étudia. Depuis, COLUZZI (1966) a démontré qu'il existe entre les espèces A et B un nombre différent de bandes géniques sur certains secteurs de l'hétérochromosome X.

Ainsi, on distingue actuellement en Afrique, cinq espèces se développant soit dans des gîtes d'eau douce : espèces A, B et C, soit en eaux saumâtres : espèces *A. melas* Theo (anciennement *A. gambiae* var. *melas* Theo) et *A. merus* Dönitz (anciennement *A. gambiae* d'eau saumâtre d'Afrique orientale, déjà redécrit avec raison comme une espèce distincte par KÜHLOW (1962) sous le nom d'*A. tangensis*). Ces nouvelles espèces sont maintenant réunies sous le nom complexe *A. gambiae*.

Ces résultats sont le fruit d'études de croisements. En effet, le croisement entre deux quelconques de ces espèces produit à la génération F1 des mâles stériles. Dans cette première génération, la proportion de mâles et de femelles est normale sauf lorsque le croisement s'effectue entre femelles des espèces *A. melas*, *A. merus* ou C et les mâles d'espèce A ou B. Dans ces croisements, la génération F1 est composée presque uniquement de mâles où leur proportion est très supérieure à la normale.

Cette méthode d'identification par croisement forcé entre parents d'espèces différentes est lente et délicate. En effet, elle exige d'abord le maintien et l'entretien en insectarium des souches de référence. Ce maintien est déjà difficile du fait de l'eurygamie de la plupart des souches des espèces du complexe (1). Ce phénomène nécessite alors l'insémination forcée, périodique de quelques jeunes femelles de chacune des souches, par des mâles de la même souche.

Les adultes supportent très mal les voyages. Il est donc nécessaire d'envoyer au laboratoire de référence les espèces à déterminer sous forme de pontes séparées obtenues des femelles à identifier.

Arrivées au laboratoire, ces pontes sont mises en élevage. Au bout d'une quinzaine de jours, les imagos mâles ou femelles obtenus sont unis par insémination forcée avec quelques individus de sexe opposé d'une ou de plusieurs souches de référence. De nouvelles pontes sont ainsi obtenues. Les nouveaux adultes mâles qui en résultent sont alors disséqués afin d'en extraire les testicules. D'après leur aspect et après examen des cellules spermatiques, on conclut à leur stérilité ou non, donc à l'identité ou la non-identité de la souche avec telle ou telle des souches de référence employées. Ces différentes manipulations auront duré environ un mois et exigé beaucoup d'ordre, de méthode, d'attention et de délicates interventions, requis un important personnel de parfaite technicité. Cette méthode est actuellement utilisée systématiquement par G. DAVIDSON au Ross Institute de Londres et, sur une moindre échelle, par J. Coz au Centre Muraz de Bobo-Dioulasso.

En 1962, L. CORONEL avait essayé de trouver des caractères morphologiques de distinction entre les espèces *A* et *B*; deux furent relevés: une différence dans la largeur de la tache pâle entre les première et deuxième taches sombres principales de la costa de l'aile (tache sectoriale) et une autre différence dans la forme et la disposition des épines de la palette natatoire de la nymphe. Mais à l'usage, ces caractères n'apparurent pas totalement spécifiques et ils furent abandonnés.

En 1963, on savait seulement reconnaître *A. melas* des quatre autres espèces grâce aux caractères morphologiques décrits depuis 1944 par RIBBANDS, puis par MUIRHEAD-THOMSON (1951), sur les œufs et le peigne abdominal des larves au 4^e stade.

En 1964, COLUZZI reprit le problème et publia le résultat d'études biométriques détaillées entreprises sur l'ensemble des cinq espèces connues. Ce chercheur travailla sur les œufs, les larves, les nymphes et les adultes de nombreuses souches des différentes espèces provenant de localités diverses d'Afrique.

De bons caractères de distinction furent trouvés entre les espèces d'eau saumâtre *A. melas* et *A. merus* et entre celles-ci et les espèces d'eau douce considérées globalement. Mais, de l'avis même de l'auteur, « l'identification de ces dernières entre elles apparaissait beaucoup plus difficile du fait de l'existence de larges gradients intraspécifiques qui, en définitive, concilient les différences interspécifiques et limitent ainsi la sûreté des observations taxonomiques » (COLUZZI, 1964). De ce fait, il lui semblait que l'étude morphologique comparée des trois espèces d'eau douce du complexe considéré dans son ensemble éthiopien était irréalisable. Par contre, il estimait possible des comparaisons dans des zones géographiques limitées.

Adoptant cette possibilité, nous avons cherché des caractères morphologiques permettant de séparer les espèces *A* et *B* dans le cadre relativement limité de l'île de Madagascar.

Il est bien évident que la découverte des critères morphologiques de distinction permettrait d'étudier plus rapidement l'écologie et les potentialités vectrices de ces deux espèces dans le cadre de Madagascar en évitant soit les envois de pontes à Londres, soit l'application locale de la technique difficile et surtout longue des croisements.

(1) Ainsi toutes les souches d'espèces *A* ou *B* de Madagascar que DAVIDSON ou nous-mêmes avons essayé d'entretenir se sont révélées être eurygames. Toutefois, parmi les nombreuses pontes reçues de tout le continent africain, G. DAVIDSON a réussi à isoler certaines souches de ces deux espèces se reproduisant en cage. Mais aucune réussite n'a encore été enregistrée pour l'espèce *C*.

2. - LE COMPLEXE *Anopheles gambiae* A MADAGASCAR

Les études que nous avons entreprises de 1964 à ce jour, en collaboration étroite avec G. DAVIDSON, ont montré l'existence des espèces A et B du complexe *A. gambiae* à Madagascar. Nous avons également découvert la présence d'*A. merus* dans une station du littoral sud-ouest (Tuléar) (2).

La répartition actuellement connue de ces espèces intéresse 120 stations dispersées sur tout le territoire. Elle est basée sur la reconnaissance de 63 stations à espèce B isolée, 30 stations à espèce A isolée, et 27 stations à espèces A et B sympatriques. L'espèce B a une large dispersion intéressant toute l'île ; elle se trouve isolée dans les plateaux du Centre et le Sud. L'espèce A, par contre, semble se localiser essentiellement dans les régions côtières et les versants. Ces dernières régions se caractérisent, par rapport à celles où l'espèce B est isolée, par leur climat plus chaud.

3. - ÉTUDE COMPARÉE DE LA CHÉTOTAXIE DES LARVES (STADE IV) DE L'ESPÈCE A ET DE L'ESPÈCE B

3.1. Protocole de travail.

Le premier point était évidemment la reconnaissance d'au moins une station à espèce A isolée et d'une station à espèce B isolée. Dans ce but nous avons recueilli les pontes individuelles de nombreuses femelles capturées dans plusieurs stations réparties le long de la route Tananarive-Brickaville. Cette route suit approximativement la latitude 19° S. Cet axe de prospection présentait de nombreux avantages : bien que relativement court (environ 160 km à vol d'oiseau), il traverse cinq zones écologiques très différentes par la végétation, la climatologie et l'altitude (différence d'altitude entre les deux extrémités de cet axe : 1.500 m) et il permet de rejoindre Tananarive dans la journée (270 km de route empierrée). G. DAVIDSON a identifié la totalité de l'important matériel recueilli. Après plusieurs envois, ses identifications nous assuraient de la présence de la seule espèce A dans la région de Ranomafana (en particulier, PK 170 : 250 m d'altitude, région boisée) et de la seule espèce B dans les environs mêmes de Tananarive dans un village du canton d'Alasora (1.300 m d'altitude, région de riziculture).

Il nous a paru également indispensable d'utiliser en priorité les caractères de chétotaxie larvaire déjà employés par COLUZZI, afin de pouvoir comparer nos observations locales aux siennes, plus générales.

Il fallait enfin faire développer parallèlement des individus de chacune des deux espèces dans des conditions de milieu différentes, naturelles ou artificielles, de façon à prendre en compte les variations éventuelles de la chétotaxie dues au milieu.

Les conditions que nous avons retenues conduisent à quatre groupes d'échantillons pour chacune des deux espèces ; tous les spécimens de l'espèce A proviennent de souches de Ranomafana et ceux de l'espèce B de souches d'Alasora.

a) Premier groupe d'échantillons

Il est formé de larves issues de deux pontes de chacune des deux espèces reçues de Londres après identification. Les femelles mères avaient été capturées à Madagascar et leurs premières pontes expédiées à Londres en janvier 1965.

(2) En 1965, de nombreuses larves d'*A. gambiae*, n'appartenant pas à l'espèce *A. melas* et vivant cependant dans une mare d'eau salée titrant 18,4 g/l de ClNa (salinité de 58 %) ont été recueillies en bordure des « Salines de Tuléar ».

Dernièrement (1967), des larves au premier stade provenant de ponte de femelles sauvages capturées dans cette même région, ont été soumises au test physiologique à l'eau salée à 75 % (MURHEAD-THOMSON, 1951) : aucune mortalité n'a été constatée au bout de six heures de contact.

L'élevage a été réalisé à Tananarive, en insectarium, à une température constante de 26° C.

Ces échantillons seront désignés par l'initiale L (Londres) : AL pour la descendance de l'espèce A, et BL pour la descendance de l'espèce B.

b) *Deuxième groupe d'échantillons*

Des larves F1 de six femelles sauvages de l'espèce A et de sept femelles de l'espèce B, toutes capturées en mars 1965, le constituent.

Elles ont été élevées en insectarium à Tananarive à la température constante de 26° C.

Les initiales de reconnaissance seront TI (Tananarive Insectarium), avec ATI pour les larves de l'espèce A, et BTI pour les larves de l'espèce B.

Le premier et le second groupe d'échantillons ont donc bénéficié des mêmes conditions de développement. Les seules différences se trouvent dans le nombre de générations écoulé entre la capture des femelles sauvages et l'étude de leur descendance, ainsi que dans la date de capture des femelles sauvages.

c) *Troisième groupe d'échantillons*

Nous avons ici des larves F1 de cinq femelles sauvages de l'espèce A et de six femelles de l'espèce B capturées en avril 1965. Elles ont été élevées à Ranomafana à l'extérieur. L'espèce B a donc été importée dans le milieu de l'espèce A. La température moyenne pendant l'élevage fut d'environ 29°,5 C.

Ces échantillons seront désignés par RE (Ranomafana Extérieur), avec ARE pour les larves de l'espèce A et de BRE pour les larves de l'espèce B.

d) *Quatrième groupe d'échantillons*

Il est constitué de larves F1 d'une femelle de l'espèce A et de quatre femelles de l'espèce B capturées en avril 1965.

Nous les avons élevées à Tananarive à l'extérieur. Ici l'espèce A a été importée dans le milieu de l'espèce B. La température moyenne pendant l'élevage fut d'environ 24° C.

Nous repérerons ces échantillons par les initiales TE (Tananarive Extérieur), avec ATE pour les larves de l'espèce A, et BTE pour les larves de l'espèce B.

L'espèce A importée de Ranomafana à Tananarive et l'espèce B importée d'Alasora-Tananarive à Ranomafana se sont mal développées ; nous n'avons obtenu que peu d'individus.

Les quatre groupes sont donc constitués d'exemplaires qui se sont développés soit dans de bonnes conditions constantes en milieu artificiel (insectarium), soit dans des conditions proches de la nature (espèces locales élevées à l'extérieur), soit enfin dans des conditions très anormales (espèces importées).

Nous avons constaté d'amples variations dans les soies étudiées suivant les échantillons considérés, ce qui était le but recherché.

3.2. **Etude de l'ensemble des soies examinées.**

Les douze soies suivantes ont été étudiées (nomenclature de PURI, 1928)

- soie clypéale interne,
- soie frontale interne,
- soie frontale médiane,
- soie frontale externe,
- soie suturale interne,
- soie suturale externe,
- soie prothoracique n° 1,
- soie prothoracique n° 2,
- soie prothoracique n° 4,

- soie mésothoracique n° 1,
- soie abdominale n° 1 du segment I,
- soie abdominale n° 2 du segment VII.

Le caractère utilisé dans chaque cas a été le nombre de branches de la soie. Sur chaque spécimen, le décompte des branches de chacune des soies a été réalisé systématiquement deux fois par deux observateurs. Un troisième examen a eu lieu lorsque les deux observations ne concordaient pas. Nous avons travaillé sur les exuvies larvaires du 4^e stade montés entre lame et lamelle dans du Polyvinyl-alcool-lactophénol (PVA) pour constituer des préparations microscopiques permanentes (3).

Chaque échantillon étant composé en général de plusieurs souches (nombre variant de 1 à 7), nous avons tout d'abord étudié, pour chacun des caractères retenus, leur distribution particulière par souche avant de cumuler les résultats d'examen de l'ensemble des souches formant le même échantillon. Cette étude nous a montré que nous étions toujours en droit de faire le cumul.

