

REPUBLIQUE DU CAMEROUN
Republic of Cameroon

UNIVERSITE DE YAOUNDE I
University of Yaoundé I



FACULTE DES SCIENCES
Faculty of Science

Département de Biochimie
Department of Biochemistry

**Variabilité et structure génétiques des
populations d'*Anopheles gambiae*
(Giles, 1902), vecteur du paludisme au
Cameroun**

THESE

Présentée et soutenue le 17 décembre 2003 en vue de l'obtention du diplôme
de Doctorat de 3^{ème} cycle en biochimie

Spécialité : Génétique et biologie moléculaire

Par

WONDJI Charles Sinclair

Maître ès Sciences filière Biochimie

Matricule 97Z863

Sous la direction de

Dr Didier FONTENILLE
Directeur de Recherche
IRD Montpellier (France)

Pr. Albert SAME EKOBO
Professeur
Université de Yaoundé I

Année académique 2002-2003

ATTESTATION DE CORRECTION DE LA THESE

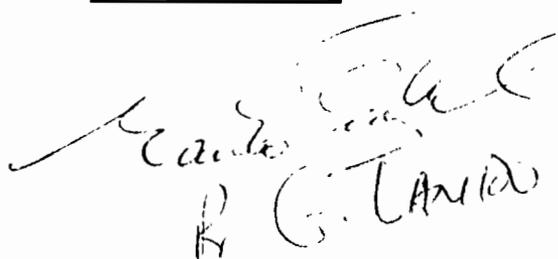
XX

Je soussigné Pr. Gabriel LANDO président du jury de thèse de Mr **WONDJI Charles Sinclair**, certifie que, la thèse présentée par WONDJI Charles Sinclair (matricule 97Z863) le 17 décembre 2003 pour l'obtention du Doctorat 3^{ème} cycle intitulé : « *Variabilité et structure génétiques des populations d'Anopheles gambiae (Giles, 1902), vecteur de Plasmodium au Cameroun* », a été corrigée selon les recommandations de tous les membres du jury.

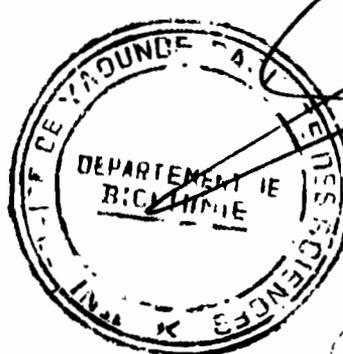
Fait à Yaoundé le 05 mars 2004, pour servir et valoir ce que de droit.

6 AVR 2004

Président du Jury



Chef du département de Biochimie

Stou François Simeon



Division de la Programmation et du Suivi des Activités Académiques

LISTE DES ENSEIGNANTS PERMANENTS AU 10 MAI 2003

ANNEE ACADEMIQUE 2002/2003

(Par Département et par Grade)

1 - DEPARTEMENT DE BIOCHIMIE (BC)			
N°	NOMS ET PRENOMS	GRADE	OBSERVATIONS
1	ETOA François Xavier	Maître de Conférences	Chef de Dépt <i>Directeur au MINESUP</i>
2	ATOGLIO Barbara Mma	Chargée de Cours	en poste
3	FOKOU Elie	Chargé de Cours	en poste
4	BENG née NINTCHOM PENLAP V	Chargée de Cours	en poste
5	BIYITI BI ESSAM née AKAM ADA L. F.	Chargée de Cours	<i>Chef de S. au MINESUP</i>
6	DEMMANO Gustave	Chargé de Cours	en poste
7	DJOKAM TAMO Rosine Rachel	Chargé de Cours	en poste
8	ESSIA NGANG Jean Justin	Chargé de Cours	en poste
9	EVEHE BEBANDOUE Marie-Solange	Chargée de Cours	en poste
10	FEKAM BOYOM Fabrice	Chargé de Cours	en poste
11	KANSCI Germain	Chargé de Cours	en poste
12	MBACHAM Wilfred	Chargé de Cours	en poste
13	MINKA Samuel	Chargée de Cours	en poste
14	MOFOR née TEUGWA Clautilde	Chargée de Cours	en poste
15	MOUNDIPA FEWOU Paul	Chargé de Cours	en poste
16	MULUH John PAYNE	Chargé de Cours	en poste
17	NGUEFACK Julienne	Chargée de Cours	en poste
18	NKENFOU NGUEFEU Céline	Chargée de Cours	en poste
19	NJAYOU MOUNDJOHOU	Chargé de Cours	<i>Absent du poste</i>
20	OBEN Julius ENYONG	Chargé de Cours	en poste
21	RIWOM SARA Honorine	Chargée de Cours	en poste
22	TCHANA NKOUATCHOUA Angèle	Chargée de Cours	en poste
23	TCHU'INMOGNE SIMO Jules	Chargé de Cours	en poste

24	WAKAM née NANA Louise	Chargée de Cours	en poste (malade)
25	BELINGA née NDOYE FOE	Assistante	en poste
26	EWOANE Célestin	Assistant	<i>Absent du poste</i>
27	SHANG Judith DZELAMBONG	Assistante	en poste
28	SOUOPGUI Jacob	Assistant	en poste

2 - DEPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE ANIMALES (B.P.A.)

29	NJIKAM NJIFUTIE	Maître de Conférences	Chef de Dépt
30	NJINE Thomas	Professeur	en poste
31	BILONG BILONG Charles Félix	Maître de Conférences	en poste
32	FOMENA Abraham	Maître de Conférences	en poste
33	KAMTCHOUING Pierre	Maître de Conférences	en poste
34	MESSI Jean	Maître de Conférences	en poste
35	NGASSAM Pierre	Maître de Conférences	en poste
36	TCHUEM TCHUENTE Louis	Maître de Conférences	en poste
37	BAPFUBUSA Benoît Alain	Chargé de Cours	en poste
38	CHUNGAG ANYE née NKEH B	Chargée de Cours	en poste
39	DIMO Théophile	Chargé de Cours	en poste
40	DJIETO Lordon Champlain	Chargé de Cours	en poste
41	ENO Anna Aret	Chargée de Cours	en poste
42	ESSOMBA née NTSAMA MBALLA	Chargée de Cours	en poste
43	FOTO MENBOHAN Samuel	Chargé de Cours	en poste
44	KAMGANG René	Chargé de Cours	en poste
45	MONKIEDJE Adolphe	Chargé de Cours	en poste
46	NJAMEN Dieudonné	Chargé de Cours	en poste
47	NJIOKOU Flobert	Chargé de Cours	en poste
48	NOLA Moïse	Chargé de Cours	en poste
49	RAKOTONIRINA Vincent Silvère	Chargé de Cours	en poste
50	BELLET EDIMO Oscar Roger	Assistant	en poste
51	MBOUDJEKA Innocent	Assistant	<i>Absent du poste</i>

3 - DEPARTEMENT DE BIOLOGIE ET PHYSIOLOGIE VEGETALES (B.P.V.)

52	AMOUGOU AKOA	Professeur	Chef de Dépt
53	NKONGMENECK Bernard A	Maître de Conférences	en poste
54	AHANDA TSOUNGUI née MBOLO M.	Chargée de Cours	en poste
55	AMBANG Zachée	Chargé de Cours	en poste
56	ATANGANA ETEME Roger	Chargé de Cours	Vice-Doyen R.C.
57	BELL Joseph Martin	Chargé de Cours	Ch. Serv. Socapalm
58	BONGEN OTTOKO Bruno	Chargé de Cours	Chef Serv. Fac.Sc.
59	BOYOMO ONANA	Chargé de Cours	en poste
60	KENLA Jean Victor	Chargé de Cours	-« -
61	DJOCGOUE Pierre François	Chargé de Cours	-« -
62	NGABA ZOGÓ Félix	Chargé de Cours	Inspecteur au MINREST
63	NWAGA Dieudonné M.	Chargé de Cours	en poste
64	YOUNBI Emmanuel	Chargé de Cours	-« -
65	ZAPFACK Louis	Chargé de Cours	-« -
66	MOSSEBO Dominique Claude	Chargé de Cours	-« -
67	NGOUO Lucas Vincent	Chargé de Cours	-« -
68	TEMGO TCHIASSEU Jean Charlem	Chargé de Cours	-« -
69	TSOATA Esale	Chargé de Cours	-« -
70	ABOMO NDONGO Solange	Assistante	<i>Absente du poste</i>
71	BIYE Elvire Hortense	Assistante	en poste
72	MBARGA BINDZI Marie A.	Assistant	-« -
73	NGALLE Hermine BILLE	Assistante	-« -

4 - DEPARTEMENT DE CHIMIE INORGANIQUE (C.I.)

74	NJOPVOUO Daniel	Professeur	Chef de Dépt. Vice-Doyen Af. Ac.
75	NOAH NGAMVENG Joseph	Professeur	en poste
76	GHOGOMU Paul MINGO	Maître de Conférences	- « -
77	NDIFON Peter TEKE	Maître de Conférences	- « -
78	NEMBA Robert	Maître de Conférences	C.E au MINREST
79	NGAMENI Emmanuel	Maître de Conférences	en poste
80	YOUNANG Elie	Maître de Conférences	- « -
81	AGWARA ONDOH	Chargé de Cours	- « -
82	AVOM Jérôme	Chargé de Cours	Chef de Division UYI

83	BABALE née DJAM DOUDOU	Chargée de Cours	en poste
84	BAIZOUMI ZOUA	Chargé de Cours	Chef Serv. MINUH
85	DJOUFAC WOUMFO Emmanuel	Chargé de Cours	en poste
86	DOUBLA AVALY	Chargé de Cours	Secret. Nat. Educ. Islamique
87	ELIMBI Antoine	Chargé de Cours	en poste
88	GWET Simon-Pierre	Chargé de Cours	- « -
89	KAMWA Jean Dagobert	Chargé de Cours	- « -
90	KETCHA MBADCAM Joseph	Chargé de Cours	- « -
91	KONG SAKEO	Chargé de Cours	C.E. au P.M.
92	LAMINSI Samuel	Chargé de Cours	en poste
93	LIBOUM	Chargé de Cours	- « -
94	MELO née CHINJE Uphie F.	Chargée de Cours	Directeur au MINREST
95	MOULIOM Christophe	Chargé de Cours	en poste
96	NANSEU Charles Péguy	Chargé de Cours	- « -
97	NDIKONTAR Maurice KOR	Chargé de Cours	- « -
98	NGOMO Horace MANGA	Chargé de Cours	C. T. au P. M.
99	PAYOM Gaston	Chargé de Cours	en poste
100	SIGNING Pierre	Chargé de Cours	- « -
101	NENWA Justin	Assistant	- « -

5 - DEPARTEMENT DE CHIMIE ORGANIQUE (CO)

102	SONDENGAM B. Lucas	Professeur	Chef de Dépt.
103	FOMUM TANEE Zacharias	Professeur	en poste
104	FON KIMBU Samuel	Professeur	- « -
105	GHOGOMU TIH Raphaël	Professeur	- « -
106	NGADJUI TCHALEU Bonaventure	Professeur	- « -
107	LONTSI David	Professeur	- « -
108	NKENFACK Augustin Ephraïm	Professeur	- « -
109	TSAMO Etienne	Professeur	- « -
110	DONGO Etienne	Maitre de Conférences	- « -
111	KAPNANG Henriette	Maitre de Conférences	- « -
112	MBAFOR Joseph	Maitre de Conférences	- « -
113	NYASSE Barthélémy	Maitre de Conférences	- « -
114	WANDJI Jean	Maitre de Conférences	- « -
115	Alex de Théodore ATCHADE	Chargé de Cours	- « -
116	FOLEFOC Gabriel NGOSONG	Chargé de Cours	- « -