TABLEAU n° 1. — Identité soie droite, soie gauche de la soie suturale interne. Moyenne, distribution de fréquence et χ^2

Echantillon	moyenne \bar{x}	effectif n	Classes			χ^2		
			1	2	3			
AL	s. droite	1,592	27	12	14	1	2,34	N.S.
	s. gauche	1,785	28	7	20	1		
ATI	s. droite	1,375	40	25	15		0,20	N.S.
	s. gauche	1,428	42	24	18			
ARE	s. droite	1,390	64	39	25		0,62	N.S.
	s. gauche	1,323	65	44	21			
BL	s. droite	1,048	41	39	2		0,77	N.S.
	s. gauche	1,103	29	26	3			
BTI	s. droite	1,054	37	35	2		2,34	N.S.
	s. gauche	1,162	43	36	7			
BTE	s. droite	1,098	61	55	6		0,14	N.S.
	s. gauche	1,105	57	51	6			

Un autre problème s'est posé ensuite : les différents caractères ont été lus a priori sur le côté droit des individus ou, à défaut, sur le côté gauche. Cette opération a été estimée nécessaire pour utiliser au mieux la faible importance numérique de quelques échantillons. Il a donc fallu nous assurer de l'identité statistique de la valeur du nombre moyen de branches des soies droites et des soies gauches ainsi que de l'identité de leur distribution de fréquence.

(3) Afin d'alléger cet exposé, nous ne donnons pas les résultats obtenus avec les soies que nous n'avons pas retenues en définitive. Mais ces résultats peuvent être obtenus sur simple demande, auprès du Service de Documentation des Services Scientifiques Centraux de l'O.R.S.T.O.M. à 93 - Bondy, France.

Les tableaux 1 et 2 donnent les résultats pour les seules soies qui ont été retenues. Ils montrent que l'on peut utiliser indifféremment les soies situées du côté droit ou du côté gauche.

Enfin, pour chaque échantillon, l'étude, la distribution de chacun des douze caractères retenus a été faite.

TABLEAU n° 2. — Identité soie droite, soie gauche de la soie mésothoracique n° 1. Moyenne, distribution de fréquence (%) et *t* de STUDENT.

Echan- tillon	effect. <i>n</i>	moy. \bar{x}	test <i>t</i>	Classes										
				22-23	24-25	26-27	28-29	30-31	32-33	34-35	36-37	38-39	40-41	
AL	s. d.	25	33,84	0,86			8,00	8,00	4,00	20,00	28,00	20,00	12,00	
	s. g.	23	33,08	N.S.				8,69	13,04	39,13	30,43	—	8,69	
ARE	s. d.	60	34,50	0,35				3,33	10,00	23,33	31,66	15,00	11,66	5,00
	s. g.	57	34,66	N.S.				3,50	8,77	28,07	19,29	22,80	10,52	7,01
BL	s. d.	29	26,48	1,00	10,34	24,13	24,13	31,03	10,34					
	s. g.	28	27,14	N.S.	10,71	10,71	35,71	25,00	14,28	3,57				
BTI	s. d.	38	29,18	0,29	2,63	—	21,05	31,57	28,94	15,78				
	s. g.	35	29,02	N.S.		8,57	8,57	37,14	28,57	11,42	5,71			
BTE	s. d.	115	30,12	1,46	1,73	2,60	12,17	32,17	19,13	20,86	6,95	0,86	3,47	
	s. g.	114	29,52	N.S.	2,63	6,14	11,40	21,05	32,45	15,78	7,89	1,75	—	0,87

Le problème devant lequel nous nous trouvons maintenant placés pour différencier les espèces est celui de l'utilisation des mesures faites sur les échantillons A et B que nous possédons pour classer dans l'une de ces espèces les individus capturés ultérieurement. C'est le problème classique de la discrimination dont la théorie est bien établie à la suite principalement des travaux de FISHER sur la fonction discriminante.

Il se trouve que ses résultats ne peuvent être utilisés dans cette étude : ils supposent en effet la normalité des observations dans un espace à *p* dimensions, *p* étant le nombre de variables observées. Cette hypothèse doit être manifestement refusée ici : certaines des variables utilisées, comme le nombre de branches de la soie suturale interne, ne montrent que deux valeurs (une ou deux dans le cas de l'espèce B).

Au cours de ces dernières années, quelques tentatives d'extension de la théorie de la discrimination au cas de variables discontinues ou même qualitatives ont été publiées, notamment : COCHRAN et HOPKINS (1961), LINDER (1963), LINHART (1959), HILLS (1967).

Mais en fait, on peut se demander si pour discriminer il est absolument nécessaire de se servir de moyens évolués mettant en jeu plusieurs variables alors que certaines d'entre elles, nous le verrons ultérieurement, peuvent être utilisées isolément. Il nous apparaît que pour être facilement utilisables, les règles de discrimination que nous comptons établir doivent être les plus simples possible. Nous allons donc, dans un premier temps tout au moins, utiliser les variables une par une, d'autant plus que cette façon de faire est justifiée par l'indépendance des deux variables retenues ici (cf. 3.4.4.).

La question est maintenant celle du choix des observations intéressantes, autrement dit : comment sélectionner les variables les plus aptes à discriminer les

TABLEAU n° 3. — Résultats des tests de comparaison A - B. 1. Valeur du *t* de STUDENT ou du *f* de COCHRAN.

Soie	Echantillon	Effectif <i>n</i>	moy. \bar{x}	variance <i>s</i> ²	tests			
					F	<i>t</i> ou <i>t'</i>		
Frontale interne	L	AL	23	16,73	1,421	1,98	0,14	
		BL	32	16,81	4,777			
	TI	ATI	39	15,92	3,763	1,44	0,92	
		BTI	86	15,53	4,086			
	RE	ARE	109	16,31	5,407	1,44	0,72	
		BRE	33	16,00	4,485			
	TE	ATE	15	15,06	4,067	1,85	3,47	
		BTE	71	16,81	2,938			
	Frontale médiane	L	AL	25	18,88	1,777	1,82	2,63
			BL	35	17,88	4,501		
TI		ATI	47	18,48	3,228	1,44	0,38	
		BTI	94	18,36	3,907			
RE		ARE	112	17,83	3,855	1,62	3,70	
		BRE	36	18,94	2,108			
TE		ATE	16	17,18	3,496	1,84	1,52	
		BTE	80	18,05	4,447			
Frontale externe		L	AL	23	23,47	4,443	1,84	6,55
			BL	29	19,93	3,226		
	TI	ATI	46	20,60	5,723	1,48	0,02	
		BTI	77	20,59	5,310			
	RE	ARE	105	20,98	5,698	1,46	2,61	
		BRE	33	19,60	7,208			
	TE	ATE	14	20,14	7,824	1,95	1,53	
		BTE	58	19,17	3,694			
	Prothoracique n° 1 ...	L	AL	30	19,00	5,267	1,69	3,25
			BL	38	17,21	4,745		
TI		ATI	51	19,50	5,269	1,48	2,52	
		BTI	87	18,52	4,042			
RE		ARE	97	19,32	4,489	1,57	0,76	
		BRE	30	18,93	6,498			
TE		ATE	11	20,81	3,963	2,30	1,46	
		BTE	112	19,66	6,292			
Mésothoracique n° 1 ..		L	AL	29	33,45	11,899	1,76	9,20
			BL	34	26,62	5,880		
	TI	ATI	60	32,75	5,513	1,00	11,40	
		BTI	120	28,02	9,806			
	RE	ARE	124	32,73	8,880	1,59	3,38	
		BRE	37	31,05	8,441			
	TE	ATE	20	34,25	5,566	1,84	5,66	
		BTE	115	30,14	10,073			

TABLEAU n° 3 (suite). — Résultats des tests de comparaison A - B. 2. Valeur de χ^2 .

Soie	Sp.	Echantillon							
		L		TI		RE		TE	
		<i>n</i>	χ^2	<i>n</i>	χ^2	<i>n</i>	χ^2	<i>n</i>	χ^2
Clypéale interne.	A	5		14	6,28	37	9,16	5	0,33
	B	23		41	2 ddl	25	3 ddl	24	1 ddl
Suturale interne (Groupement 1, 2 et 3)	A	32	27,22	64	38,76	120	12,34	17	55,34
	B	44	1 ddl	120	1 ddl	32	1 ddl	124	1 ddl
Suturale externe.	A	29	47,15	65	56,42	122	16,94	20	0,04
	B	44	2 ddl	117	2 ddl	31	2 ddl	89	2 ddl
Prothoracique n°1	A	27	2,82	51	4,70	123	0,95	16	12,60
	B	36	3 ddl	110	3 ddl	40	4 ddl	122	2 ddl
Prothoracique n°2	A	24	6,18	49	24,09	112	1,50	19	17,13
	B	38	2 ddl	95	4 ddl	39	3 ddl	129	3 ddl
Abdominale 1 Seg. I	A	30	19,37	65	24,15	128	2,12	18	8,02
	B	39	3 ddl	129	5 ddl	39	4 ddl	129	4 ddl
Abdominale 2 Seg. VI	A	30	6,80	59	23,20	114	2,01	30	12,42
	B	38	2 ddl	112	3 ddl	39	3 ddl	127	2 ddl

Suturale interne. prop. de porteurs de soie à 2 bran- ches (groupement 1 et 3, 2	22,64	36,44	10,81	41,71
	1 ddl	1 ddl	1 ddl	1 ddl

deux espèces ? Le critère que nous avons retenu pour décider de leur valeur est le suivant : seules seront prises en considération les variables pour lesquelles les quatre différences observées entre A et B et correspondant aux quatre conditions d'élevage, sont significatives. Nous n'avons pas, bien sûr, la certitude d'être arrivés ainsi à sélectionner toutes les variables intéressantes : certaines différences non significatives dans notre étude auraient pu l'être par utilisation d'échantillons de taille plus grande, donc de tests plus puissants. Mais notre critère a au moins le mérite de ne retenir que les variables suffisamment puissantes pour montrer des différences avec les tailles d'échantillon dont il est matériellement possible de se servir de façon courante.

Nous avons donc, pour les douze variables mesurées, comparé les deux échantillons A et B dans chacune des quatre conditions dans lesquelles ils ont été placés. Les tests utilisés ont été :

- a) Lorsque l'hypothèse de normalité est admissible, le test F, puis :
— soit le test *t* de STUDENT si les deux variances théoriques peuvent être admises égales ;

— soit l'approximation t' de COCHRAN s'il n'est pas possible d'admettre l'hypothèse d'égalité des variances.

b) Dans les autres cas : le test χ^2 de comparaison de deux répartitions observées.

Le tableau 3 donne le résultat de ces différents tests pour toutes les variables. Le tableau 4 récapitule les conclusions.

TABLEAU n° 4. — Différences entre A et B (tous tests à 5 % par test ; un astérisque indique une différence).

Soie	Echantillon							
	L		TI		RE		TE	
	variance	moy.	variance	moy.	variance	moy.	variance	moy.
Frontale interne.	*							*
Frontale médiane	*	*			*	*		
Frontale externe.		*				*		
Prothoracique n° 4		*		*				
Mésothoracique n° 1		*	*	*		*		*
Clypéale interne.	(1)		*		*			
Suturale interne.	*		*		*		*	
Suturale externe.	*		*		*			
Prothoracique n° 1							*	
Prothoracique n° 2	*		*				*	
Abdominale 1 seg. I	*		*					
Abdominale 2 seg. VII	*		*				*	

(1) Test exact de randomisation étant donné la faiblesse de l'effectif de l'échantillon A ($n = 5$).

Comme on le constate aisément à sa lecture, deux soies seulement satisfont à la condition que nous nous sommes fixée : la soie mésothoracique n° 1 et la soie suturale interne. Ce sont les deux seules soies que nous utiliserons dans la suite.

3.3. Utilisation de la soie suturale interne.

3.3.1. — NOMBRE DE BRANCHES DE LA SOIE.

Le nombre de branches de la soie suturale interne est une variable qui pourra être utilisée dans la mesure où :

a) les distributions dans les populations correspondant aux différents échantillons de la même espèce sont identiques, autrement dit dans la mesure où les quatre distributions observées pour chaque espèce ne diffèrent pas significativement ;

b) un modèle théorique peut être trouvé pour ces distributions de façon à ce que la distribution d'échantillonnage du nombre de branches puisse être étudiée et, par conséquent, des limites fixées aux observations échantillonnées ultérieurement.

Les distributions utilisées pour A et pour B sont, rappelons-le :

Espèce A		Echantillon			
		L	TI	RE	TE
Nombre de branches ..	1	12	33	64	3
	2	18	30	53	12
	3	2	1	3	2
Total....		32	64	120	17

Espèce B		Echantillon			
		L	TI	RE	TE
Nombre de branches ..	1	41	110	28	113
	2	3	10	4	11
Total....		44	120	32	124

Si pour l'espèce B le point a) ne soulève pas de difficultés ($\chi^2 = 0,798$ à 3 ddl, non significatif), il n'en n'est pas de même pour l'espèce A : le χ^2 calculé sur le tableau ci-dessus ($\chi^2 = 10,5$ à 6 ddl, non significatif) est sujet à critique : les effectifs attendus à la dernière ligne sont assez faibles, en particulier celui de la dernière colonne ($c = 0,5387$). On peut, bien sûr, envisager de regrouper cette troisième ligne avec une autre, mais laquelle choisir : la première ou la seconde ? Les conclusions sont d'ailleurs différentes : le groupement 1, 2 et 3 conduit à $\chi^2 = 9,38$ à 3 ddl, significatif ; le groupement 1 et 3, 2 conduit à $\chi^2 = 5,06$ à 3 ddl, non significatif.

En définitive, le test retenu pour la comparaison globale est le test de KOLMOGOROV-SMIRNOV : comparaison des deux échantillons extrêmes, RE et TE, au seuil $5/6 \% = 0,83 \%$ (4).