117	KEUMEDJIO Félix	Chargé de Cours	- « -
118	KOUAM Jacques	Chargé de Cours	- « -
119	MBAZOA née DJAMA Céline	Chargée de Cours	- « -
120	MESSANGA Bernard Blaise	Chargé de Cours	- « -
121	MOFO née NGOUNOU Fernande	Chargée de Cours	- « -
122	NGOUELA Silvère Augustin	Chargé de Cours	- « -
123	NOUNGOUE TCHAMO Diderot	Chargé de Cours	- « -
124	NTEDE NGA Hypolite	Chargé de Cours	- « -
125	NYEMBA née ETOUNOU Anne-M.	Chargée de Cours	- « -
126	PEGNYEMB Dietidonné Emmanuel	Chargé de Cours	- « -
127	TCHOUANKEU Jean-Claude	Chargé de Cours	- « -
128	THI née NGO BILON E. Anastasie	Chargée de Cours	- « -
129	YANKEP Emmanuel	Chargé de Cours	- « -

6 - DEPARTEMENT D'INFORMATIQUE (IN)

130	KAMGNIA Emmanuel	Chargé de Cours	Chef de Dépt .
131	TCHUENTE Maurice	Professeur	MINESUP
132	FOTSO Pauline Laure	Maître de Conférences	Chef Service UYI
133	BOYOM SOP Flaubert .	Chargé de Cours	en poste
134	FOUDA NDJODO Marcel	Chargé de Cours	Chef de Service MINESUP
135	KAMDOUM Basile	Chargé de Cours	en poste
136	LOUKA Basile	Chargé de Cours	- « -
137	NDOUNDAM René	Chargé de Cours	-« -
138	NZALI Jean-Pierre	Chargé de Cours	-« -
139	TINDO Gilbert	Chargé de Cours	- « -
140	ATSA ETOUNDI Roger	Assistant	-« -
141	TAPAMO	Assistant	-« -

7 - SERVICE DES LANGUES (LA)

8 - DEPARTEMENT DE MATHÉMATIQUES (MA)

1

142	BEKOLLE David	Professeur	Chef de Dépt
143	NOUTCHEGUEME Norbert	Professeur	en poste

144	DOSSA COSSY Marcel	Maître de Conférences	en poste
145	NGUETSENG Gabriel	Maitre de Conférences	Chef Centre Calcul UYI
146	NJOCK Georges Edward	Maître de Conférences	en poste
147	TAYOU SIMO	Maître de Conférences	- « -
148	WAMON François	Maître de Conférences	- « -
149	WOUAFO KAMGA Jean	Maître de Conférences	Doyen Fac. Sc.
150	BATE EYA Hans MBI	Chargé de Cours	En poste
151	BINZOULI Etienne Jean-Jacques	Chargé de Cours	- « -
152	EMVUDU WONO Yves S	Chargé de Cours	Chef Serv. MINESUP
153	FOMEKONG Christophe	Chargé de Cours	en poste
154	KIKI Maxime Armand	Chargé de Cours	- « -
155	MBAKOP Guy Merlin	Chargé de Cours	- « -
156	MBIANDA Gilbert	Chargé de Cours	- « -
157	MEWOLI Boulchard	Chargé de Cours	Chef Divis. MINESUP
158	NDAKBO Victor	Chargé de Cours	en poste
159	NGUIMTSA Charles	Chargé de Cours	- « -
160	NKUIMI JUGNIA Célestin	Chargé de Cours	- « -
161	NOUNDJEU Pierre	Chargé de Cours	- « -
162	TCHANGANG Roger Duclos	Chargé de Cours	- « -
163	TCHAPNDA NJABO Sophonie Blaise	Chargé de Cours	-« -
164	TIAYA TSAGUE N. Anne-Marie	Chargée de Cours	- « -
165	TONGA Marcel	Chargé de Cours	- « -
166	ZAME Alfred	Chargé de Cours	Chef Ser. MINCOM
167	TAN AHANDA Barnabé	Assistant	Chef Ser. MINCOM
168	TSEBO Alexis	Assistant	en poste
9 - DEPARTEMENT DE PHYSIQUE (PH)			
169	OWONO ATEBA	Maître de Conférences	Chef de Dépt. Vice-Recteur UYI
170	DOMINGANG Samuel	Professeur	en poste
171	KOFANE Timoléon Crépin	Maître de Conférences	Chef Division UYI
172	MANGUELLE-DICOUM Eliézer	Maître de Conférences	en poste
173	MIKANKAM KAMGA François	Maître de Conférences	- « -
174	WOAFO Paul	Maître de Conférences	- « -
175	EKOBENA FOUA Henri Paul	Chargé de Cours	- « -
176	ESSIMBI ZOBO Bernard	Chargé de Cours	-« -

177	KWENDE MBANWI Teddy	Chargé de Cours	- « -
178	MAGA Emire Mondésir	Chargé de Cours	- « -
179	MBANE BIOUELE	Chargée de Cours	- « -
180	MBONO SAMBA Yves Christian U	Chargé de Cours	- « -
181	MOUSSA Idolko	Chargé de Cours	Sous- Directeur MINUH
182	NDJAKA Jean Marie Bienvenu	Chargé de Cours	en poste
183	NDOHI Raphaël	Chargé de Cours	-« -
184	NIOMO Donatien	Chargé de Cours	-« -
185	NOUAYOU Robert	Chargé de Cours	-« -
186	PEMHA Elkana	Chargé de Cours	-« -
187	TABOD Charles TABOD	Chargé de Cours	- « -
188	TCHAWOUA Clément	Chargé de Cours	- “ -
189	TCHOFFO Fidèle	Chargé de Cours	- « -
190	ZEKENG Serge Sylvain	Chargé de Cours	-« -
191	BEN-BOLIE Germain Hubert	Assistant	- « -
192	BIYA MOTTO Frédéric	Assistant	Chef Service UYI
193	NDOP Joseph	Assistant	en poste

10- DEPARTEMENT DES SCIENCES DE LA TERRE (S.T.)

194	TCHOUA M Félix	Professeur	en poste
195	BILONG Paul	Maitre de Conférences	Chef Dépt p.i. Vice-Doyen Sc.
196	EKODECK Georges E.	Maitre de Conférences	Vice-Recteur UYII
197	NDJENG Emmanuel	Maitre de Conférences	Directeur UYI
198	NZENTI Jean-Paul	Maitre de Conférences	en poste
199	ABOSSOLO née ANGUE Monique	Chargée de Cours	- « -
200	BEKOA Etienne	Chargé de Cours	- « -
201	BITOM Dieudonné	Chargé de Cours	- « -
202	BONDJE Charles	Chargé de Cours	- « -
203	DJEUDA TCHAPNGA Henri Bosco	Chargé de Cours	- « -
204	GHIOMU Richard TANWI	Chargé de Cours	- « -
205	KAMGANG Pierre	Chargé de Cours	- « -
206	LAMILLEN BILLA Daniel	Chargé de Cours	- « -
207	MOUAFU Lucas	Chargé de Cours	- « -
208	MOUNDI Amidou	Chargé de Cours	Chef Service UYI
209	MIVONDO ONDOA Joseph	Chargé de Cours	Chef Serv. MINVILLE

210	NDAM NGOUPAYOU Jules-Remy	Chargé de Cours	en poste
211	NDJIGUI Paul -- Désiré	Chargé de Cours	- « -
212	NGONGE Emmanuel Donald	Chargé de Cours	- « -
213	NGOS III Simon	Chargé de Cours	-« -
214	NJILAH Isaac KONFOR	Chargé de Cours	- « -
215	NJOM Bernard de Lattre	Chargé de Cours	- « -
216	NKONGUIN NSIFA Emmanuel	Chargé de Cours	- « -
217	NKOUMBOU Charles	Chargé de Cours	- « -
218	NYECK Bruno	Chargé de Cours	- « -
219	TCHOUANKOUE Jean-Pierre	Chargé de Cours	- « -
220	TEMDJIM Robert	Chargé de Cours	- « -
221	YENE ATANGANA Joseph Q	Chargé de Cours	Chef Serv. MINEF
222	YONGUE nce FOUATEU Rose	Chargée de Cours	- « -
223	BISSO Dieudonné	Assistant	- « -
224	ESSONO Jean	Assistant	- « -
225	MINYEM Dieudonné	Assistant	-« -
226	NGO BIDJECK Louise Marie	Assistante	- « -
227	TCHAKOUNTE Jacqueline	Assistante	- « -
228	ZO'O ZAME Philémon	Assistant	- « -

**REPARTITION DES ENSEIGNANTS PERMANENTS
DE LA FACULTE DES SCIENCES
PAR DEPARTEMENTS
+ SERVICE DES LANGUES**

DEPARTEMENT	NOMBRE D'ENSEIGNANTS				
	PR	MC	CC	ASS	Total
BIOCHIMIE	0	1	23(10)	4(3)	28 (13)
BIOLOGIE ET PHYSIOLOGIE ANIMALES	1	7	13 (3)	2 (1)	23 (4)
BIOLOGIE ET PHYSIOLOGIE VEGETALES	1	1	16 (1)	4(3)	22 (4)
CHIMIE INORGANIQUE	2	5	20(2)	1 (0)	28 (2)
CHIMIE ORGANIQUE	8	5(1)	15 (4)	0 (0)	28 (5)
INFORMATIQUE	1	1(1)	8 (0)	2 (0)	12 (1)
MATHEMATIQUES	2	6	17 (1)	2 (0)	27 (1)
PHYSIQUE	1	5	16 (1)	3 (0)	25 (1)
SCIENCES DE LA TERRE	1	4	24 (2)	6 (2)	35 (4)
SERVICE DES LANGUES					
Total	17 (0)	35 (2)	152 (24)	24 (9)	228 (35)

Soit un total de :

- Professeurs **17**
- Maîtres de Conférences **35** dont **1** non Camerounais
- Chargés de Cours **152** dont **3** non Camerounais
dont **1** absent du poste
- Assistants **24** dont **1** non Camerounais et **3** absents du poste
- () = Nombre de femmes

*À celui qui m'aime (Jésus-Christ), qui m'a délivré de tous mes péchés par son sang, à lui
la gloire et le pouvoir aux siècles des siècles | Amen | (Apo 1 : 5-6)*

Dédicaces

Je dédie ce travail à

- Mes parents Mr et Mme **Jérémie & Clémentine KOMBOU** pour leur amour et pour tous les sacrifices qu'ils ont consentis à mon égard. Trouvez ici l'expression de ma profonde gratitude.
- Mon oncle, Grand-père et même père, **Dr Flobert NJIOKOU** pour qui les mots me manquent pour exprimer ma reconnaissance. Merci pour la grande sollicitude que tu as sans cesse manifesté à mon égard sur tous les plans.
- Ma sœur aînée, madame **Suzie Ntadoun** en souvenir de tous les durs et bons moments que nous avons passés ensemble depuis la tendre enfance à la maison comme sur les routes d'école ou tout simplement dans la vie.

A mon épouse

Murielle Judith

Remerciements

Les recherches effectuées durant cette thèse ont bénéficié d'un soutien financier du MIM/TDR sous la forme d'un projet de recherche N° 980101 accordé aux Drs John Vululé et Timoléon Tchuinkam, ainsi qu'une bourse RTG N°A00797 que j'ai obtenue du TDR/OMS. J'ai également bénéficié d'une allocation de recherche de l'IRD (Institut de Recherche pour le Développement).

Mes remerciements s'adressent à tous ceux qui de près ou de loin ont contribué à la réalisation de ce travail. Je pense particulièrement à :

Dr Auguste Bilongo Manéné ancien Secrétaire Général de L'OCEAC, pour m'avoir accueilli à l'OCEAC où tout ce travail a été mené. Merci pour l'engagement que vous avez eu dans la formation des jeunes chercheurs de nos pays d'Afrique Centrale.

Monsieur le Secrétaire Général de l'OCEAC, le **Dr Jean Léon Moka** pour avoir permis que ce travail puisse continuer dans son institution.

Monsieur le Représentant de l'IRD au Cameroun, le **Dr François Rivière** pour sa disponibilité vis à vis des sollicitations des allocataires de recherche de l'IRD. Je suis particulièrement sensible à l'action que l'IRD mène pour la formation et l'encadrement des jeunes chercheurs des communautés du Sud.

Pr Albert Samé Ekobo pour avoir accepté de m'encadrer malgré ses multiples occupations. Je vous suis reconnaissant pour toute l'attention que vous avez constamment porté sur mon travail ainsi que pour vos conseils. Je garde un très bon souvenir de toutes les discussions que nous avons eues sur ce travail et sur d'autres sujets d'ordre scientifique. Je vous remercie également pour votre disponibilité vis à vis de mes multiples sollicitations sur le plan administratif (allocation IRD, projet TDR/OMS).

Dr Didier Fontenille pour m'avoir accueilli dans son équipe de recherche. Les mots me manquent pour vous signifier ma profonde gratitude pour la formation à la recherche que j'ai obtenue à vos côtés durant toutes ces 4 années. Trouvez ici ma profonde reconnaissance pour les multiples stages de formation que vous m'avez permis de bénéficier en France et aux

Etats-Unis. Ces stages m'ont permis de découvrir d'autres dimensions sur le plan scientifique et humain. Je garde de vous ces qualités que vous cultivez si bien à savoir le sérieux, l'application, le dynamisme et la bonne gestion des hommes et des projets de recherche. J'ai l'espoir que notre collaboration ne s'arrêtera pas à ce niveau mais qu'elle pourra s'approfondir pour une plus grande production scientifique. Merci une fois de plus.

Dr Frédéric Simard pour sa grande disposition à mon égard tant sur le plan scientifique qu'humain. Je garderai toujours présent en moi toutes les riches discussions que nous avons eues à propos de ce travail. Tu m'as fait bénéficier de tes connaissances en génétique de populations tout comme en biologie moléculaire sans lesquelles la clarté de ce travail aurait sûrement souffert. J'ai bon espoir de voir notre collaboration aller de l'avant.

Au chef du département de Biochimie, **Pr François-Xavier Etoa**, pour tout le sérieux avec lequel vous veillez à la formation des étudiants dans notre département malgré vos multiples occupations.

Dr Tovi Lehmann pour m'avoir accueilli dans son laboratoire au CDC d'Atlanta aux Etats-Unis. Je te dis merci pour ton aide précieuse dans l'analyse des données de ce travail.

Au **Pr. Vincenzo Petrarca** et au **Dr Alessandra Della Torre** de l'Université «la Sapienza» de Rome en Italie pour le soutien qu'ils m'ont accordé dans l'analyse cytogénétique des échantillons d'*An. gambiae* du Cameroun.

Aux enseignants du Département de Biochimie qui par leurs enseignements et encadrements ont contribué pour beaucoup à l'aboutissement de ce travail.

Dr Timoléon Tchuinkam pour tous les conseils de grand frère qu'il m'a fait profiter. Je n'oublierai pas que la grande partie de ce travail a été financée par le projet MIM/TDR dont tu étais le co-responsable. Je te remercie de m'y avoir associé.

Dr Gustave Simo pour avoir bien voulu m'initier au laboratoire à mes débuts. Le goût du travail propre et bien fait, de l'assiduité au «labo de biomol» me viennent de l'apprentissage à tes côtés. Trouve ici l'expression de ma grande reconnaissance.

Dr Pierre Kengne pour son aide dans la mise en route des protocoles de biologie moléculaire lors de ce travail et pour tous ses conseils de grand frère sur divers aspects de la vie. Je te signifie ici ma profonde gratitude.

Aux collègues de laboratoire **Drs Christophe Antonio Nkondjio, Parfait Awono et Josiane Etang** pour la collaboration qui a été la nôtre durant toute cette période. J'ai bénéficié énormément des discussions et échanges que nous avons eus sans cesse. La qualité de ce travail y est redevable. Je crois que ce socle de fraternité que nous avons bâti ne s'arrêtera pas à ce niveau mais que de beaux jours sont devant nous tant sur les plans scientifique qu'humain.

Aux cadets du laboratoire que sont **Elysée Nchoutpouen, Chouaïbou Mohammadou, Honoré Kengné**. Je vous souhaite bon courage.

Jean-Claude Toto, Sylvie Kemleu, Rose Nyamban, Françoise Eboumbou, Jean Pierre Agbor, Serge Donfack, Isaac Tchikangwa et Elie Guelsenatia pour tout l'appui technique que vous m'avez apporté ainsi que pour les chaleureux rapports que nous avons eus tout au long de ce travail.

Aux camarades de promotion de Doctorat à savoir **Guy Collince Kamdem** (je garde un très bon souvenir de notre relation qui d'académique s'est transformée en amicale. Je te souhaite bon courage dans tes entreprises), **Daniel Fotio, Roger Ponka, Hermann Mbouobda, Thérèse Metoh, Cathérine Tazong...**

Mes amis et frères en Christ, **Guy Serges et Thérèse Tchakoumegne, Alain Wandji, Merlin et Lise Nkouako, Irène et Michaël Kotzine, Luc Fotsing, Elsa Gwonkam, Calvin Tchoualong, Sévérin Kwayep, Jules Bruno Tchokobou** et tout le groupe UCJG de la paroisse EEC de Biyem-Assi, pour toutes vos prières à mon égard devant le Seigneur qui certainement ont aidé à l'aboutissement heureux de ce travail. Puisse l'Eternel vous le rendre au centuple.

TABLE DE MATIERES

	Dédicace	i
	Remerciements	iv
	Table de matières	vii
	Résumé	xv
	Abstract	xvi
	Liste des figures	xvii
	Liste des tableaux	xix
	INTRODUCTION	1
	PREMIERE PARTIE : GENERALITES	3
	Chapitre I : Le paludisme	4
I.1	Le paludisme : son épidémiologie	4
I.1.1	Les relations Homme- <i>Plasmodium</i>	4
I.1.2	Les relations <i>Plasmodium</i> -Anophèles	4
I.1.3	Les relations Hommes-Anophèles	5
I.2	Description et biologie des vecteurs du paludisme : les anophèles	5
I.2.1	Taxonomie et rôle vecteur	5
I.2.2	Cycle de développement et biologie des différents stades de l'anophèle	6
	Chapitre II : Les études de génétique des populations	8
II.1	La génétique des populations	8
II.1.1	Définition	8
II.1.2	Modèles mathématiques	9
II.1.2.1	Le modèle de Hardy-Weinberg	9
II.1.2.2	Les modèles démographiques	11
II.1.2.2.1	Les modèles en îles	11
II.1.2.2.2	Les modèles dits "Stepping stone"	12
II.1.2.2.3	Les modèles de voisinage ou isolement par la distance	12
II.1.2.2.4	Le modèle de populations subdivisées et transitoires ou métapopulations	12
II.1.3	Approches expérimentales et traitement des données	13
II.1.3.1	Méthode directe : étude de la dispersion	13

II.1.3.2	Méthodes indirectes en génétique	14
II.2	Les marqueurs moléculaires	16
II.2.1	Définition et propriétés générales	16
II.2.2	Les ADN microsatellites	16
II.2.2.1	Caractéristiques générales	16
II.2.2.2	Propriétés	17
II.2.2.3	Contraintes associées aux microsatellites	20
II.2.2.4	Applications	23
II.2.3	Autres marqueurs ou techniques utilisés en génétique des populations	24
II.3	Espèces et spéciation	24
II.3.1	La spéciation géographique ou allopatrique	25
II.3.2	La spéciation sympatrique	26
II.3.3	La spéciation par remaniements chromosomiques	26
	Chapitre III : Le complexe <i>An. gambiae</i>	26
III.1	Historique	26
III.2	Biologie et distribution géographique des membres du complexe	28
III.3	Structure génétique et spéciation au sein du complexe <i>An. gambiae</i>	30
III.3.1	Cytologie d' <i>An. gambiae s.l.</i>	30
III.3.2	Etudes biochimiques à base d'isoenzymes	33
III.3.3	Les études de génétique moléculaire	33
III.3.4	Apport des microsatellites	34
III.4	La lutte antivectorielle	35
III.4.1	Lutte antilarvaire	36
III.4.2	Lutte imagocide	36
III.4.3	Voie nouvelle de lutte par la génétique	37
III.5	Présentation de l'étude	38
	DEUXIEME PARTIE : METHODOLOGIE GENERALE	39
1	Description des sites pour l'étude de la structuration génétique chez <i>An. gambiae</i> s.s.	40
1.1	Simbock	40
1.2	Mbalmayo	40
1.3	Dschang	40

1.4	Santchou	41
1.5	Nkoteng	41
2	Echantillonnage des spécimens	41
2.1	Capture sur volontaires	41
2.2	Capture par pulvérisation intradomiciliaire d'insecticide	43
3	Identification morphologique	43
4	Caractérisation moléculaire des spécimens	44
4.1	Extraction d'ADN des spécimens	44
4.2	Principe de la PCR	45
4.3	Identification des membres du complexe <i>An. gambiae</i>	46
4.3.1	Déroulement de la PCR complexe <i>An. gambiae</i>	46
4.3.2	Electrophorèse des échantillons issus de la PCR	47
4.4	Détermination de la forme moléculaire des spécimens d' <i>An. gambiae</i>	48
4.4.1	Méthode de PCR-RFLP	48
4.4.2	Méthode par PCR uniquement	48
5	Détermination de la forme chromosomique	50
6	Détermination des paramètres entomologiques de la transmission	51
6.1	Détermination de la préférence trophique par ELISA repas de sang	51
6.1.1	Principe	51
6.1.2	Déroulement de la manipulation	52
6.2	Estimation du taux d'infestation par ELISA	53
6.2.1	Principe	53
6.2.2	Déroulement de la manipulation	54
7	Détermination des génotypes des spécimens aux différents loci d'ADN microsatellites	54
7.1	Amplification des différents loci	54
7.2	Détermination des génotypes	55
7.2.1	Préparation des gels de polyacrylamide	55
7.2.2	Migration	55
7.2.3	Révélation des génotypes sur gel de polyacrylamide	55
7.2.4	Construction des échelles de taille pour locus	56
10	Analyse des données	57
10.1	Test de conformité à l'équilibre de Hardy-Weinberg	57