(4) Lorsqu'on compare globalement k quantités, on se trouve devant un échantillon de $c = C_k^2$ différences. La plus grande des c différences observées, d_o , doit être comparée à la valeur critique d_M définie par son seuil α' tel que :

$$\Pr(D > d_M) = \alpha'$$

ou encore :

$$\Pr(D \leq d_M) = 1 - \alpha'$$

α' étant défini par (α est le seuil global désiré) : $(1 - \alpha')^c \leq 1 - \alpha$

L'égalité n'aurait lieu que si les c différences étaient indépendantes, en fait, elles ne le sont pas. En écrivant que $(1 - \alpha')^c \neq 1 - c\alpha'$ (α' est toujours petit), on a :

$$1 - c\alpha' \leq 1 - \alpha$$

$$\alpha' \geq \alpha/c$$

En prenant $\alpha' = \alpha/c$, on a un test un peu conservatif du fait de la non-indépendance des différences. RYAN (1959) indique d'ailleurs que l'effet de cette non-indépendance est dans la pratique peu important.

Ils ne diffèrent pas : différence maximale observée entre les fréquences cumu-
lées : 0,3563, non significative au seuil 1 % ($\delta = 0,4224$), donc *a fortiori* au seuil 0,83 %.
La faiblesse de l'effectif TE ($n = 17$) permettrait de douter quelque peu de la validité de
cette conclusion, mais le fait que la différence observée atteigne juste la différence théo-
rique au seuil 5 %, nous rassure et en définitive nous admettons l'identité des lois de dis-
tributions pour A.

Reste maintenant le problème du modèle à adopter. En fait peu de modèles
apparaissent *a priori* convenir à de telles distributions. De plus la validité de l'ajuste-
ment ne sera pas aisée à tester étant donné le faible nombre de classes : deux ou trois :
pour B par exemple on ne peut même pas envisager d'ajuster un modèle à un paramètre,
car l'ajustement de ce paramètre absorbera le seul degré de liberté dont on pourrait
disposer. Toutefois le modèle de POISSON a été examiné pour A. La variable est le nom-
bre de branches diminué de une unité, puisqu'il faut au moins une branche pour que
la soie existe. Les caractéristiques de la distribution sont résumées ci-dessous :

x_i	n_i	
0	112	$\bar{x} = 0,5536$ $s^2 = 0,31715$
1	113	
2	8	
Total..	233	

Remarquons tout de suite que la moyenne vaut 1,75 fois la variance : le modèle
de POISSON ne convient très probablement pas. C'est ce que confirme le χ^2 d'ajustement à
une loi de paramètre $\mu = 0,554$ ($\chi^2 = 35,8$ à 1 ddl, significatif). Il n'est pas possible de
songer à l'ajustement d'un autre modèle à deux paramètres : il resterait zéro ddl pour le
 χ^2 d'ajustement.

3.3.2. — PROPORTION DE LARVES PORTEUSES DE SOIE A DEUX BRANCHES.

Le nombre de branches de la soie suturale interne est donc une variable que
nous ne pouvons utiliser : nous ne possédons pas le modèle de distribution, nous ne
pouvons donc pas fixer de limites à la moyenne.

Le nombre de branches n'est toutefois pas la seule variable : il en existe
d'autres. La première qui vient à l'esprit est la proportion d'individus porteurs de
soie à un nombre donné b de branches dans un échantillon de n larves observées. La
population devient alors une population à deux catégories, schématisée par une urne
contenant des boules de deux couleurs dans laquelle on tire un échantillon. Ce modèle
est classique.

Nous avons examiné ci-dessus l'homogénéité de ces proportions lorsque nous
avons, pour l'espèce A, procédé aux deux regroupements : 1, 2 et 3 d'une part, et 2,
1 et 3 d'autre part. Dans le premier cas, nous avons examiné l'homogénéité des pro-
portions de porteurs de soie à une branche, nous avons vu qu'elles différaient
($\chi^2 = 9,38$ à 3 ddl, significatif).

Dans le second cas, nous avons comparé les proportions de porteurs de soie
à deux branches et trouvé qu'elles étaient homogènes ($\chi^2 = 5,06$ à 3 ddl, non significatif).

Pour B le résultat est toujours le même : $\chi^2 = 0,798$ à 3 ddl, non significatif.

Si donc nous voulons utiliser une information tirée de tous les échantillons
réunis, nous utiliserons la proportion d'individus porteurs de soie à deux branches qui
s'avère constante d'un échantillon à l'autre. D'autre part, du moment que les tirages
des unités échantillons se font indépendamment et au hasard de façon non exhaustive,
ce qui peut être considéré comme vérifié, la distribution d'échantillonnage d'une pro-
portion est connue : c'est la loi binomiale qui peut être approchée par la loi de GAUSS
si n est suffisamment grand ($np > 5$, ou mieux 10).

Pour A au total 113 larves, sur 223 examinées, ont montré une soie à deux branches, soit une proportion de 0,4850 ; pour B : 28 sur 320, soit une proportion de 0,08750. Ces deux proportions diffèrent notablement. En fait, ce ne sont pas les proportions vraies : ce sont des valeurs échantillonnées dans des populations dont les véritables proportions sont inconnues. Nous avons donc à déterminer des limites de confiance pour ces proportions, limite inférieure pour A et limite supérieure pour B (puisque le risque d'erreur pour A vient des faibles proportions et pour B des fortes proportions).

Ces limites sont données par les formules classiques :

$p_0 - z\sigma_{p_0}$ pour la limite inférieure, $p_0 + z\sigma_{p_0}$ pour la limite supérieure, z étant la variable de GAUSS centrée réduite. En choisissant le niveau de confiance 0,95 ($z = 1,645$) nous trouvons la limite inférieure pour A, appelée p_A : 0,4311 et la limite supérieure pour B, appelée p_B : 0,1135.

Nous pouvons donc dire, au niveau de confiance 95 %, que pour A la véritable proportion de larves porteuses de soie à deux branches est supérieure ou égale à 0,4311 et que pour B, au même niveau de confiance, la proportion de larves porteuses de soie à deux branches est inférieure ou égale à 0,1135 (5).

A remarquer qu'il faudra une taille minimale d'échantillon de 12 individus pour A et de 45 pour B lorsque par la suite nous utiliserons l'approximation gaussienne de la loi binomiale. En fait, l'étude exacte des tailles inférieures à 45 pour B nous montrera que la limite peut être placée plus bas que 45.

Le problème que nous avons à résoudre maintenant est donc le suivant : supposant que dans les populations les proportions vraies sont $p_A = 0,4311$ et $p_B = 0,1135$ trouver pour les proportions observées sur des échantillons de taille n une limite inférieure pour A et une limite supérieure pour B, limites qui seront dans chacun des cas dépassées avec une probabilité de 5 % (6).

Si les proportions vraies sont supérieures à 0,4311 et (ou) inférieures à 0,1135, les limites calculées ci-dessus restent *a fortiori* valables, seuls les risques auront changé : ils seront plus faibles.

En fait, dans le double but de faciliter le travail ultérieur des observateurs et de calculer plus commodément les probabilités dans le cas de petits échantillons, nous allons déterminer non pas les proportions limites, mais les nombres limites d'exemplaires porteurs de soie à deux branches. Ceci est aisément réalisé en utilisant l'approximation normale : nous écrivons que l'écart réduit correspondant à la limite est égal à l'écart normal réduit qui a la probabilité 5 % d'être dépassé

- en moins pour A,
- en plus pour B.

En appelant n_{IA} le nombre limite inférieur pour A et n_{SB} le nombre limite supérieur pour B nous avons, en tenant compte de la correction de continuité (n'_{IA} et n'_{SB} sont les limites sur l'échelle continue) :

$$\frac{n'_{IA} - 0,5 - np_A}{\sqrt{np_A(1-p_A)}} = -1,645 \quad \text{et} \quad \frac{n'_{SB} + 0,5 - np_B}{\sqrt{np_B(1-p_B)}} = +1,645$$

soit, en remplaçant p_A et p_B par leur valeur et en résolvant :

$$n'_{IA} = 0,4311 n + 0,5 - 0,81461 \sqrt{n}$$

$$n'_{SB} = 0,1135 n - 0,5 + 0,52176 \sqrt{n}$$

Les nombres n_{IA} et n_{SB} seront :

n_{IA} : l'entier immédiatement inférieur à n'_{IA}

n_{SB} : l'entier immédiatement supérieur à n'_{SB}

(5) Il ne nous a pas paru utile de prendre un niveau de confiance simultané de 0,95, ce qui nous aurait conduit à adopter pour chacun des niveaux de confiance, supposés égaux : $(1-\alpha)(1-\alpha) = 0,95$.

$1-\alpha = 0,9747 \neq 0,975$

Le niveau de confiance 0,95 pour chaque assertion prise séparément nous a paru suffisant.

(6) Encore ici il ne nous a pas paru utile d'être stricts au point d'exiger une probabilité de 5 % de dépassement simultané des limites.

L'interprétation des nombres n_{IA} et n_{SB} est la suivante :

Il y a une probabilité de 5 % pour que le nombre de larves porteuses de soie suturale interne à deux branches dans un échantillon de taille n ,

- soit inférieur à n_{IA} dans un échantillon provenant de l'espèce A ,
- soit supérieur à n_{SB} dans un échantillon provenant de l'espèce B ,

ou encore :

Il y a une probabilité de 95 % pour que le nombre de larves porteuses de soie suturale interne à deux branches dans un échantillon de taille n ,

- soit supérieur ou égal à n_{IA} dans un échantillon provenant de l'espèce A ,
- soit inférieur ou égal à n_{SB} dans un échantillon provenant de l'espèce B .

L'utilisation pratique que nous ferons de ce résultat est évidente : si dans un échantillon de taille n nous observons un nombre de larves à deux branches supérieur ou égal à n_{IA} nous le classerons comme provenant de A car la probabilité pour que n_{IA} soit égal ou dépassé vaut 95 % dans le cas d'échantillons extraits de cette espèce.

Le même raisonnement tient pour B .

Les nombres n_{IA} et n_{SB} donnés par les formules ci-dessus ont été calculés et les probabilités exactes recherchées directement pour connaître la validité de l'approximation (donc rectifier la limite si besoin) et posséder les probabilités de mauvaise classification lorsqu'elles dépassent 5 % (pour des tailles d'échantillons inférieures ou égales à 21).

Le tableau 5 fournit les nombres limites pour des tailles allant de 14 à 50. Les probabilités de mauvaise classification non signalées sont inférieures ou égales à 5 %. Lorsqu'elles dépassent ce seuil, elles sont signalées dans la colonne « observations » :

P_{MA} : probabilité de mal classer un échantillon provenant de A ,

P_{MB} : Probabilité de mal classer un échantillon provenant de B .

La signification de ces probabilités est la suivante : la probabilité pour que dans un échantillon de taille n provenant de A , le nombre de larves porteuses de soie à deux branches soit inférieur ou égal à n_{SB} est P_{MA} .

Il conviendra évidemment dans la pratique d'essayer d'obtenir des échantillons de taille au moins égale à 30. Les limites pour des tailles inférieures sont fournies ici surtout à titre indicatif.

La façon de voir que nous venons de développer pour la fixation des limites n_{IA} et n_{SB} relève du simple point de vue de l'échantillonnage dans deux populations : nous avons en définitive fixé des limites aux valeurs qui seront échantillonnées ultérieurement en établissant en quelque sorte un intervalle de pari pour les observations à venir. Ce que nous avons fait a le sens suivant : si une observation arrive dans l'intervalle de pari pour A , nous dirons qu'elle provient de A , si elle arrive dans l'intervalle de pari pour B , nous dirons qu'elle provient de B .

En fait, d'autres points de vue peuvent être adoptés.

Il est clair, en particulier, que la situation dans laquelle nous nous trouvons correspond à l'application du théorème de BAYES : nous capturons des femelles dans un ensemble composé de deux populations, espèce A et espèce B avec des probabilités *a priori* de tirage Π_A et Π_B ($= 1 - \Pi_A$). Ces populations sont caractérisées par les proportions p_A et p_B de larves porteuses de soie à deux branches. Le problème est de préciser après examen, donc connaissance du nombre d'individus porteurs de soie à deux branches dans un échantillon de taille n , la probabilité *a posteriori* pour que l'espèce soit A ou B .

La probabilité pour que l'espèce soit A sachant qu'on a observé x exemplaires porteurs de soie à deux branches sur n s'écrit :

$$\text{Pr} \{ A / xsn \} = \frac{1}{1 + \text{SR} \cdot R_p^x \cdot R_q^{n-x}}$$

en appelant SR le " species ratio " : Π_B / Π_A , R_p et R_q les rapports p_B / p_A et q_B / q_A .

La probabilité $Pr \{ B/xsn \}$ est bien sûr donnée par : $Pr \{ B/xsn \} = 1 - Pr \{ A/xsn \}$. Nous pouvons dans chaque cas échantillonné calculer après observation $Pr \{ A/xsn \}$ et $Pr \{ B/xsn \}$ et affecter l'échantillon soit à *A*, soit à *B*, d'après les valeurs des probabilités ci-dessus, la probabilité de mauvaise classification étant alors le complément à 1 de la probabilité attachée à l'espèce qui a été choisie, c'est-à-dire l'espèce la plus probable.

En fait, cette façon de faire se heurte à l'inconnnaissance des probabilités Π_A et Π_B que nous avons au stade actuel de déroulement du travail : les deux probabilités $Pr \{ A/xsn \}$ et $Pr \{ B/xsn \}$ ne sont calculables que moyennant des hypothèses supplémentaires sur la valeur du rapport SR.