10.2	Déséquilibre de liaison entre différentes paires de loci	58
10.3	Test de différenciation entre populations	58
10.4	Les statistiques F	58
TROISIEME PARTIE : RESULTATS & DISCUSSIONS		59
CHAPITRE I : Caractérisation moléculaire, cytologique des formes moléculaires M et S d'<i>An. gambiae</i> au Cameroun et leurs rôles épidémiologiques		60
I.1	Rappel de la problématique	61
I.2	Résultats	61
I.2.1	Distribution géographique des formes moléculaires M et S au Cameroun	61
I.2.2	Comparaison des deux méthodes (PCR-RFLP et PCR directe) d'identification des formes moléculaires M et S	63
I.2.3	Caractérisation chromosomique des populations d' <i>An. gambiae</i>	67
I.2.3.1	Test de Hardy-Weinberg pour les populations de l'Adamaoua	72
I.2.4	Variations saisonnières des formes moléculaires M et S dans la localité de sympatrie de Simbock	73
I.2.4.1	Hypothèse de deux saisons climatiques à Simbock	74
I.2.4.2	Hypothèse de 4 saisons climatiques à Simbock	75
I.2.5	Implications des formes moléculaires M et S d' <i>An. gambiae</i> dans la transmission du paludisme	75
I.2.5.1	Indice circumsporozoïtique	78
I.2.5.2	Taux d'inoculation entomologique	78
I.2.5.3	Préférence trophique	79
I.3	Discussion	80
I.3.1	Facteurs influençant la distribution des formes moléculaires M et S	80
I.3.1.1	Exclusion compétitive	80
I.3.1.2	Facteurs éco-climatiques	81
I.3.1.3	Le phénomène d'isolement écologique ou «niche partitionning»	81
I.3.2	Relations formes chromosomiques/formes moléculaires au Cameroun	82
I.3.2.1	La forme moléculaire S	83

I.3.2.2	La forme moléculaire M	84
I.3.3	Isolement reproductif entre les deux formes moléculaires	84
I.3.3.1	Hypothèse d'isolement post-zygotique	85
I.3.3.2	Hypothèse d'isolement prézygotique	85
I.3.4	Impact sur la transmission du paludisme	86
CHAPITRE II : Structure génétique d'<i>An. gambiae</i> au Cameroun		88
II.1	Rappel de la problématique	89
II.2	Résultats	90
II.2.1	Sélection des loci ADN microsatellites pour notre étude	90
II.2.1.1	Résultats des tests sur l'ensemble des 26 loci	91
II.2.1.1.1	Le polymorphisme des loci	91
II.2.1.1.2	Qualité de l'amplification	91
II.2.1.1.3	Taux d'hétérozygotie	92
II.2.1.1.4	Localisation cytologique des loci sur le génome	93
II.2.1.1.5	Taille des allèles	93
II.2.2	Structure génétique au sein des populations d' <i>An. gambiae s.s.</i>	95
II.2.2.1	Caractérisation génétique des populations d' <i>An. gambiae</i> étudiées	95
II.2.2.2	Polymorphisme des loci microsatellites	96
II.2.2.3	Variabilité génétique intra-populationnelle	101
II.2.2.4	Test de la panmixie	101
II.2.2.4.1	Test de panmixie dans chacune des populations	103
II.2.2.5	Différenciation génétique entre populations	106
II.2.2.5.1	Différenciation génétique au sein de chacune des formes moléculaires	106
II.2.2.5.1.1	Forme moléculaire S	106
II.2.2.5.1.2	Forme moléculaire M	106
II.2.2.5.2	Différenciation génétique entre populations de formes moléculaires différentes	108
II.2.3	Tests d'assignation pour évaluation de la différenciation génétique entre populations	108
II.2.3.1	Principe	110
II.2.3.1.1	Méthodes basées sur la probabilité	110
II.2.3.1.1.1	La méthode des fréquences	110

	chromosomique Forest	
II.2.4	Etude de la corrélation entre les tests d'assignation et les statistiques F	114
II.2.5	Etude de l'isolement par la distance entre populations	116
II.3	Discussions	119
II.3.1	Les loci ADN microsatellite comme outil d'estimation de la variabilité génétique	119
II.3.2	Variabilité génétique des populations	120
II.3.3	Structure génétique des populations d' <i>An. gambiae</i> au Cameroun	123
II.3.3.1	Phénomène de renforcement entre M et S à Simbock	123
II.3.3.2	Rôle de la distance dans les profils de différenciation observés	124
II.3.3.3	Les tests d'assignation comme indicateurs de différenciation génétique entre M et S	125
II.3.4	Statut taxonomique des formes moléculaires M et S : Représentent-elles deux espèces ?	127
	CONCLUSION ET PERSPECTIVES	130
1	Apport des loci ADN microsatellites dans la détermination de la structure génétique des populations d' <i>An. gambiae</i>	131
2	Spéciation chez <i>An. gambiae</i>	133
3	Apport des études de structure génétique des populations du complexe <i>An. gambiae</i> sur les mesures de contrôle anti-vectoriel	135
	REFERENCES BIBLIOGRAPHIQUES	138
	ANNEXES	153

Annexe 1

Charles Wondji, Frédéric Simard & Didier Fontenille (2002). Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. *Insect Molecular Biology* 11(1), 11-19.

Annexe 2

Antonio Nkondjio C., Awono-Ambene P., Toto J.C., Nyamban R., Kemleu-Zébazé S. Meunier J.Y., **Wondji C.**, Tchuinkam T & Fontenille D (2002). High malaria transmission in a village close to Yaoundé, the capital city of Cameroon. *Journal of Medical Entomology*. 39 (2) : 350-355.

Annexe 3 : Protocoles détaillés des différentes PCR**Annexe 4 : Préparation des tampons et solutions stocks**

RESUME

Il est indispensable de bien connaître les systèmes vectoriels responsables de la transmission du paludisme en vue d'une meilleure compréhension de l'épidémiologie de la maladie ainsi que pour la mise en place des moyens de contrôle efficaces. Dans cette perspective, nous avons conduit une étude de génétique des populations d'*An. gambiae*, vecteur principal du paludisme au Cameroun. Nous avons choisi les loci d'ADN microsatellites comme principal outil moléculaire pour cette étude. Nous avons caractérisé les populations d'*An. gambiae* au Cameroun tant sur le plan moléculaire que sur le plan chromosomique. Les formes moléculaires M et S sont retrouvées avec chaque fois une nette prédominance d'une des deux formes sur l'autre. L'absence d'hybrides M/S dans cette étude indique l'existence d'un isolement reproductif entre ces deux formes moléculaires. Les formes chromosomiques Forest, Savanna et Mopti ont été trouvées.

En utilisant 10 loci microsatellites, nous avons observé une différenciation génétique (symbolisée par les valeurs de F_{st}) plus importante entre les populations de formes moléculaires différentes alors que les populations appartenant à la même forme moléculaire étaient génétiquement homogènes. Cette hétérogénéité entre les forme M et S est confirmée par les tests d'assignation qui montrent que les spécimens d'une forme moléculaire sont préférentiellement assignés dans une population de la même forme moléculaire. Une forte corrélation a été trouvée entre les résultats obtenus par les statistiques F et ceux obtenus par les tests d'assignation. Dans cette nous n'avons pas mis en évidence un modèle d'isolement par la distance entre les différentes populations.

L'ensemble de ces résultats nous a amené à suggérer que les formes M et S constitueraient deux espèces distinctes ayant divergé récemment.

Mots-clés : *An. gambiae* génétique, vecteur, paludisme, formes moléculaires M et S, ADN microsatellite, tests d'assignation, PCR

Abstract

Better knowledge of vectorial systems involved in malaria transmission is crucial to understand the epidemiology of the disease in the aim of a development of new tools for the control. With respect to this rationale, we have conducted a study of genetic structure of *An. gambiae*, the most important malaria vector, member of the *An. gambiae* complex. We have chosen microsatellite loci as main molecular tool for this study. We have characterized *An. gambiae* populations in Cameroon molecularly and cytogenetically. We have found the two molecular forms M and S with each time one form that is predominant over the other. The absence of M/S hybrids in this study is a strong evidence of the reproductive barrier between M and S molecular forms. We have also found the Savanna, Mopti and Forest chromosomal forms within different *An. gambiae* populations during this study.

By using 10 microsatellites loci, we have observed a more greater genetic differentiation between populations belonging to different molecular forms while populations belonging to same molecular form were genetically homogeneous. This genetic heterogeneity between M and S is confirmed by assignment tests. Results obtained from this method showed that specimens are assigned preferentially to a population of same molecular form than the population from where it's originated. A high correlation have been observed between results from F-statistics and those from assignment tests. No evidence of a model of isolation by distance between the different populations have been observed.

By taking into account all these results, we think that these two molecular forms are probably two different species that have diverged recently.

Keywords : *An. gambiae*, genetic, vector, malaria, M and S molecular forms, microsatellite DNA, assignment tests, PCR.

Liste des figures

Figure 1 : Cycle de développement des anophèles	7
Figure 2 : Femelle d' <i>An. gambiae</i> s.l. prenant un repas de sang sur homme	27
Figure 3 : Répartition des espèces du complexe <i>Anopheles gambiae</i> en Afrique	29
Figure 4 : Inversions et gradient climatique caractérisant les formes chromosomiques d' <i>An. gambiae</i>	31
Figure 5 : Carte du Cameroun avec les différentes localités où les échantillons ont été collectés durant notre étude	42
Figure 6 : Schéma du principe de la PCR	45
Figure 7 : Structure de l'ADNr des diptères	46
Figure 8 : Profil de migration sur gel d'agarose des spécimens du complexe d' <i>An. gambiae</i> après PCR d'identification d'espèces selon Scott <i>et al.</i> (1993)	47
Figure 9 : Localisation des amorces utilisées pour la détermination des formes moléculaires d' <i>An. gambiae</i> sur un fragment de 2,3 kb dans la région des IGS de l'ADNr d'après Favia <i>et al.</i> (2001)	49
Figure 10 : Profil de migration sur gel d'agarose des spécimens de formes moléculaires M et S d' <i>An. gambiae</i>	50
Figure 11 : Chromosomes polytènes d' <i>An. gambiae</i>	51
Figure 12 : Schéma du principe de l'ELISA repas de sang	52
Figure 13 : Principe de l'ELISA Circumsporozoïtique (CSP)	53
Figure 14 : Carte du Cameroun illustrant les différentes localités où la détermination des formes moléculaires M et S a été effectuée	65
Figure 15 : Fréquences mensuelles des formes moléculaires M et S d'octobre 1998 à mars 2001 à Simbock	76
Figure 16 : Variation mensuelle des formes moléculaires M et S à Simbock en fonction de la pluviométrie	77
Figure 17 : Relation entre les formes moléculaires et les formes chromosomiques au Cameroun	83
Figure 18 : Localisation cytologique des 10 loci microsatellites d' <i>An. gambiae</i> choisis	94
Figure 19 : Profil de migration du locus H417 sur gel de polyacrylamide	94
Figure 20 : Fréquences alléliques des 10 loci au sein des différentes populations d' <i>An. gambiae</i> forme Forest	98

Figure 21 : Histogramme des proportions d'assignation des spécimens des différentes localités	112
Figure 22 : Droite de corrélation entre les valeurs de F_{st} et les résultats des tests d'assignation	116
Figure 23 : Droite de régression illustrant l'isolement par la distance entre toutes les populations	118
Figure 24 : Droite de régression de l'isolement par la distance entre populations de forme S	118

Liste des tableaux

Tableau 1 : Amorces utilisées pour la PCR complexe <i>An. gambiae</i>	47
Tableau 2 : Amorces utilisées dans la PCR pour identification des formes moléculaires d' <i>An. gambiae</i>	49
Tableau 3 : Distribution des formes moléculaires M et S d' <i>An. gambiae</i> au Cameroun du Nord au Sud	64
Tableau 4 : Paramètres éco-climatiques des localités étudiées (en allant du Nord au Sud)	66
Tableau 5 : Détermination des formes chromosomiques	68
Tableau 6 : Fréquence relative des inversions chromosomiques de 187 spécimens d' <i>An. gambiae</i> dans les localités étudiées au Cameroun	69
Tableau 7 : Nombre d'individus des formes M et S identifiés mensuellement à Simbock d'octobre 1998 à mars 2001	73
Tableau 8 : Distribution des deux formes M et S en fonction des saisons	74
Tableau 9 : Distribution des formes M et S en fonction des 4 saisons climatiques à Simbock	75
Tableau 10 : Récapitulatifs des valeurs des paramètres entomologiques des formes moléculaires M et S d' <i>An. gambiae</i> dans 4 localités du Cameroun	79
Tableau 11 : Caractéristiques des loci sélectionnés pour l'étude de la variabilité génétique d' <i>An. gambiae</i>	96
Tableau 12 : Polymorphisme des loci microsatellites	97
Tableau 13 : Variabilité génétique des différentes populations au niveau des 10 loci utilisés	102
Tableau 14 : Test de déviation par rapport à l'équilibre de Hardy-Weinberg : Fis et valeurs de la probabilité P des tests d'équilibre de Hardy-Weinberg, basés sur l'hypothèse nulle d'un déficit d'hétérozygotes pour toutes les populations d' <i>Anopheles gambiae</i>	105
Tableau 15 : Différenciation génétique (estimée par Fst) entre les populations De la même forme moléculaire d' <i>An. gambiae</i>	107
Tableau 16 : Différenciation génétique entre populations de formes moléculaires différentes	109
Tableau 17 : Résultats des tests d'assignation	113
Tableau 18 : Corrélation entre proportions d'assignation et valeurs de Fst	115

Liste des Abréviations

ADN :	Acide Désoxyribonucléique
ADNr :	ADN ribosomal
ADNmt :	ADN mitochondrial
ARN :	Acide ribonucléique
ARNr :	ARN ribosomal
CSP :	Circumsporozoite Protein
DDT :	Dichlorodiphényl Trichloroéthane
DNTP :	Desoxynucléotides Triphosphate
ELISA :	Enzyme Linked Immuno-Sorbent Assay
IAM :	Infinite Allele Model
IGS :	Intergenic Spacer
ITS :	Internal Transcribed Spacer
KAM :	k-allele Model
kdr :	knock down resistance
OBET :	Odour Baited Entry Trap
OMS :	Organisation Mondiale de la Santé
Pb :	Paire de bases
PBS :	Phosphate Buffered Saline
PCR :	Polymerase Chain Reaction
QTL :	Quantitative Trait Loci
RAPD :	Random Amplified Polymorphic DNA
RFLP :	Restriction Fragment Length Polymorphism
SSCP :	Single Strand Conformation Polymorphism
SMM :	Stepwise Mutation Model
SOSUCAM :	Société Sucrière du Cameroun
STR :	Short Tandem Repeat
SSR :	Simple Sequence Repeats
TBE :	Tris-Borate-EDTA
TE :	Tris-EDTA
TEMED :	Tetra-méthyl-éthylène-diamine
TIE :	Taux d'Inoculation Entomologique
TR :	Tandem Repeats
VNTR :	Variable Number of Tandem Repeats

INTRODUCTION

En ce début du 21^{ème} siècle, le paludisme demeure de loin la parasitose la plus importante dans le monde intertropical et particulièrement en Afrique au sud du Sahara. Il concerne 36% de la population mondiale et est présent dans près de 102 pays avec plus de 2 milliards de personnes vivant dans les zones de transmission. (Danis & Mouchet, 1991 ; Rogier *et al.*, 2001). La plupart des données statistiques estime entre 300 et 500 millions le nombre de personnes atteintes par an avec environ 1,5 à 3 millions de décès notamment chez les enfants âgés de moins de 5 ans et chez les femmes enceintes. 90% des cas sont recensés en Afrique sub-saharienne (OMS, 1998). Le paludisme est causé par quatre espèces de parasites du genre *Plasmodium* qui sont transmis à l'homme par des moustiques femelles du genre *Anopheles*.

De nombreux obstacles jalonnent l'effort de la communauté scientifique à contrôler le paludisme. La campagne d'éradication des années 50-60 a connu un échec dû entre autres à la résistance des vecteurs aux insecticides notamment au DDT et à la mauvaise connaissance des vecteurs. La biologie du parasite comme celle des vecteurs et la complexité de leurs interrelations restent encore obscures sur bien de points. La grande diversité des faciès épidémiologiques de la maladie, le faible niveau des moyens alloués à la recherche dans ce domaine constituent des obstacles majeurs à la mise sur pied des programmes de contrôle efficaces. De plus, les infrastructures de santé inadéquates dans la plupart des pays, l'insuffisance du personnel de santé qualifié et de chercheurs en paludologie, rendent plus difficile le contrôle du paludisme (Touré *et al.*, 1999). Les nombreuses tentatives pour mettre au point un vaccin n'ont pas encore été couronnées de succès, d'où le besoin urgent de mettre sur pied de nouvelles stratégies de lutte moins coûteuses.

L'OMS dans son initiative "Roll Back Malaria" a pour objectif une diminution de moitié de l'impact du paludisme d'ici 2010 (Nabarro, 1999). Le succès de cette initiative passe par la mise sur pied de nombreuses mesures pour améliorer le diagnostic, la prise en charge des malades et le contrôle des populations de vecteurs. Les moyens de la lutte antivectorielle préconisés sont l'utilisation des moustiquaires imprégnées d'insecticides et les pulvérisations intra-domiciliaires pour diminuer le contact homme/moustique pour faire baisser de façon significative le taux de transmission du paludisme. Ce contrôle du paludisme nécessite une connaissance approfondie de la biologie, de l'écologie et la génétique des populations de vecteurs.

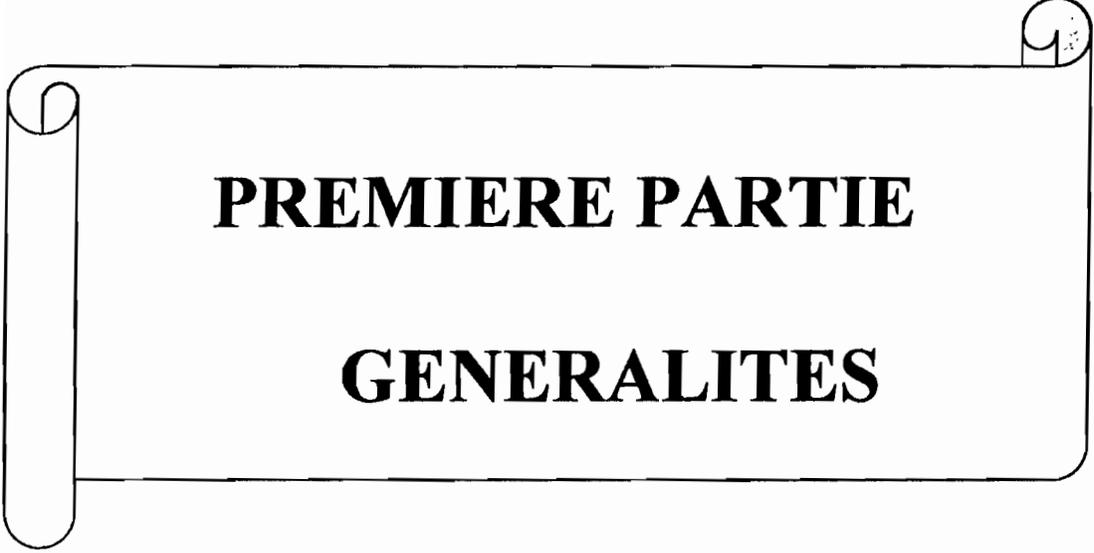
Pour mener une lutte adéquate contre les vecteurs, il est nécessaire de connaître la diversité spécifique des populations d'anophèles, leur distribution géographique et leur rôle vectoriel. Depuis la mise en évidence du complexe *An. maculipennis* en Europe dans les années 30, il a été montré à plusieurs reprises d'une part que les vecteurs du paludisme appartiennent en majorité à des complexes d'espèces jumelles et d'autre part que ces espèces souvent sympatriques, présentent des différences comportementales, écologiques et génétiques qui influencent leur rôle vecteur (Coluzzi,

1970 ; Coluzzi *et al.*, 1979). Le succès des programmes de contrôle des vecteurs dépend de la capacité à pouvoir reconnaître les espèces vectrices des espèces non vectrices. Les études conduites jusqu'à présent montrent une grande complexité du système vectoriel en Afrique tropicale. Trois espèces d'anophèles y sont considérés comme vecteurs majeurs. Ce sont *Anopheles gambiae*, *Anopheles arabiensis* et *Anopheles funestus*. *Anopheles moucheti* et *Anopheles nili* sont considérés comme de bons vecteurs, mais d'importance locale. *An. gambiae* et *An. arabiensis*, tous deux membres du complexe *Anopheles gambiae*, sont responsables à plus de 90% de la transmission du paludisme en Afrique (Robert & Carnevale, 1984). Pour cette raison, de nombreuses études et de nombreux programmes de lutte sont menés sur ces deux espèces de vecteurs. *An. gambiae s.s.*, est subdivisé en cinq formes chromosomiques ; Mopti, Savanna, Bamako, Bissau et Forest (Coluzzi *et al.*, 1979 ; Favia *et al.*, 1997). Certaines de ces formes vivant quelques fois en sympatrie dans certaines régions, ne présentent que de faibles taux d'échanges géniques (Lanzaro *et al.*, 1998). Un autre niveau de structuration génétique a été mis en évidence chez *An. gambiae* avec la définition de deux formes moléculaires dites M et S (Favia *et al.*, 1997, 2001). Ces deux formes moléculaires ne présentent pas toujours une même sensibilité face aux insecticides. En effet des études menées au Burkina Faso et en Côte d'Ivoire ont montré que la forme moléculaire S présentait la mutation kdr (Knock down resistance) et donc développait une résistance aux pyréthrinoïdes alors que la forme moléculaire M ne présentait pas (ou à faible fréquence) cette mutation et par conséquent était sensible à ces insecticides (Chandre *et al.*, 1999a). Cette complexité du système vectoriel impose une meilleure connaissance des populations de vecteurs à un niveau local. Pour améliorer ces connaissances, les études de génétique des populations des vecteurs en utilisant des marqueurs moléculaires puissants comme les ADN microsatellites, sont aujourd'hui indispensables. En effet, ces études peuvent aider à une meilleure caractérisation des populations de vecteurs et à une estimation du niveau de flux de gènes (notamment les gènes de résistance aux insecticides) entre ces populations.