Nous pourrions admettre que $SR = 1$, mais ceci est arbitraire et probablement non représentatif de la réalité dans de nombreux cas. Nous pourrions aussi songer à

TABLEAU n° 5. — Valeurs critiques du nombre de larves porteuses de soie suturale interne à deux branches.

Taille de l'échantillon <i>n</i>	Nombre de larves à deux branches			Observations
	<i>o</i>		<i>n</i>	
	n_{SR}		n_{TA}	
	Conclusion			
	Espèce B		Espèce A	
14	3		4	PMA = 0,10
15	3		4	PMA = 0,06
16	3		4	PMA = 0,11 PMB = 0,08 PMB = 0,11
17	4		5	PMA = 0,11
18	4		5	PMA = 0,08
19	4		5	
20	4		5	PMB = 0,06
21	5		6	PMA = 0,07
22, 23, 24	5		6	
25, 26, 27	6		7	
28, 29, 30	6		8	
31	6		9	
32	7		9	
33, 34, 35	7		10	
36, 37, 38	7		11	
39, 40	8		12	
41, 42, 43	8		13	
44	8		14	
45, 46	9		14	
47, 48, 49	9		15	
50	9		16	

Au-delà de 50 on prendra :

$n_{SR} =$ entier immédiatement supérieur à $0,1135 n - 0,5 + 0,52176 \sqrt{n}$

$n_{TA} =$ entier immédiatement inférieur à $0,4311 n + 0,5 - 0,81461 \sqrt{n}$

Les probabilités ne sont données qu'à deux chiffres après la virgule car elles ont été obtenues graphiquement à partir de l'abaque de la fonction B incomplète des tables de PEARSON et HARTLEY (1958), p. 156.

considérer un intervalle de valeurs possibles pour SR, par exemple $1/19 \leq SR \leq 19$, par simple analogie avec le risque 5 % accepté pour tous les tests, et en déduire des limites par utilisation de la règle suivante : un échantillon sera déclaré appartenir à une espèce donnée (A ou B) lorsque la probabilité *a posteriori* correspondante est supérieure ou égale à 95 % dans les conditions les plus défavorables pour cette espèce, c'est-à-dire lorsqu'elle est en nette minorité ($SR = 19$ ou $SR = 1/19$ selon le cas).

Le calcul montre que dans ce cas, la probabilité d'une affectation erronée est inférieure à 1 %, mais que le domaine de non-décision n'est suffisamment peu peuplé (moins de 5 % des échantillons de A ou de B) qu'au-delà de $n = 50$: c'est une taille d'échantillon déjà bien grande pour la pratique.

Un autre critère de fixation des limites est la considération du rapport des probabilités d'arrivée de x larves porteuses de soie à deux branches sur n sous les deux hypothèses : l'échantillon provient de A et l'échantillon provient de B.

Soit R ce rapport : $R = \Pr \{ xsn/A \} / \Pr \{ xsn/B \}$; nous pourrions prendre comme règle d'affectation :

- $R > 1$: espèce A.
- $R < 1$: espèce B.

Il semble toutefois plus sage de laisser un intervalle de non-décision entre les deux bornes de façon à ce que le risque de mauvaise affectation soit limité : cet intervalle sera défini par $L_{SB} \leq R \leq L_{IA}$ et la règle de classification deviendra alors :

- pour A : $R > L_{IA}$.
- pour B : $R < L_{SB}$.

En fait, il est aisé de montrer que la règle ci-dessus n'est pas indépendante de la valeur du « species ratio » SR :

$R > L_{IA}$ est équivalent à $\Pr \{ A/xsn \} > \frac{L_{IA}}{SR + L_{IA}}$, et

$R < L_{SB}$ est équivalent à $\Pr \{ B/xsn \} > \frac{SR}{SR + L_{SB}}$ (7).

Nous rencontrerons donc dans l'utilisation de R les difficultés déjà rencontrées dans l'utilisation du théorème de BAYES : inconnnaissance de SR et domaine de non-décision trop peuplé pour les tailles courantes d'échantillon.

Nous nous en sommes donc tenus à la règle développée ci-dessus et qui peut s'écrire avec la notation utilisée ici :

- pour l'affectation à l'espèce A :

n_{IA} défini par :

$$\Pr \{ n_{IA}sn/A \} + \Pr \{ (n_{IA} + 1)sn/A \} + \dots + \Pr \{ nsn/A \} \geq 0.95$$

- pour l'affectation à l'espèce B :

n_{SB} défini par :

$$\Pr \{ 0sn/B \} + \Pr \{ 1sn/B \} + \dots + \Pr \{ n_{SB}sn/B \} \geq 0.95$$

De toute façon, lorsque nous aurons accumulé un nombre suffisamment élevé de déterminations, nous pourrons tenter une critique expérimentale de la règle choisie et voir s'il convient de la modifier ou d'en choisir une autre.

(7) Nous pouvons constater en particulier que $R > 19$ est équivalent à $\Pr \{ A/xsn \} > 19/(SR + 19)$, ou encore $\Pr \{ A/xsn \} > 19/20$ avec $SR = 1$. Si nous voulons comme ci-dessus nous fixer $\Pr \{ A/xsn \} > 0.95$ dans les conditions les plus défavorables pour A ($SR = 19$), nous devons prendre $L_{IA} = 19^2$.

Le problème se pose de la même façon pour L_{SB} .

3.3.3. — AUTRES CARACTÉRISTIQUES

Un fait paraît frappant sur les observations recueillies pour cette étude : les larves de l'espèce *B* n'ont jamais montré de soie à trois branches. On pourrait donc être tenté d'utiliser cette absence comme critère de discrimination. En réalité il est sans beaucoup de valeur car le fait que nous n'ayons pas observé de porteuses de soie à trois branches dans les échantillons *B* ne signifie pas que de telles larves sont totalement absentes : une proportion $\pi'_B = 0,00932$ peut encore, avec une probabilité de 5 %, nous donner zéro porteur sur 320 individus observés. De même pour *A*, une proportion $\pi'_A = 0,01722$ peut, toujours avec une probabilité de 5 %, laisser apparaître huit ou plus de huit individus porteurs de soie à trois branches dans un échantillon de 233 individus (8).

Il n'est évidemment pas question de se servir ici de la proportion observée de larves porteuses de soie à trois branches : les proportions théoriques π'_A et π'_B sont trop faibles. Il faut envisager une autre variable aléatoire. Par exemple, du fait que la probabilité π'_A est à peu près le double de la probabilité π'_B , il y a plus de chances de voir apparaître tôt un porteur de soie à trois branches dans *A* que dans *B*.

Nous allons donc nous fixer à 0,95 la probabilité de voir apparaître dans un échantillon issu de *A* au moins une larve porteuse de soie à trois branches. Ceci étant fait, nous allons calculer la taille d'échantillon nécessaire et voir quelle est la probabilité pour qu'avec cette taille un échantillon issu de *B* ne présente aucune larve porteuse de soie à trois branches.

Nous arrivons ainsi à $n = 173$ pour qu'un échantillon *A* présente au moins une porteuse de soie à trois branches. C'est une taille impraticable dans les conditions ordinaires d'observation. De plus, la probabilité pour qu'un échantillon *B* de cette taille n'ait alors présenté aucune porteuse de soie à trois branches n'est que de 0,20, bien inférieure à 0,95. Cette variable est donc inutilisable. Nous avons songé également à nous servir du rang d'apparition de la première porteuse de soie à trois branches. Les rangs moyens sont : pour *A* 57,07 et pour *B* 106,30. Les calculs n'ont pas été poussés plus avant étant donné l'importance de ces valeurs moyennes : il aurait fallu des tailles d'échantillons non réalistes pour arriver à la quasi-certitude que nous nous sommes fixée : 0,95.

3.3.4. — RÉSUMÉ ET CONCLUSIONS RELATIVES A L'UTILISATION DE LA SOIE SUTURALE INTERNE

Nous avons vu que, faute de modèle de distribution, il n'a pas été possible de fixer de limites au nombre moyen de branches de cette soie. La proportion de larves porteuses de soie à deux branches dans un échantillon de larves s'est au contraire révélée intéressante : les quatre proportions sont homogènes et le modèle binomial (et son approximation gaussienne) peut être utilisé. La proportion vraie pour *A* est au moins égale à 0,4311 ; pour *B*, elle est au plus égale à 0,1135. Chacune de ces limites est fixée au niveau de confiance 0,95.

Nous avons tiré de ces proportions des nombres critiques ayant une probabilité 5 % d'être dépassés, en moins pour *A*, en plus pour *B*. Ces nombres n_{IA} et n_{SB} sont donnés dans le tableau 5.

Nous avons ensuite considéré d'autres critères de fixation des limites tels que l'utilisation du théorème de BAYES ou l'utilisation du rapport des probabilités d'arrivée du résultat observé sous les deux hypothèses origine *A* et origine *B*. Ils ont été écartés par suite de l'inconnaitance des probabilités *a priori* d'obtention d'une femelle de l'espèce *A* ou de l'espèce *B* ou du peuplement important de l'intervalle de non-décision avec les tailles courantes d'échantillons.

(8) Il est justifié de prendre la somme des quatre échantillons *A*, car ces derniers ne sont pas avérés différents pour le nombre de larves porteuses de soie à trois branches (test exact de FISHER pour les comparaisons d'échantillons à effectifs très faibles : $P = 0,11$ pour les deux échantillons extrêmes).

Enfin, nous avons vu que le fait qu'aucun individu porteur de soie à trois branches ne soit apparu dans *B* ne peut être utilisé : les probabilités correspondantes dans les populations sont trop proches l'une de l'autre pour que ce critère nous conduise à une conclusion suffisamment sûre avec des échantillons de taille praticable.

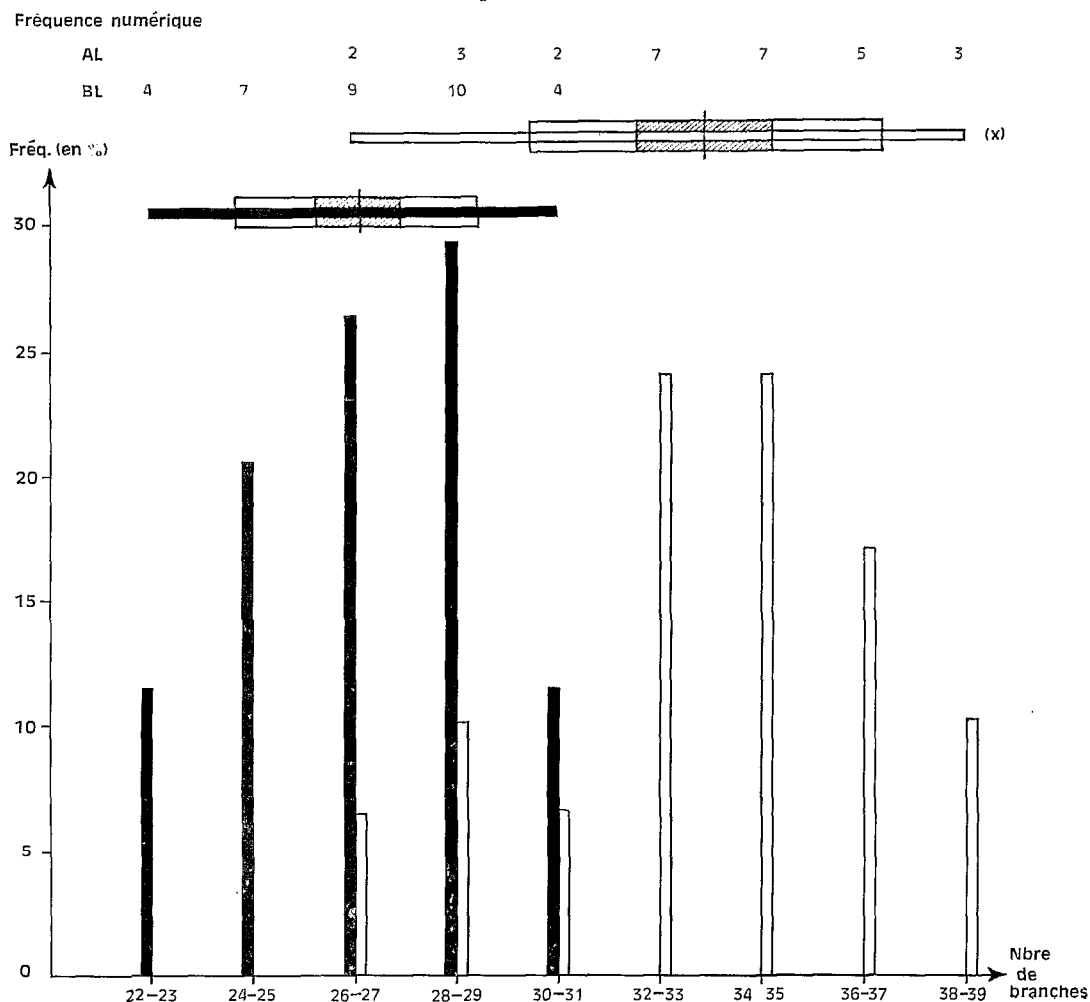
3.4. Utilisation de la soie mésothoracique n° 1.

3.4.1. — NOMBRE DE BRANCHES. MODÈLE DE DISTRIBUTION ET HOMOGENÉITÉ

Les figures 1, 2, 3 et 4 représentent les huit distributions observées.