Dans cette perspective, nous avons mené une étude de génétique des populations d'*Anopheles gambiae* à l'aide des marqueurs ADN microsatellites. Ce travail sera présenté en deux grands chapîtres avec

Premièrement : La caractérisation moléculaire et cytologique des formes moléculaires M et S d'*An. gambiae* du Cameroun ainsi que leur implication dans la transmission du paludisme.

Secondairement : La structure génétique des populations d'*An. gambiae* au Cameroun.



PREMIERE PARTIE

GENERALITES

CHAPITRE I : Le paludisme

I.1- Le paludisme : son épidémiologie

Le paludisme est une maladie complexe au niveau de son épidémiologie. Celle-ci varie d'un continent à un autre, d'une région à une autre, d'un village à un autre et parfois au sein d'un même village. Les raisons de cette grande hétérogénéité ne sont que partiellement connues. Elle dépend de facteurs qui pourraient être rassemblés en grands groupes : 1) facteurs affectant la distribution des anophèles et leur capacité vectorielle, 2) caractéristiques génétiques et biologiques des parasites, 3) facteurs de susceptibilité de l'homme et 4) développement socio-économique et interventions de la lutte anti-paludique. Nous présenterons l'épidémiologie du paludisme en définissant trois niveaux de relations impliquant le *Plasmodium*, l'anophèle et l'homme.

I.1.1- Les relations Homme - *Plasmodium*

Sur une centaine de *Plasmodium* actuellement connus, quatre sont responsables du paludisme humain. Il s'agit de *Plasmodium falciparum* Welch 1897, *P. malariae* Laveran, 1881, *P. vivax* Grassi & Feletti, 1880 et *P. ovale* Stephens 1922. Parmi ces quatre espèces, *P. falciparum* est de loin la plus importante car pouvant causer des accès aigus mortels en l'absence de traitement médical. *P. vivax* n'entraîne que très rarement la mort, mais provoque une morbidité importante avec des répercussions économiques sérieuses, *P. malariae* moins fréquent, a été accusé de troubles rénaux alors que *P. ovale* est considéré comme peu pathogène. Face à l'agression des parasites et, en particulier de *P. falciparum*, les sujets vivant en zone d'endémie développent une immunité d'autant plus solide et précoce qu'ils sont régulièrement soumis aux piqûres infectantes des anophèles. Cette prémunition se manifeste par la présence d'un nombre élevé de porteurs de parasites sans signes cliniques, dans les zones de forte transmission. Elle s'établit au prix d'une mortalité infanto-juvénile élevée alors que la mortalité est moins élevée chez les adultes.

I.1.2- Les relations *Plasmodium* – Anophèles

L'anophèle entre en contact avec le *Plasmodium* lors d'un repas de sang pris sur un homme infecté. Le moustique prélève alors des trophozoïtes et des schizontes, qui seront rapidement détruits, et éventuellement les gamétocytes, éléments sexués du parasite chez l'homme et forme infectante du *Plasmodium* pour l'anophèle. Au niveau de l'estomac du moustique, les gamétocytes se différencient en gamètes mâles et femelles. Il s'en suit une

fécondation qui aboutit à la formation d'un zygote mobile appelé ookinète qui traverse la membrane intestinale. Il se loge au niveau de la membrane basale des cellules épithéliales où il se différencie en oocyste. Ceux-ci après maturation et mitoses donnent des sporozoïtes (haploïdes) qui ensuite gagnent les glandes salivaires. Le moustique est alors infectant et lors des repas de sang ultérieurs, les sporozoïtes, forme infectante du *Plasmodium*, seront injectés avec la salive de l'anophèle à l'homme.

I.1.3- Les relations Hommes – Anophèles

Les anophèles viennent au contact de l'homme pour prendre un repas de sang pour la maturation de leurs œufs. L'aptitude de l'anophèle à transmettre le parasite à l'homme est fonction de divers facteurs biologiques, écologiques et génétiques très spécifiques qui déterminent la capacité vectorielle de l'espèce. Le degré de contact entre une espèce anophélienne et une population humaine, dépend aussi de nombreux facteurs et ceux-ci peuvent être évalués par l'établissement d'indices très classiques notamment :

- **Indices traduisant la densité anophélienne** : Ceux-ci concernent particulièrement la fraction de la population d'anophèles prenant leur repas de sang sur homme (anthropophilie). Cet indice permet d'évaluer le contact homme – anophèle (nombre d'anophèles par case, nombre de piqûres par homme et par nuit).

- **Indices de préférences trophiques** : Il s'agit du degré d'anthropophilie ou de la zoophilie (comportement du moustique à se nourrir sur homme ou sur animal respectivement), du degré d'endo ou d'exophagie (comportement du moustique à prendre son repas de sang à l'intérieur et à l'extérieur respectivement) ou de la durée du cycle gonotrophique (période allant de la prise du repas de sang à la ponte des œufs).

- **Indices bio-écologiques** importants sur le plan épidémiologique : abondance et nature des gîtes larvaires et, en particulier, leur situation par rapport aux habitats humains, dispersion des adultes, degré d'endo ou d'exophilie (comportement du moustique à se reposer à l'intérieur ou à l'extérieur des maisons) et la longévité.

I.2- Description et biologie des vecteurs du paludisme : les anophèles

I.2.1- Taxonomie et rôle vecteur

L'importance des anophèles dans les zones tropicales du globe tient principalement à leur rôle de vecteur de *Plasmodium*. Toutefois, en Afrique, l'on a montré que *An. gambiae* et *An. funestus* jouent un rôle majeur dans les zones rurales dans la transmission de *Wuchereria bancrofti*, agent de la filariose lymphatique. En outre, ces anophèles (notamment *An. funestus*

et *An. gambiae*) ont joué un rôle majeur dans les flambées épidémiques des fièvres dues au virus o'nyong-nyong en Afrique de l'Est (Gillies & De Meillon, 1968).

Les anophèles sont des diptères nématocères appartenant à la famille des *Culicidae*, de la sous-famille des *Anophelinae* et du genre *Anopheles*. Sur plus de 500 espèces d'anophèles connues, près d'une cinquantaine seulement peuvent assurer la transmission du paludisme de manière plus ou moins efficace. Mais seules environ 20 espèces sont considérées comme vecteurs majeurs du paludisme, les autres jouant un rôle principal ou secondaire de manière localisée (Rhodain & Perez, 1985).

1.2.2- Cycle de développement et biologie des différents stades d'anophèles

Les anophèles comme tous les diptères sont holométaboles c'est-à-dire qu'ils présentent des métamorphoses complètes. Leur développement comprend quatre stades de vie : œuf, larve, nymphe et adulte ou imago. Les stades jeunes, pré-imaginaux sont aquatiques alors que les adultes ont une vie aérienne (voir figure 1).

Les œufs sont pondus isolément par les femelles à la surface des collections d'eau. Les pontes peuvent comprendre entre 200 et 300 œufs, chacun mesurant entre 0,6 et 0,8 mm de long. Ces œufs se maintiennent à la surface des eaux à l'aide de flotteurs dont ils sont munis latéralement. La durée du cycle ovulaire est de 2 à 3 jours au bout duquel intervient l'éclosion. Les variations de forme, de taille et de coloration des œufs sont parfois utilisées en taxonomie pour distinguer des espèces d'anophèles très proches.

On distingue quatre stades larvaires successifs séparés par des mues. Les gîtes larvaires sont constitués par des collections d'eau dont le type est extrêmement variable suivant l'espèce. Dans l'eau du gîte, les larves se trouvent au repos, sous la surface de l'eau, respirant l'air atmosphérique grâce à leurs spiracles. Elles plongent en profondeur lorsqu'elles font face à une menace. Leur nourriture est constituée d'éléments planctoniques (bactéries, algues microscopiques, protozoaires...). Les pièces buccales adaptées au régime saprophytique, sont de type broyeur. La durée des quatre stades larvaires est environ de 8 à 12 jours. A chaque mue, l'exuvie (tégument) du stade précédent, se retrouve abandonnée dans l'eau et la dernière mue transforme la larve du 4^e stade en nymphe.

La nymphe est aquatique, mobile et ne se nourrit pas. C'est un stade (environ 48 h) de transition au métabolisme extrêmement actif, au cours duquel l'anophèle subit de très profondes modifications morphologiques et physiologiques qui l'amènent du stade larvaire, aquatique et saprophyte, à la forme adulte, aérienne et hématophage chez la femelle. L'adulte

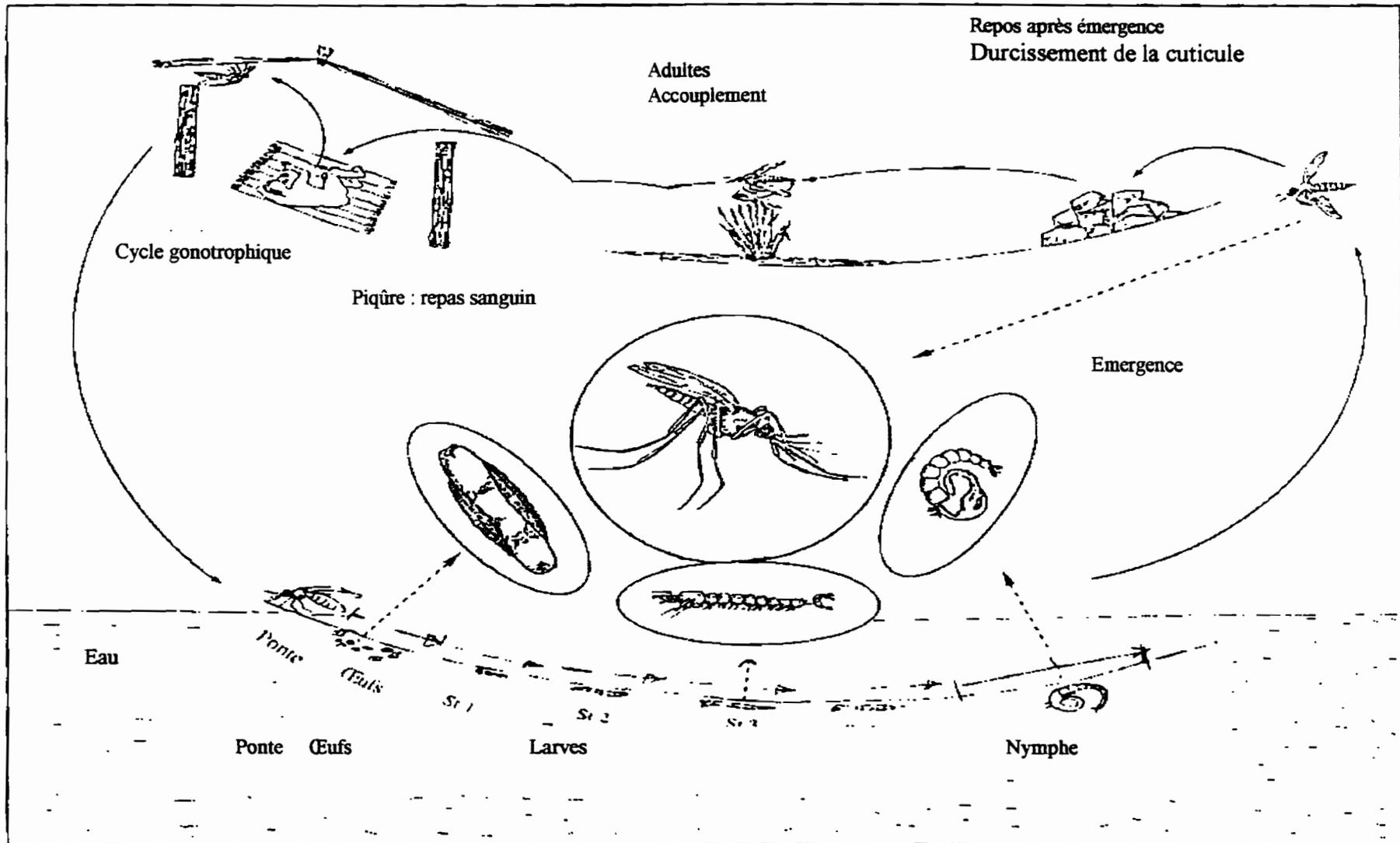


Figure 1 : Cycle de développement des anophèles (d'après Danis & Mouchet, 1991)

émerge de la nymphe (l'émergence des mâles ayant lieu en moyenne 24 h avant celle des femelles).

Chez les adultes, les mâles et les femelles se différencient quant à leur éthologie et leur écophysiologie. Les anophèles mâles ne sont pas hématophages et se nourrissent de sucs d'origine végétale. De ce fait, ils se déplacent assez peu de leur gîte d'émergence. Leur longévité est aussi relativement faible. Quant aux femelles, bien qu'absorbant des jus sucrés d'origine végétale, elles ont un régime hématophage. Les substances résultant de la digestion du sang absorbé sont utilisées en partie pour la nutrition, mais surtout pour assurer la maturation des œufs. Le régime hématophage amène les femelles à rechercher un hôte convenable d'où une plus grande dispersion à partir du gîte d'origine.

La faune anophélienne africaine est caractérisée par la présence d'excellents vecteurs du paludisme dont les principaux sont *An gambiae*, *An. arabiensis* et *An. funestus*. Partout sont présentes une ou plusieurs de ces espèces accompagnées parfois d'*An. nili* et *An. moucheti* qui sont d'autres bons vecteurs localisés. Cette ubiquité de bons vecteurs fait de l'Afrique un énorme foyer presque ininterrompu de paludisme, de l'océan Atlantique à l'océan Indien, du Sahara au Kalahari. Ces vecteurs et particulièrement les plus dangereux, sont associés étroitement à l'homme qui constitue leur principale source de repas de sang.

CHAPITRE II : Les études de génétique des populations

II.1- La génétique des populations

II.1.1- Définition

La génétique des populations a pour objet l'étude des forces génétiques, écologiques ou environnementales qui sont à l'origine de la variabilité génétique et de sa répartition entre et au sein des populations d'êtres vivants. Ces forces nombreuses et interactives, sont notamment la mutation, la recombinaison, la sélection, la migration, la distance géographique entre individus et le système de reproduction. Reconnaître et évaluer ces facteurs sont essentiels pour mieux comprendre et éventuellement prévoir l'évolution des populations et des espèces.

Ces études de génétique des populations font appel à de nombreux modèles mathématiques reposant sur un certain nombre de postulats qui permettront l'analyse des structures génétiques existantes.

II.1.2- Modèles mathématiques

Un modèle consiste en une simplification volontaire d'une situation complexe en vue d'éliminer les détails superflus pour s'intéresser seulement à l'essentiel de la question. En génétique des populations, on doit tenir compte de facteurs tels que la taille de la population, les modes de croisement, la répartition géographique des individus, les mutations, la migration et la sélection. Le but final est de comprendre les effets combinés de tous ces facteurs et d'autres encore, mais ceux-ci sont si nombreux et interagissent de façon si complexe qu'ils ne peuvent pas être pris en considération tous à la fois. Compte tenu de ceci, on a imaginé des situations plus simples dans laquelle il est possible d'identifier les facteurs les plus importants et de négliger les autres.

Un modèle mathématique consiste donc en un ensemble d'hypothèses décrivant les relations mathématiques entre des quantités mesurées ou mesurables (les paramètres) dans un système ou un processus. La description de ces modèles nécessite au préalable la définition de certaines notions clés qui figurent dans l'encadré 1.

II.1.2.1- Le modèle de Hardy-Weinberg

Ce modèle qui a été établi au début du siècle dernier (1908) par G.H. Hardy et W. Weinberg, est le plus simple et sert de base à la génétique des populations. Ce modèle décrit les relations entre fréquences alléliques et génotypiques dans une population idéale (population infinie au sein de laquelle les croisements sont aléatoires). Au nombre des hypothèses permettant de simplifier les calculs, nous pouvons citer :

- L'organisme considéré est diploïde et a une reproduction sexuée.
- Les générations sont non chevauchantes et les croisements se font au hasard (panmixie).
- La taille de la population est très grande et peut être considérée comme infinie (absence de dérive génétique).
- La fréquence des gènes n'est pas modifiée d'une génération à l'autre par les forces telles que la mutation, la sélection ou l'apport extérieur de nouveaux gènes (absence de migration).

Lorsque les croisements se font au hasard, les fréquences alléliques restent stables au cours des générations et les fréquences des génotypes se déduisent des fréquences alléliques par la relation :

Homozygotes $A_i A_i = p_i^2$

Hétérozygotes $A_i A_j = p_i p_j$

Où p_i et p_j sont les fréquences des allèles A_i et A_j respectivement dans la population. Cette population est dite en **équilibre de Hardy-Weinberg**.

Encadré 1

Définition de quelques notions importantes en génétique des populations

- **Génome** : Ensemble du matériel génétique présent dans un jeu de chromosomes.
- **Gène**: En génétique des populations, un gène est un fragment d'ADN non recombinant, qui peut être codant ou non.
- **Allèle**: Ce terme représente l'état d'un gène sur un chromosome. Chez les organismes diploïdes, un gène possède deux allèles, chacun sur l'un des deux chromosomes homologues. Si les deux copies sont identiques, l'individu est dit **homozygote**; si les deux copies sont différentes, l'individu est **hétérozygote**.
- **Hétérozygotie** : Mesure de la variation génétique dans une population, par rapport à un locus ; cela représente aussi la fréquence d'hétérozygotes à ce locus
- **Locus**: Emplacement d'un gène sur le chromosome (= localisation cytologique).
- **Hybride** : Descendant de tout croisement impliquant des parents de génotypes différents
- **Population**: ensemble d'individus de la même espèce entre lesquels les croisements sont libres (c'est la **panmixie**).
- **Sous-population** (synonymes: dème, population locale): subdivision de la population.
- **Taille effective** d'une population (Wright 1931, 1938): Taille (en nombre d'individus) d'une population idéale, non soumise à la migration, ayant les mêmes propriétés que la population naturelle étudiée. Le sex-ratio est 1:1, les croisements sont libres et tous les individus ont la même capacité de reproduction. De ce fait, la taille effective de la population est très souvent inférieure à la taille réelle de la population naturelle car elle correspond au nombre d'individus qui vont effectivement avoir une descendance.
Symbole conventionnel: N_e .
- **Dérive génétique**: C'est une force évolutive qui traduit l'échantillonnage aléatoire des allèles de la population parentale qui seront représentés dans la descendance. Elle est directement liée à la taille finie des populations naturelles: plus les populations seront de taille restreinte, plus l'influence de la dérive génétique sera importante. Elle agit en faveur de l'augmentation de la différenciation entre populations.
- **Flux de gènes**: Cette appellation regroupe toutes les formes de migration des gamètes, des individus ou de groupes d'individus d'une sous-population à une autre. Cette force s'oppose à la dérive génétique et favorise l'homogénéisation de la population.

En comparant les valeurs observées dans l'échantillon de la population qu'on étudie à celles attendues sous l'hypothèse de l'équilibre de Hardy-Weinberg, on peut estimer si les croisements se font au hasard ou non au sein d'une population donnée. Le test couramment employé en génétique de population est le test du chi-carré (X^2) qui a pour nombre de degré de liberté le nombre d'allèles -1 (nombre d'allèles indépendants) observé à un locus.

Une déviation significative à cet équilibre peut-être due au système de reproduction (autogamie, consanguinité, homogamie...), au mélange d'au moins deux sous-populations qui sont reproductivement isolées : C'est l'effet Wahlund qui se manifeste par un déficit en hétérozygotes au sein de l'échantillon étudié.

II.1.2.2- Les modèles démographiques

Deux grandes catégories sont définies pour les modèles décrivant la structure des populations : les modèles de populations composées de sous-populations discrètes et ceux de populations distribuées de façon continue. L'étude de l'équilibre entre la dérive génétique et les flux de gènes est au centre de ces modèles. Les populations sont supposées persister suffisamment longtemps pour que l'équilibre soit atteint. Quelques-uns de ces modèles sont définis ci-dessous.

II.1.2.2.1- Les modèles en îles

i- Nombre infini d'îles

Ce modèle qui fut décrit par Wright (1931), est caractérisé par une population d'individus diploïdes à générations discrètes, subdivisée en un nombre infini d'îles de taille finie dans lesquelles les croisements entre individus sont aléatoires. Chaque île échange des migrants avec toutes les autres. Les fréquences alléliques ne changent pas sur l'ensemble de la population. Mais chaque île est soumise à la dérive génétique et à la migration.

Wright en utilisant le modèle infini d'îles, a montré que si une fraction m d'une population est remplacée par des immigrants, il n'y aura aucune différenciation génétique significative si $Nm > 1$ où N est la taille effective de la population. Ces résultats ont amené Wright à stipuler que l'échange d'un seul individu par génération entre deux sous-populations est suffisant pour prévenir la différenciation due à la dérive génétique et maintenir ainsi l'homogénéité de la variabilité génétique au sein de la population.

ii- Nombre fini d'îles ou modèles à n îles (Latter 1973 ; Slatkin 1985)

Les éléments de base de ce modèle sont identiques au modèle précédent à la seule grande différence que le nombre d'îles est fini et que toutes les populations ont le même effectif

(N). Dans ce cas, la population globale est elle aussi soumise à la dérive génétique et les fréquences alléliques peuvent varier d'une génération à l'autre.

Il n'y a aucune géographie spécifique entre les îles pour chacun de ces modèles, car chaque sous-population est ouverte à l'arrivée de migrants de chacune des autres îles. Dans ces modèles en îles, les effets de la dérive sont d'autant plus marqués que le taux de migration m et l'effectif efficace N_e seront petits.

II.1.2.2.2- Les modèles dits "Stepping stone" (Kimura, 1953)

Ces modèles font référence à un système géographique bien structuré au sein duquel les sous-populations échangent réciproquement des migrants uniquement avec leurs voisins immédiats. Plusieurs variants de ces modèles existent à une, deux ou trois dimensions en fonction de la disposition des sous-populations dans l'espace.

II.1.2.2.3- Les modèles de voisinage ou isolement par la distance (Wright, 1943)

Contrairement aux modèles précédents, aucune subdivision de la population n'est supposée et les individus sont librement dispersés dans l'espace de manière continue. Ce modèle suppose que la dispersion des individus, ou plus généralement des gènes est une fonction décroissante de la distance. Le nombre de migrants échangés entre sous-populations n'est plus le paramètre clé mais plutôt la taille du voisinage.