La validité du modèle gaussien a été examinée pour chacune des quatre distributions *A* et chacune des quatre distributions *B*. Ce modèle peut être accepté pour toutes les distributions. Même celle qui s'en écarte le plus, celle de BTE, montre un

FIGURE 1
Distribution de fréquence du nombre de branches
de la soie mésothoracique n° 1 des échantillons AL et BL



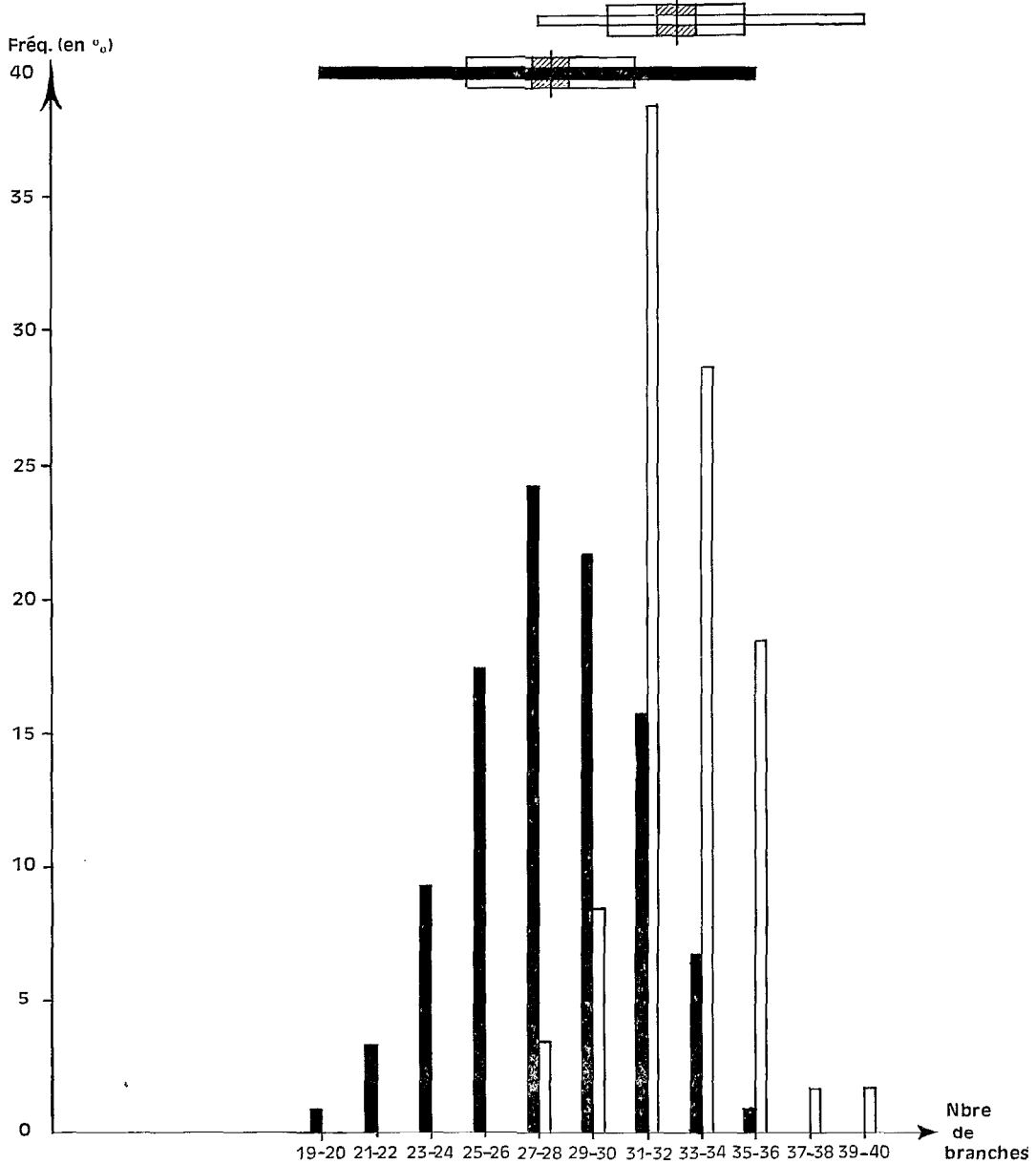
* Pour cette figure et les suivantes (2, 3 et 4)

- le trait vertical représente la moyenne arithmétique,
- le rectangle hachuré son intervalle de confiance au niveau de confiance 95 %,
- le rectangle blanc l'écart-type de l'échantillon autour de la moyenne ($\pm s$),
- la ligne épaisse l'étendue de la distribution,

FIGURE 2
 Distribution de fréquence du nombre de branches
 de la soie mésothoracique n° 1 des échantillons ATI et BTI

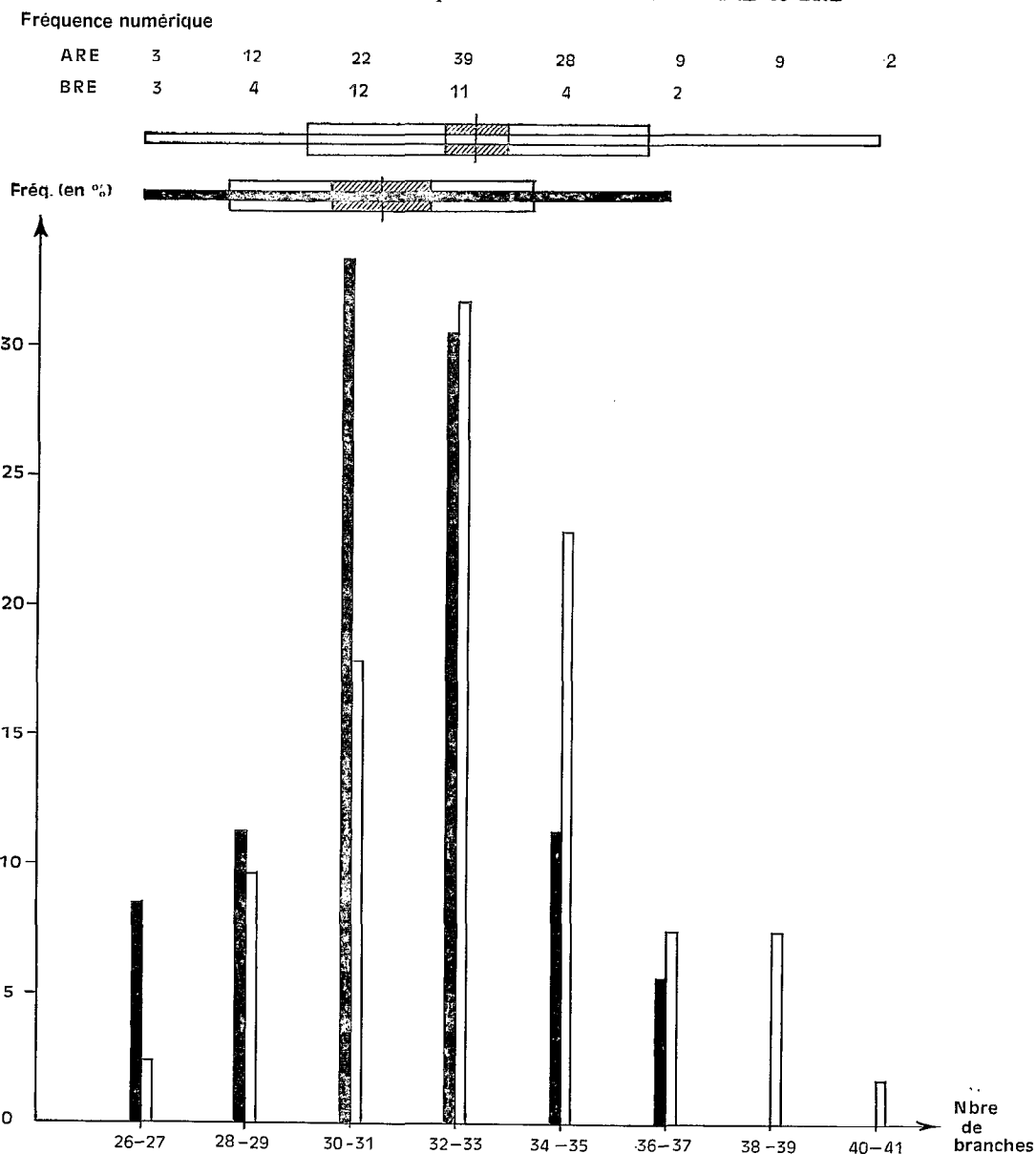
Fréquence numérique

ATI					2	5	23	17	11	1	1
BTI	1	4	11	21	29	26	19	8	1		



écart maximal entre la fonction de répartition théorique et les fréquences cumulées égal à 0,105, inférieur à 0,127 valeur critique au seuil 5 % (test de KOLMOGOROV-SMIRNOV). La différence est non significative pour une distribution prise isolément.

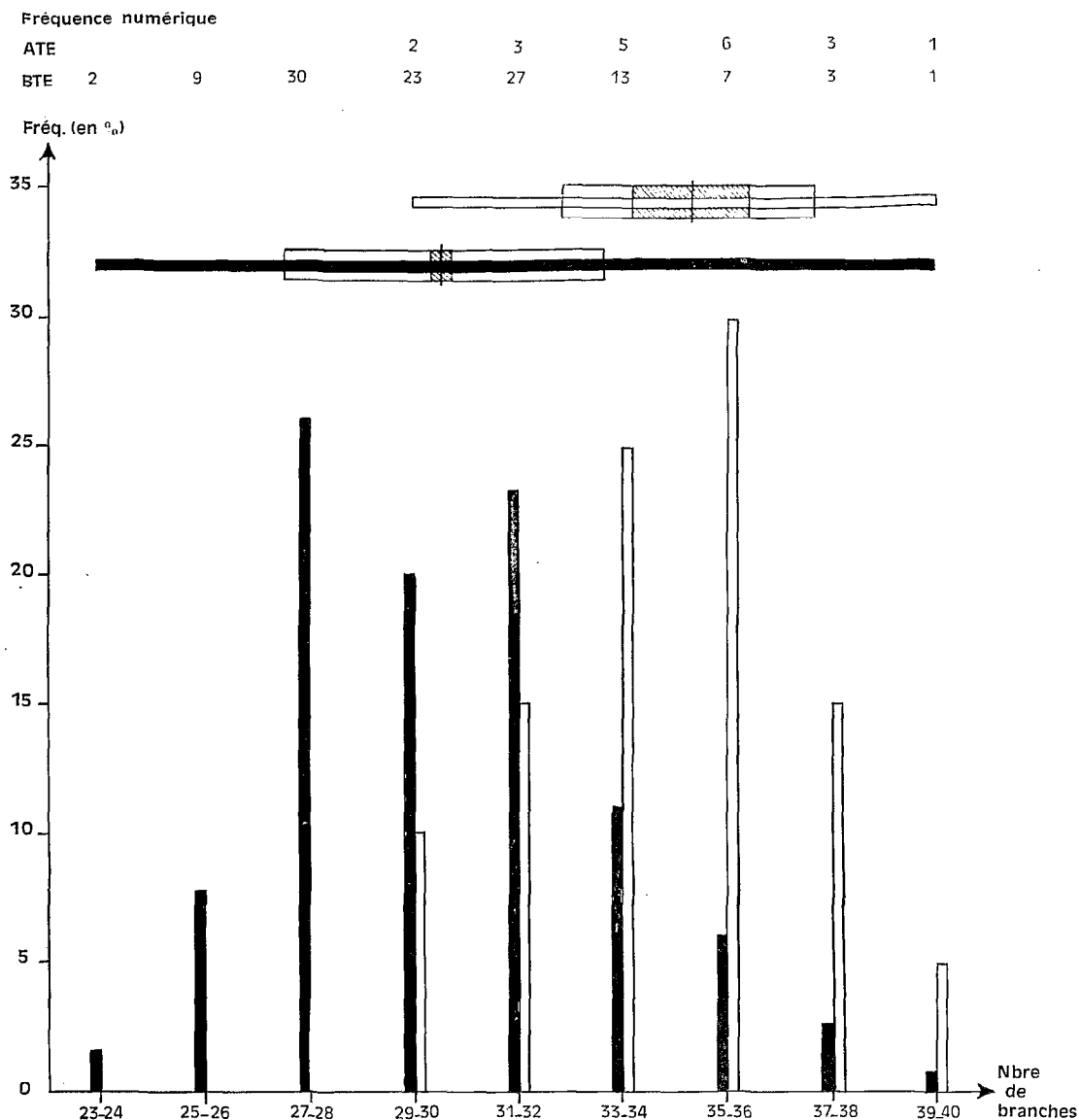
FIGURE 3
Distribution de fréquence du nombre de branches
de la soie mésothoracique n° 1 des échantillons ARE et BRE



donc *a fortiori* pour une distribution prise dans l'ensemble des distributions. Le nombre de branches de la soie mésothoracique n° 1 peut donc être considéré comme distribué de façon gaussienne. Ceci n'est évidemment pas exact en toute rigueur puisque le nombre de branches est discontinu, mais les valeurs moyennes sont suffisamment élevées pour que le modèle puisse être adopté.

Il reste maintenant le problème de l'homogénéité des échantillons : puisque toutes les distributions peuvent être admises gaussiennes et que les deux seuls paramètres moyenne et variance définissent complètement une distribution de GAUSS, notre

FIGURE 4
Distribution de fréquence du nombre de branches
de la soie mésothoracique n° 1 des échantillons ATE et BTE



problème se ramène à celui de l'homogénéité des quatre moyennes et des quatre variances.

Ces caractéristiques, ainsi que les effectifs, sont notés dans le tableau 3.

L'homogénéité des variances a été examinée au moyen du test de BARTLETT (test global de comparaison de variances).

Les résultats sont les suivants :

- pour A : $\chi^2 = 7,94$ à 3 ddl, juste significatif (χ^2 au seuil 5 % : 7,81),
- pour B : $\chi^2 = 3,71$ à 3 ddl, non significatif.

Pour A , nous nous trouvons pratiquement au seuil. En définitive, nous refusons l'homogénéité car il n'y a aucune raison *a priori* pour que la variabilité soit restée constante dans les conditions très différentes où l'espèce A a été placée.

Pour B , il n'y a aucun problème : la valeur observée pour χ^2 nous conduit à admettre l'homogénéité des variances.

Il reste l'homogénéité des moyennes : pour B , l'analyse de variance à une dimension peut être utilisée, toutes les hypothèses à la base de la méthode sont vérifiées. Cette analyse a été faite : elle conduit au rejet de l'hypothèse d'homogénéité (F calculé = 21,9 à $\nu_1 = 3$ et $\nu_2 = 302$ ddl, significatif). Les moyennes de B diffèrent donc.

Pour A , nous ne pouvons pas utiliser l'analyse de variance classique puisque nous venons de refuser l'égalité des variances, d'autant plus que les effectifs sont très inégaux.

La seule possibilité qui nous reste pour comparer les tendances centrales des distributions est la comparaison des médianes. En fait, nous parlons de médiane puisque nous avons admis le modèle gaussien, mais dans la réalité il ne nous sera pas possible de définir une médiane observée : la variable est discontinue et on ne trouvera aucun nombre de branches tel qu'il y ait la moitié de l'effectif total au-dessus et la moitié en dessous d'elle. Si elle existait (c'est-à-dire sur l'échelle continue), la médiane des échantillons groupés serait comprise entre 32 et 33 branches. Nous allons donc comparer, pour les quatre échantillons, la proportion d'individus ayant au plus 32 branches. Les données sont résumées dans le tableau suivant :

		AL	ATI	ARE	ATE
Nombre de branches ..	≤ 32	10	29	65	5
	≥ 33	19	31	59	15
Total		29	60	124	20

$\chi^2 = 7,21$ à 3 ddl, bien proche de sa valeur critique à 5 % : 7,81. Nous arrivons donc à la situation résumée ci-dessous :

		A	B
Homogénéité des	variabilités	non	oui
	tendances centrales	oui ?	non

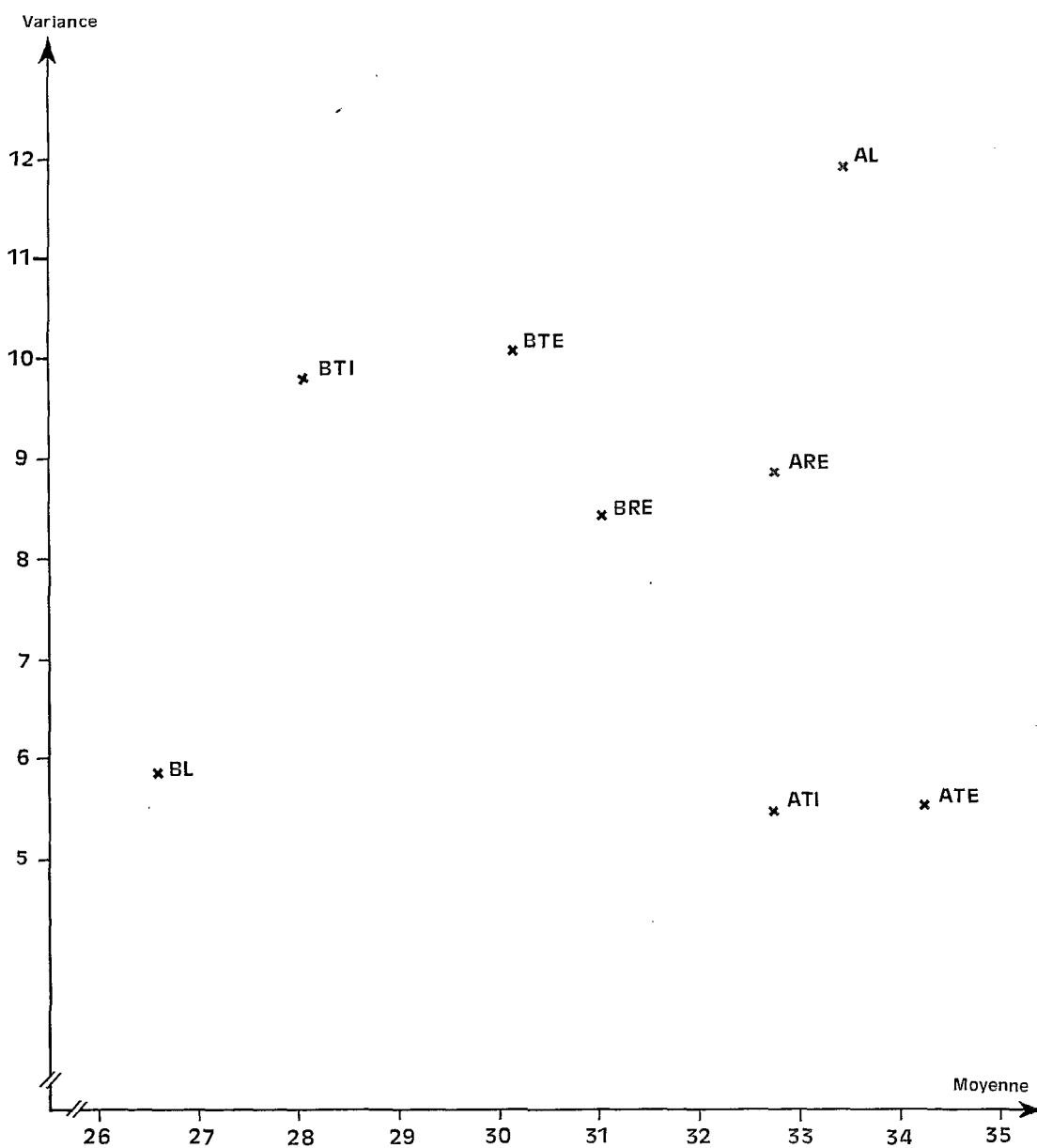
Devant une telle situation, nous sommes obligés de renoncer au premier principe que nous avons énoncé en 3.3.1., à savoir que les distributions théoriques ne doivent pas différer. Notre but en posant ce principe était l'utilisation de l'information la plus grande possible, en l'occurrence de celle fournie par la réunion des quatre échantillons. Une telle façon d'opérer n'est plus possible ici, nous devons l'abandonner.

En définitive, ceci n'est pas tellement surprenant : nous avons élevé les échantillons dans des conditions très différentes pour obtenir des résultats variés. Nous

les avons obtenus, il ne nous reste plus qu'à utiliser les extrêmes les plus proches. c'est-à-dire pour *A* aussi bien que pour *B* l'échantillon RE dont les caractéristiques sont, rappelons-le :

	<i>A</i>	<i>B</i>
moyenne	32,73	31,05
variance	8,880	8,441
effectif	124	37

FIGURE 5
Variance en fonction de la moyenne pour les huit échantillons *A* et *B*



Toutefois avant d'entamer l'exploitation proprement dite de ces résultats, il convient d'examiner un point : celui de la possibilité d'une liaison moyenne-variance qui nous permettrait éventuellement par une transformation de stabiliser les variances de A . Un simple examen de la figure 5 où sont portées en abscisse les moyennes et en ordonnée les variances pour chacun des huit échantillons nous montre l'indépendance de ces deux caractéristiques. Il est donc vain de chercher une transformation et nous utiliserons les deux résultats A et B donnés ci-dessus : ARE et BRE.

3.4.2. — LIMITES DE TOLÉRANCE POUR DES MOYENNES D'ÉCHANTILLONS DE n SUJETS ISSUS DE A OU DE B .

Comme nous l'avons fait pour A et pour B dans le cas de la soie suturale interne, nous pouvons chercher des limites de confiance pour les moyennes : limite inférieure pour A et limite supérieure pour B . Ces limites nous sont fournies par les formules classiques :

$$\bar{x} - t_v s_{\bar{x}} \text{ pour } A \text{ et } \bar{x} + t_v s_{\bar{x}} \text{ pour } B$$

soit

$$32,734 - 0,440 = 32,294 \text{ pour } A$$

$$31,054 + 0,786 = 31,840 \text{ pour } B$$

Nous pourrions, toujours comme nous l'avons fait pour la soie suturale interne, tenter de trouver les limites pour des moyennes d'échantillons de taille n .

En fait, le problème n'est pas aussi simple dans le cas présent : deux paramètres, moyenne et variance, et non plus un seul, proportion, caractérisent la population.

Nous possédons une limite pour la moyenne, nous pourrions, bien sûr, chercher une limite supérieure pour la variance et opérer comme ci-dessus. Nous ne l'avons pas fait car le problème de la fixation de limites lorsque ni μ ni σ^2 ne sont connus a été résolu dans le cas d'une loi de Gauss : les limites ainsi trouvées s'appellent les limites de tolérance. Elles s'obtiennent à partir de la moyenne et de l'écart type observé de la même manière que les limites de confiance :

$$\bar{x} + Ks \text{ et } \bar{x} - Ks$$

Le coefficient K est évidemment différent de t_v utilisé pour les limites de confiance.

Avant de fixer ces limites dans le cas qui nous intéresse, voyons d'abord ce qu'elles sont, c'est-à-dire le sens qu'on doit leur attacher.

a) Limites de tolérance pour une mesure isolée

Supposons que nous prélevions un échantillon dans une loi de Gauss de moyenne μ et de variance σ^2 inconnues ; cet échantillon nous fournit deux estimations \bar{x} et s^2 de ces deux paramètres. Nous nous demandons quelle est la valeur x_T de x telle qu'il y ait 95 % (p. ex.) de la population en dessous de cette valeur. Le problème pour une limite inférieure (95 % au-dessus) est symétrique. Manifestement, nous ne pouvons répondre de façon certaine à une telle question car nous ne connaissons ni μ ni σ^2 mais deux estimations de ces quantités qui sont des variables aléatoires. Nous allons donc devoir répondre en probabilité, associant un risque à notre affirmation ou, ce qui est plus courant, affirmant avec une quasi-certitude égale au complément à 1 de notre risque. Dans notre exemple, nous dirons : avec une quasi-certitude de 0,95 (p. ex.) x_T est égal à telle valeur. Autrement dit, nous venons de nous fixer, à partir des caractéristiques observées sur un échantillon, une limite dont on peut dire, avec une probabilité suffisamment grande, qu'elle laisse en dessous d'elle une proportion suffisamment importante de la population.

b) Limites de tolérance pour les moyennes d'échantillons

Le problème est exactement le même : nous considérons ici non plus la population des valeurs isolées mais la population des moyennes d'échantillons de taille n . La solution est la même que précédemment à cette différence près que là où nous

trouverons s nous le remplacerons par s/\sqrt{n} et les limites données ci-dessus seront remplacées par :

$$\bar{x} - K \frac{s}{\sqrt{n}} \text{ et } \bar{x} + K \frac{s}{\sqrt{n}}$$

Il convient de bien faire la différence entre limites de confiance et limites de tolérance : les limites de confiance s'adressent à un paramètre, moyenne par exemple, de la population, alors que les limites de tolérance s'adressent à des moyennes qui seront observées sur des échantillons à venir. On peut dire qu'il s'agit de limites de pari sur des échantillons extraits d'une population qu'on ne connaît pas de façon exhaustive, mais seulement par échantillonnage.

Si on connaît, pour une population de Gauss, les deux paramètres μ et σ^2 , on peut affirmer que la moyenne d'un échantillon de n individus sera inférieure à $\mu + 1,645 \frac{\sigma}{\sqrt{n}}$ avec une probabilité de 0,95.

A partir du moment où ni μ ni σ^2 ne sont connus, mais estimés par \bar{x} et s^2 déjà observés, on peut encore dire, mais cette fois ce n'est plus une certitude (on n'affirme plus à 1 mais à 0,95) que la moyenne d'un échantillon de taille n sera inférieure à $\bar{x} + K \frac{s}{\sqrt{n}}$ avec une probabilité au moins égale à 0,95.

c) *Etablissement de la valeur du coefficient K*

La théorie établissant la valeur de ce coefficient est assez compliquée puisqu'elle doit tenir compte du fait que \bar{x} et s^2 sont des variables aléatoires. Elle est exposée dans différents ouvrages de statistique, p. ex. HALD (1962), p. 311. Pour des limites unilatérales, celles que nous utilisons ici, les formules les plus simples sont fournies par M. G. NATRELLA (1963), p. 2-15 : si γ est le coefficient de confiance, P la proportion de la population, et N la taille de l'échantillon que l'on a utilisé, la valeur de K est donnée par les formules suivantes :

$$a = 1 - \frac{z_\gamma^2}{2(N-1)}$$

$$b = z_P^2 - \frac{z_\gamma^2}{N}$$

$$K = \frac{z_P + \sqrt{z_P^2 - ab}}{a}$$

z est la variable de GAUSS centrée réduite et z_α a le sens suivant :

$$\text{Pr} \{ z \leq z_\alpha \} = \alpha$$

Appliquant ici ces formules, nous trouvons ($\gamma = 0,95$; $P = 0,95$) :

- pour A
 $K_A = 1,892$
- pour B
 $K_B = 2,140$

Les limites de tolérance pour des moyennes d'échantillons de taille n sont donc :

- limite inférieure pour A : $I_A = 32,734 - \frac{5,638}{\sqrt{n}}$
- limite supérieure pour B : $S_B = 31,054 + \frac{6,218}{\sqrt{n}}$

Un effectif intéressant à connaître est celui pour lequel $I_A = S_B$. En effet au-dessus nous aurons $I_A > S_B$ et nous pourrions conclure, en dessous nous aurons $I_A < S_B$ et nous ne pourrions pas conclure.

Nous écrivons que $I_A = S_B$:

$$32,734 - \frac{5,638}{\sqrt{n}} = 31,054 + \frac{6,218}{\sqrt{n}}$$

et nous en tirons

$$n = 49,80, \text{ soit arrondi : } 50.$$

En dessous de $n = 50$, nous ne pourrions pas conclure si nous nous en tenons aux risques de 5 % que nous nous sommes fixés.

TABLEAU n° 6. — Limites de tolérance, inférieure pour A : I_A , supérieure pour B : S_B , pour quelques tailles d'échantillons : n

n	S_B	I_A
10	33,02	30,95
11	32,93	31,03
12	32,85	31,11
13	32,78	31,17
14	32,72	31,23
15	32,66	31,28
20	32,44	31,47
25	32,30	31,61
30	32,19	31,70
35	32,10	31,78
40	32,04	31,84
45	31,98	31,89
50	31,93	31,94
60	31,86	32,01
70	31,80	32,06
80	31,75	32,10
90	31,71	32,14
100	31,68	32,17
110	31,65	32,20
120	31,62	32,22
130	31,60	32,24
140	31,58	32,26
150	31,56	32,27
200	31,49	32,33

Le tableau 6 fournit les valeurs de I_A et S_B pour quelques tailles d'échantillons. Si des points supplémentaires sont désirés, ils peuvent aisément être calculés à partir des formules données ci-dessus.

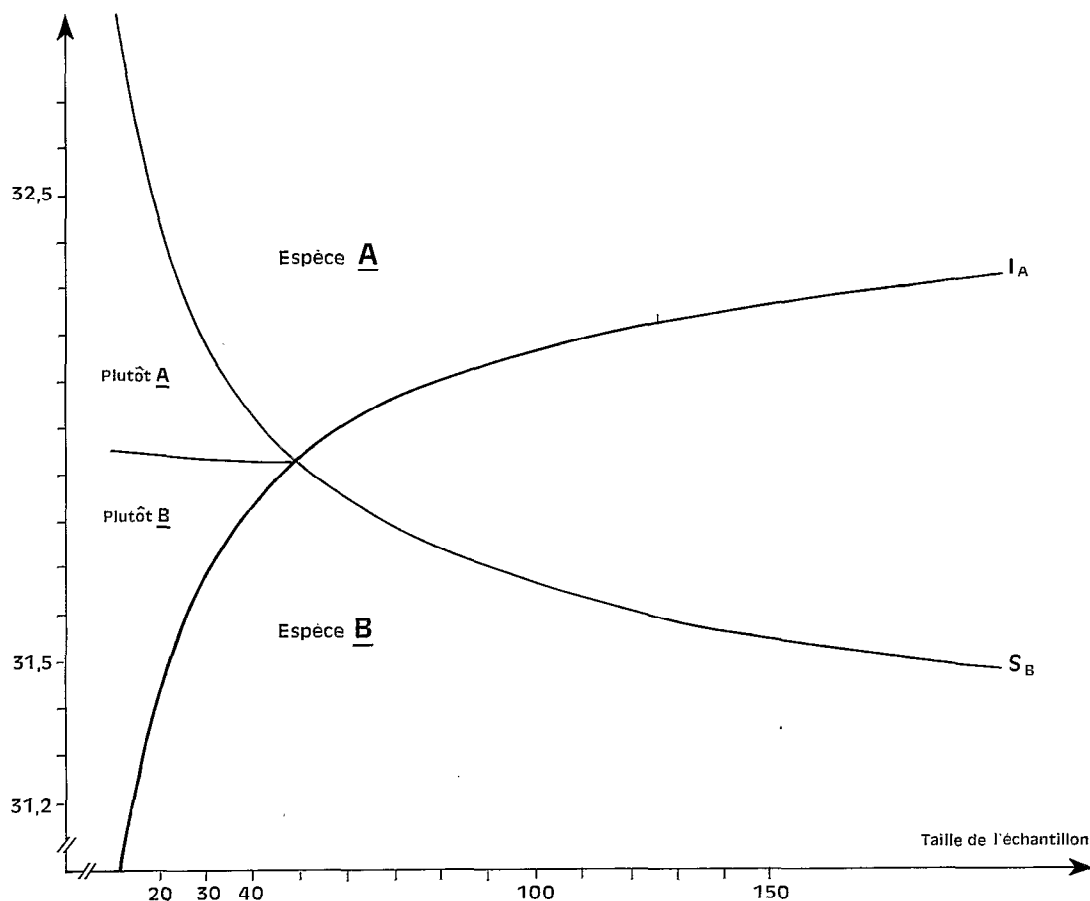
Le graphique de la figure 6 traduit ce tableau. Pour les besoins pratiques, il sera fait sur papier millimétré à partir du tableau 6.

Son utilisation (pour n supérieur ou égal à 50) est la suivante :

Pour un échantillon de taille n , nous déterminerons le nombre moyen de branches de la soie mésothoracique n° 1. Si pour cette taille le point représentatif est en dessous de S_B nous conclurons qu'il s'agit de l'espèce B, s'il est au-dessus de I_A nous conclurons qu'il s'agit de l'espèce A. En principe, le point « observation » ne

tombera qu'exceptionnellement entre I_A et S_B puisque ces limites n'ont chacune qu'une chance sur 20 d'être dépassées (cette affirmation étant, comme nous l'avons vu, faite à 0,95). La raison pour laquelle nous portons ces conclusions a été exposée à propos de la soie suturale interne.

FIGURE 6
Limites I_A et S_B en fonction de n



La situation est donc claire pour des effectifs supérieurs à 50.

Pour les tailles d'échantillons inférieures à 50, la majorité des cas pratiques. S_B est supérieur à I_A et il apparaît sur le graphique un domaine où la conclusion est « A » ou « B ». Si nous nous en tenons aux risques 5 %, il faudra renoncer à conclure dans les cas où le point observé arrivera dans ce domaine. Mais si nous sommes moins stricts sur les risques, acceptant des probabilités plus grandes que 5 %, nous pourrions encore conclure : c'est pour cette raison que ce domaine a été divisé en deux par une courbe : partie supérieure « plutôt A », partie inférieure « plutôt B », obtenue de la façon suivante :

Ne nous fixant plus $P = 0,95$, nous avons K_A et K_B en fonction de z_P :

$$K_A = 1,0111z_P + \sqrt{0,011245z_P^2 + 0,022063}$$

$$K_B = 1,03905z_P + \sqrt{0,040571z_P^2 + 0,075982}$$

toujours à la quasi-certitude 0,95.

Nous allons chercher le lieu des points pour lesquels $I_A = S_B$ (en gardant la symétrie pour P : $P_A = P_B = P$) :

$$\text{avec } I_A = \bar{x}_A - K_A \frac{s_A}{\sqrt{n}} \quad \text{et} \quad S_B = \bar{x}_B + K_B \frac{s_B}{\sqrt{n}}$$

nous obtenons

$$\sqrt{n} = \frac{K_A s_A + K_B s_B}{\bar{x}_A - \bar{x}_B} \quad \text{et} \quad I_A = S_B = \frac{\bar{x}_B K_A s_A + \bar{x}_A K_B s_B}{K_A s_A + K_B s_B}$$

K_A et K_B étant fonction de P par l'intermédiaire de z_P .

Nous pouvons donc, en remplaçant \bar{x}_A , \bar{x}_B , s_A et s_B par leur valeur, obtenir les relations qui lient n , I_A (ou S_B) et P. Deux des variables sont fonction de la troisième. Les calculs sont numériquement un peu compliqués. Nous avons choisi de fixer P et d'en déduire l'abscisse n et l'ordonnée $I_A = S_B$. Le tableau 7 donne les résultats pour quelques valeurs de P. Nous pensons que ces quelques repères suffisent pour les besoins pratiques.

TABLEAU n° 7. — Quelques points du lieu géométrique $I_A = S_B$ et risque associé aux courbes $I_A(n)$ et $S_B(n)$ passant par ces points.

P %	risque %	n	$I_A = S_B$
95	5	49,80	31,935
93	7	40,74	31,937
90	10	31,49	31,939
85	15	21,59	31,944
80	20	15,12	31,949
75	25	10,50	31,956

Comme nous le constatons, et c'était évident *a priori*, plus les tailles sont petites plus le risque augmente. La signification de ce risque est claire : si le point « observé » tombe dans le domaine « plutôt A », nous porterons la conclusion « espèce A », mais sans perdre de vue au moment de l'interprétation des résultats qu'un échantillon issu de B peut donner un point dans ce domaine avec la probabilité qui a été appelée risque.

3.4.3. — RÉSUMÉ ET CONCLUSIONS RELATIVES A L'UTILISATION DE LA SOIE MÉSOTHORACIQUE n° 1.

Nous avons vu au cours de l'examen des résultats relatifs à cette soie que le modèle gaussien pouvait être admis. Toutefois, l'ensemble des résultats n'a pu être utilisé car les différentes distributions se sont révélées hétérogènes à un point de vue ou à un autre.

Partant des échantillons extrêmes les plus proches observés, nous avons calculé des limites de tolérance pour des moyennes d'échantillons de taille n issus de A et issus de B. Il s'agit de valeurs critiques, inférieure pour A, appelée I_A , supérieure pour B, appelée S_B , qui ne seront dépassées qu'avec une probabilité de 0,05 au plus. Ces affirmations de dépassement ne sont évidemment pas certaines puisque ni les moyennes, ni les variances théoriques ne sont connues. Elles sont faites avec une quasi-certitude de 0,95.

Les valeurs I_A et S_B ont été calculées pour différentes tailles n d'échantillons et sont données dans le tableau 6. Nous avons constaté que des conclusions respectant les risques de 5 % que nous nous sommes fixés ne sont possibles qu'avec des échantillons de taille au moins égale à 50. En dessous de $n = 50$, nous pouvons conclure « plutôt A » ou « plutôt B » avec des risques plus grands, d'autant plus grands que les tailles sont petites. Ils peuvent atteindre 25 % environ pour $n = 11$. La soie mésothoracique n° 1 sera donc dans la pratique moins intéressante que la soie suturale interne.

3.4.4. — INDÉPENDANCE DES CRITÈRES UTILISÉS.

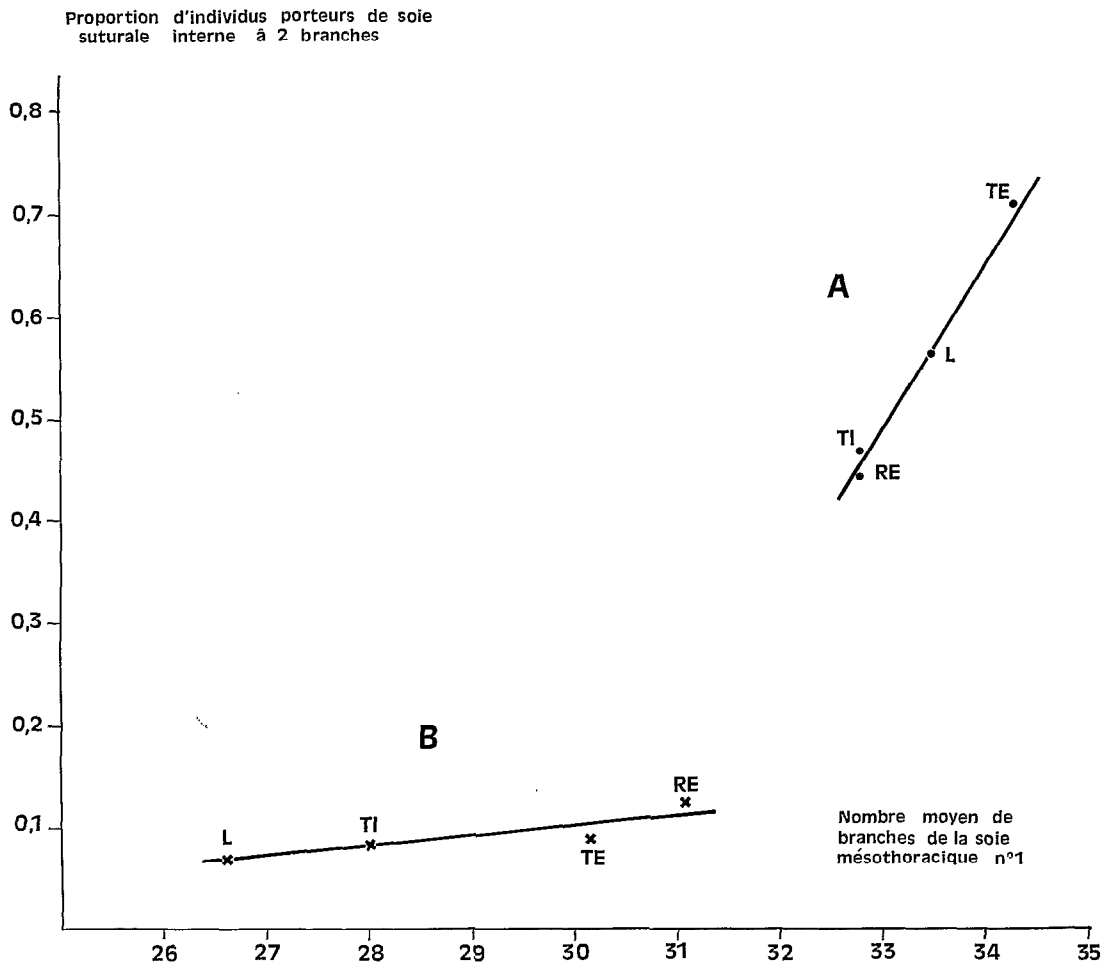
Ces critères sont : proportion d'individus porteurs de soie suturale interne à deux branches et nombre moyen de branches de la soie mésothoracique n° 1.

Il est très malaisé d'étudier directement l'indépendance de ces deux variables puisque nous ne possédons que huit échantillons. D'autre part, ces échantillons ne sont pas tous de même taille, il s'en faut de beaucoup. La question de l'indépendance se pose toutefois ne serait-ce qu'à cause du fait que *A* s'est trouvé au-dessus de *B* aussi bien pour la soie suturale interne que pour la soie mésothoracique n° 1.

La figure 7 résume les huit observations faites : les deux variables apparaissent comme très liées, presque de façon fonctionnelle, telle est la première constatation. La seconde est qu'il faut renoncer à l'utilisation de l'analyse discriminante classique sur ces données, les deux droites n'étant manifestement pas parallèles. Le problème reste donc de séparer le plan en deux régions, l'une « espèce *A* », l'autre « espèce *B* » de telle façon que les risques de mauvaise classification soient égaux et les plus faibles possible. Il resterait de plus à fixer les tailles d'échantillons de telle façon que ces risques soient acceptables.

FIGURE 7

Liaison entre le nombre moyen de branches de la soie mésothoracique n° 1 et la proportion d'individus porteurs de soie suturale interne à deux branches



Ce problème de séparation du plan en deux n'a en définitive pas été abordé ici car nous espérons être arrivés à des règles de discrimination suffisamment simples par l'utilisation des variables prises isolément.

Comme nous venons de le constater, les deux critères apparaissent à première vue fortement liés. En fait, il convient d'être prudent car il peut s'agir d'une fausse liaison, les deux variables étant influencées dans le même sens par les conditions variées dans lesquelles ont été élevés les échantillons. C'est la raison pour laquelle la liaison a été recherchée dans les échantillons eux-mêmes au moyen du test χ^2 d'indépendance, ou du test exact de FISHER lorsque les effectifs étaient trop faibles pour que le test χ^2 puisse être utilisé.

TABLEAU n° 8. — Indépendance des nombres de branches des soies suturale interne et mésothoracique n° 1.

Echantillon	Valeur du χ^2 ou probabilité (test de FISHER)		Concl.
AL	3,15	2 ddl	N. S.
BL	P = 0,52		N. S.
ATI	2,65	3 ddl	N. S.
BTI	0,45	1 ddl	N. S.
ATE	P = 0,56		N. S.
BTE	0,40	1 ddl	N. S.
ARE	17,58	8 ddl	S.
BRE	P = 0,28		N. S.

Pour la soie suturale interne, le nombre de classes utilisé a toujours été le même : « deux branches » et « différent de deux branches » ; pour la soie mésothoracique n° 1, le nombre de classes a été fixé en fonction des effectifs de telle façon que la limite de validité du test χ^2 soit respectée. Le tableau 8 donne les résultats de ces tests : valeur du χ^2 et nombre de ddl ou valeur de la probabilité dans le cas d'utilisation du test de FISHER. La conclusion est portée au seuil 5 % par test. Seul l'échantillon ARE montre un χ^2 significatif. Encore faut-il ne pas tellement en tenir compte dans la conclusion globale car il n'est pas significatif à 1 % et le seuil qu'on devrait adopter par test, pour huit tests simultanés est d'environ 0,7 % si l'on fixe le risque global à 5 %.

La conclusion générale de l'ensemble des tests est donc l'indépendance des nombres de branches des deux soies. Il s'ensuit que d'une part la liaison constatée sur la figure 7 est une fausse liaison, et d'autre part la séparation du plan, envisagée ci-dessus, en domaines « A » et « B » se fera au moyen de droites parallèles aux axes de référence, ce qui, en définitive, revient à étudier les variables séparément, ce que nous avons fait.

4. - ÉTUDE COMPARÉE DE LA CHETOTAXIE DES NYMPHES DE L'ESPÈCE A ET DE L'ESPÈCE B

Nous avons repris quatre des caractères déjà étudiés par COLUZZI (1964) parmi ceux qui semblaient les plus prometteurs.

Les soies retenues et étudiées biométriquement furent les soies n° 4 et n° 5 du segment II et la soie n° 1 du segment III et du segment IV.

La dénomination de ces soies correspond à la nomenclature de BELKIN (1952),

Chaque type de soie a été étudié individuellement, puis nous avons groupé, comme COLUZZI l'avait fait, les soies 4 et 5 du segment II et les soies 1 des segments III et IV.

Près de 800 exuvies nymphales ont été montées en préparation microscopique permanente (P.V.A.) dont 558 ont pu être utilisées.

L'étude de ces exuvies a été menée en ce qui concerne : groupes d'échantillonnages correspondants entre espèces, analyses biométriques, principe de différenciation interespèces, suivant les mêmes méthodes que celles que nous avons adoptées pour l'étude des larves.

Aucune des soies, considérées isolément ou groupées comme il a été dit ci-dessus, ne permet de distinguer une espèce de l'autre (9).

5. - CONCLUSIONS

Les conditions variées de développement imposées aux deux espèces dans nos élevages nous ont permis de noter d'amples variations chétotaxiques dans les soies étudiées. Cependant, nous ne pouvons affirmer que nous avons obtenu par cet artifice l'amplitude extrême de variation chétotaxique qui peut se rencontrer dans la nature. La répartition de *A. gambiae* A et B couvre en effet des aires étendues et diverses alors que les souches que nous avons utilisées pour l'une comme pour l'autre avaient dans chaque cas une seule origine. Seule l'application des résultats obtenus à l'ensemble de Madagascar (10) nous le confirmera ou non.

Néanmoins, comme premier résultat, il apparaît possible de différencier une souche d'espèce A d'une souche d'espèce B au moins entre les deux régions de Madagascar que nous avons considérées. Cette distinction peut se faire en examinant le nombre de branches de la soie suturale interne et de la soie mésothoracique n° 1. Notons au passage que nous n'avons pas retenu la soie prothoracique n° 1 signalée comme la plus prometteuse par COLUZZI pour séparer nos deux espèces. Cette non-concordance peut provenir de ce que notre critère de discrimination est très sévère. A partir de cette constatation, nous pouvons d'ailleurs espérer que les deux variables retenues seront suffisamment puissantes pour montrer des différences entre les deux espèces à l'échelon de Madagascar avec des tailles d'échantillons qu'il est pratiquement possible d'obtenir de façon courante. La vérification est en cours et elle semble parfaitement confirmer cet espoir.

6. - REMERCIEMENTS

Tous nos remerciements vont à la Division de l'Eradication du Paludisme de l'Organisation Mondiale de la Santé pour son aide financière, et à M. le Dr BRUCE-CHWATT, chef du Service « Recherches et Informations techniques » de cette division pour l'aide critique constante qu'il nous a apportée.

Ils vont également à M. le Dr ESTRADE, chef du service de Lutte contre les Grandes Endémies de la République malgache, pour le personnel de qualité qu'il a bien voulu mettre à notre disposition. Parmi ce personnel, MM. RASOLONIAINA L. de G. (pour la partie calcul) et RAVAONJANAHARY Ch. (pour la partie lecture) sont à mentionner pour leur sérieux et l'intérêt personnel qu'ils ont accordé à ce travail.

(9) Afin d'alléger cet exposé, nous ne donnons pas les résultats obtenus avec ces soies. Mais ils peuvent être obtenus sur simple demande, auprès du Service de Documentation des Services Scientifiques Centraux de l'O.R.S.T.O.M. à 93 - Bondy (France).

(10) Un article ultérieur (en coll. avec G. DAVIDSON), donnera les résultats de cette application sur le terrain.

Ils vont encore à G. DAVIDSON, du Ross Institute de Londres, sans lequel cette étude n'aurait pu avoir lieu puisque ce sont ses déterminations d'espèces qui sont à la base de ce travail ; ils vont enfin à M. le P^r L'HÉRITIER et à P. LAZAR pour les suggestions qu'ils nous ont faites.

Nous ne saurions oublier M. VÉRENT, technicien d'entomologie O.R.S.T.O.M, qui fut le second lecteur et a collationné les résultats, ainsi que M^{mes} ANDRÉ et SCHWARTZ qui ont réalisé la plupart des calculs numériques exécutés aux Services scientifiques centraux de l'O.R.S.T.O.M.

BIBLIOGRAPHIE

1. Entomologie

- BELKIN (N. J.), 1952. — The homology of the chaetotaxy of immature mosquitoes and a revised nomenclature for the chaetotaxy of the pupa. *Proc. Ent. Soc. Wash.*, 54, 115-130.
- BURGESS (R. W.), 1962. — Preliminary experiments on the hybridization of *Anopheles gambiae* Giles and *Anopheles melas* Theobald. *J. trop. Med. Hyg.*, 11, 702-704.
- CORONEL (L. T.), 1962. — *Morphological variations in Anopheles gambiae* Giles, W.H.O. document : W.H.O./Mal/328.
- COLUZZI (M.), 1964. — Morphological divergences in the *Anopheles gambiae* complex. *Riv. Malariol.*, 43, 197-232.
- COLUZZI (M.), 1966. — Osservazioni comparative sul cromosoma X nelle specie A e B del complesso *Anopheles gambiae*. *Accad. naz. Lincei, R.C.Sc. fis., mat. e nat.* vol. 40, ser. 8, fasc. 4, 671-678.
- DAVIDSON (G.), 1956. — Insecticide resistance in *Anopheles gambiae* Giles, a case of simple mendelian inheritance. *Nature (London)*, 178, 863-864.
- DAVIDSON (G.) et JACKSON (C. E.), 1962. — Incipient speciation in *Anopheles gambiae* Giles. *Bull. Org. Mond. Santé*, 27, 303-305.
- DAVIDSON (G.), 1963. — Speciation in *Anopheles gambiae* Giles. *Abstracts of the Seventh International Congresses on Tropical Medicine and Malaria*. Rio de Janeiro, 438-439.
- KUHLLOW (F.), 1962. — Studies on the bionomics and the morphology of the salt-water breeding *Anopheles gambiae* on the coast of Tanganyika. *Riv. Malariol.*, 41, 187.
- MASON (G. F.), 1964. — The causes of males sterility in *Anopheles gambiae* A - B group crosses. *Riv. Malariol.*, XLIII, 185-191.
- MUIRHEAD-THOMSON (R. C.), 1945. — Studies of the breeding places and control of *Anopheles gambiae* and *A. gambiae* var. *melas* in coastal districts of Sierra Leone. *Bull. Ent. Res.*, 36, 185-252.
- MUIRHEAD-THOMSON (R. C.), 1948. — Studies on *Anopheles gambiae* and *A. melas* in and around Lagos. *Bull. Ent. Res.*, 38, 527-558.
- PATERSON (H. E.), 1962. — On the status of the East African salt-water breeding variant of *Anopheles gambiae* Giles. *Nature (London)*, 195, 469-470.
- PATERSON (H. E.), 1963. — The species, species control and antimalarial spraying campaigns. Implications of recent work on the *Anopheles gambiae* complex. *S. Afr. J. med. Sci.*, 28, 33-44.
- PATERSON (H. E.), PATERSON (J. S.) et VAN EEDEN (G. S.), 1963. — A new member of the *Anopheles gambiae* complex : Preliminary report. *Medical Proc.*, 9, 414-418.
- PURI (I. M.), 1928. — The relationship of certain morphological characters of anophelines larvae to the classification of Indian anopheline mosquitoes. *Ind. J. Med. Res.* 16, 519-528.

RIBBANDS (C. R.), 1944. — Differences between *Anopheles melas* (*A. gambiae* var. *melas*) and *Anopheles gambiae*. I. The larval pecten. *Ann. trop. Med. Parasit.*, 38, 85-86.

2. Statistique

COCHRAN (W. G.) et HOPKINS (C. E.), 1961. — Some classification problems with multivariate qualitative data. *Biometrics*, 17, 10-32.

HALD (A.), 1962. — *Statistical theory with engineering applications*, New York. J. Wiley & Sons, XII + 783 p.

HILLS (M.), 1967. — Discrimination and allocation with discrete data. *Applied statistics*, 16, 237-250.

LINDER (A.), 1963. — Trennverfahren bei qualitativen Merkmalen. *Metrika*, 6, 76-83

LINHART (H.), 1959. — Techniques for discriminant analysis with discrete variables. *Metrika*, 2, 138-149.

NATRELLA (M. G.), 1963. — *Experimental Statistics*, N.B.S. Handbook, 91, 542 p.

PEARSON (E. S.) et HARTLEY (H. O.), 1958. — *Biometrika tables for statisticians*, Vol. 1 Cambridge, The University Press, XIV + 240 p.

RYAN (T. A.), 1959. — Comparisons in psychological research. *Psychol. Bull.*, 56, 26-47.