II.1.2.2.4- Le modèle de populations subdivisées et transitoires ou métapopulations

Le terme de métapopulation désigne une population composée de petites populations locales ou demeures, qui sont établies par des individus colonisateurs, qui survivent pendant un certain temps, envoient des migrants et, éventuellement, disparaissent. La persistance d'une population dans une région donnée dépend de la balance entre le taux de colonisation et le taux d'extinction. Du point de vue génétique, les questions posées sont relatives à la distribution de la variabilité génétique : quelles sont les relations entre les phénomènes d'extinction et colonisation et les flux de gènes? Comment se répartit la variabilité entre et au sein des populations? Les études menées par de nombreux auteurs (Wright, 1931 et 1940 ; Slatkin, 1977...) sur la base de modèles en nombre fini d'îles, ont abouti à la conclusion selon laquelle la dynamique d'extinction et de colonisation des populations locales peut conduire à une augmentation ou à une diminution de la différenciation génétique en fonction des paramètres suivants : le nombre k d'individus établissant une nouvelle population, le nombre N d'individus constituant chacune des populations locales à l'équilibre, le taux m de migration et le coefficient Φ qui décrit l'origine des migrants (qui peuvent provenir d'une seule population ou de l'ensemble). Si tous les migrants sont issus d'une seule population parentale, les phénomènes d'extinction et de recolonisation renforcent la structure génétique des populations si $k < N$.

Mais si chaque population contribue de façon équivalente à la migration, la structure génétique peut se trouver accentuée ou réduite selon que la composante de la dérive génétique (nombre k de colonisateurs) prendra le pas ou non sur la composante liée au flux de gènes (origine des colonisateurs).

La variation de taille des populations au cours du temps affecte aussi la composition génétique des populations. Si en dehors de toute migration, la population subit un goulot d'étranglement démographique (une réduction d'effectif ou «bottleneck» en anglais), les effets de la dérive génétique seront accentués et à l'origine d'un goulot d'étranglement génétique caractérisé dans un premier temps par une diminution de la diversité génétique et de l'hétérozygotie (Nei *et al.*, 1975 ; Maruyama & Fuerst 1984 et 1985 ; Tajima 1989).

Ces modèles nous offrent un certain nombre de prédictions et d'attendus théoriques qu'il est possible de confronter avec les données issues des expériences biologiques. Ils sont à l'origine d'outils d'analyse de données de la structure génétique des populations.

II.1.3.- Approches expérimentales et traitement des données

Deux méthodes peuvent être utilisées dans le but d'étudier la structure génétique des populations et plus particulièrement les flux de gènes.

II.1.3.1- Méthode directe : étude de la dispersion

L'observation directe de la dispersion des individus à différents stades de vie ou les expériences de capture-marquage-recapture permettent d'estimer les flux de gènes entre populations. Cette méthode a été appliquée à l'étude de la dispersion des moustiques, particulièrement chez *Ae. aegypti* et quelques expériences ont été menées avec *An. gambiae* (Service, 1993 ; costantini *et al.*, 1996). Les moustiques sont marqués par des poudres fluorescentes. Grâce à des modèles spécifiques, on estime la capacité de dispersion des individus, leur espérance de vie (taux de mortalité journalier), la densité de la population et la taille efficace (Fisher & Ford, 1947 ; Service, 1993). Cette méthode fournit des informations très importantes sur la capacité vectorielle des populations, leur biologie et leur importance sur le plan épidémiologique. Ces mesures directes estiment la dispersion des individus uniquement sous les conditions dans lesquelles elles sont réalisées et ne fournissent donc pas une estimation moyenne. Ces mesures sont limitées dans le temps et dans l'espace et indiquent les dispersions à un moment particulier, et non à travers une longue échelle de temps où une variété d'événements auraient pu survenir mais pas au moment de l'observation. Une autre limite de la méthode est le fait que la capacité de dispersion ne reflète pas toujours les flux de gènes car n'étant pas automatiquement équivalent au succès d'établissement ou de

reproduction des individus ayant migré. Chez les moustiques, le taux de recapture est généralement faible (1 à 5%) et cette méthode ne peut pas être appliquée en saison sèche puisque les densités de populations sont très faibles (Touré *et al.*, 1998 ; Costantini *et al.*, 1996).

II.1.3.2- Méthodes indirectes en génétique

Ce sont des méthodes qui utilisent la distribution géographique des fréquences alléliques obtenues à l'aide de plusieurs marqueurs notamment isoenzymatiques mais de plus en plus des marqueurs moléculaires ciblant directement le polymorphisme de l'ADN. Parmi ces marqueurs moléculaires, on peut citer les RAPD, le RFLP, les microsatellites ou des séquences d'ADN. Des segments chromosomiques ou des traits phénotypiques, sont aussi utilisés pour estimer le taux et les schémas de flux de gènes entre populations. Il y a une différence fondamentale entre cette méthode et la précédente car on se place ici en amont des flux de gènes dans le but d'estimer le taux de flux génétiques qui ont dû se produire pour justifier le profil de répartition de la variabilité génétique observée.

Cette approche nécessite des modèles mathématiques afin de rapprocher l'observation brute (comme les génotypes) des conséquences en termes de structure génétique ou de flux de gènes. Parmi ces modèles, la méthode des statistiques de Wright est la plus utilisée. Elle estime la variance standardisée des fréquences alléliques parmi les populations locales. Wright a en effet mis au point un ensemble d'outils mathématiques, les statistiques F, permettant de décrire la répartition de la variabilité génétique au sein et entre les populations. Au sein de la population subdivisée on distingue trois niveaux de complexité : les individus (I), les sous-populations (S) et la population totale (T). A chacun de ces niveaux correspond un coefficient :

- **Fis** qui mesure les déviations par rapport aux proportions attendues sous l'hypothèse de l'équilibre de Hardy-Weinberg au sein de chaque sous-population. Il est aussi appelé coefficient de consanguinité des individus dans la sous-population à laquelle ils appartiennent.
- **Fst** encore appelé indice de fixation, permet de mesurer la différenciation génétique entre sous-populations.
- **Fit** mesure la déviation par rapport aux proportions attendues sous l'équilibre de Hardy-Weinberg sur l'ensemble des populations.

Ces trois paramètres sont reliés par l'équation :

$$(1-Fit) = (1-Fis)(1-Fst)$$

Différentes méthodes ont été développées pour estimer ces paramètres (le paramètre G_{st} , estimateur de F_{st} de Nei (1973), les paramètres f , F et θ de Weir & Cockerham (1984)). Les estimateurs de Weir & Cockerham dépendent de la taille des populations et de leur histoire, mais sont indépendants de la stratégie d'échantillonnage (nombre d'individus par population, nombre de populations, nombre d'allèles par locus). Pour des allèles neutres :

$$F_{st} = 1/(1+4Nm)$$

où N est la taille de la population locale et m le taux moyen de migrants dans un modèle en îles. La dérive génétique entraîne une différenciation génétique locale significative si $Nm < 1$. Toutes ces estimations sont faites sous l'hypothèse d'un équilibre mutation/dérive, d'un modèle en nombre infini d'îles, d'un taux de mutation négligeable par rapport à la taille efficace N_e .

Comme autre méthode d'estimation des flux géniques, nous citerons celle des allèles privés (ou rare) de Slatkin (1985b) qui dépend de la fréquence des allèles rares présents dans une ou très peu de populations locales. Ces allèles apparus par mutation varient en fréquence sous l'action de la dérive génétique. Ils peuvent perdre leur état d'allèles privés s'ils sont introduits dans d'autres populations par migration. Pour les modèles en îles et de «stepping stone», il existe une corrélation linéaire entre le logarithme de la fréquence moyenne des allèles privés et le logarithme du flux de gènes (Nm). Cette méthode ne peut être utilisée que s'il existe de tels allèles dans les populations. Elle n'est pas adéquate si le taux de mutation est élevé.

D'autres méthodes existent aussi pour l'estimation des flux géniques, tels que les analyses d'auto-corrélations spatiales (Slatkin & Arter, 1991 ; Epperson, 1993), ou la méthode du maximum de vraisemblance (Rannala & Hartigan, 1996). Le paramètre de la distance génétique D est aussi souvent utilisé pour estimer la différenciation génétique entre sous-populations (Nei, 1987) et il est basé sur l'analyse de la covariance des fréquences alléliques. Wang *et al.* (2001) ont défini de nouveaux paramètres D' et D_k à partir de ce paramètre D pour estimer le taux effectif de migration Nm entre populations et détecter ainsi les flux de gènes. Le paramètre R_{st} quant à lui, utilise la variance de la taille des allèles pour estimer la différenciation génétique (Slatkin, 1995) et suppose un modèle de SMM (Stepwise Mutation Model) pour le processus de mutation du marqueur. Nous pouvons encore citer le paramètre $(\delta\mu)^2$ qui est le carré de la différence entre la taille moyenne des allèles (Goldstein *et al.*, 1995b) et qui sert également d'estimateur de la différenciation génétique entre populations.

Les analyses indirectes de flux de gènes et de structure des populations se font à l'aide de données (génotypes, allèles) obtenues à partir d'outils informationnels tels que les marqueurs moléculaires.

II.2- Les marqueurs moléculaires

II.2.1- Définition et propriétés générales

Un marqueur moléculaire est un fragment d'ADN (information génétique) ou son expression moléculaire (ARN, protéines) permettant la caractérisation génétique d'un individu ou d'une population. Ces marqueurs ont plusieurs domaines d'utilisation entre autres les études d'apparement, les études de systématique (la définition des espèces jumelles), les études phylogénétiques (relations entre espèces) mais surtout les études de structure génétique des populations (estimation des effets de la dérive génétique, des flux de gènes) (Raymond, 1996).

Les techniques d'électrophorèse des protéines ont été les premières mises au point dès les années 50. Mais ce n'est qu'en 1966 que leur potentiel a été révélé (par des travaux menés chez l'homme et la drosophile). Depuis lors, de milliers d'espèces ont été étudiées.

Les techniques d'étude de l'ADN ont progressivement été mises en place dans les années 70 (extraction d'ADN, production en masse d'enzymes de restriction, clonage et séquençage). Ainsi le polymorphisme de l'ADN a été étudié à l'aide de la technique de RFLP. Mais la véritable révolution vient de la mise au point de la technique de PCR en 1985 par K. Mullis. Cette technique est un moyen efficace pour augmenter la quantité initiale de matériel (ADN). Elle a autorisé l'analyse ciblée de l'ADN, et a donné naissance à un grand nombre de marqueurs moléculaires ou de techniques associées (RAPDs, AFLPs, microsatellites, SSCP...). Ici on ne s'intéresse plus à des enzymes, produits de l'expression des gènes mais à des séquences d'ADN.

II.2.2- Les ADN microsatellites

II.2.2.1- Caractéristiques générales

Les microsatellites sont des séquences d'ADN répétées en tandem dont l'unité de répétition fait de deux à six paires de bases (pb). La limite supérieure de six pb n'est nullement absolue, mais au-delà de cette limite on parle plutôt de minisatellites. Le nom de ce type de marqueur a donné lieu à plusieurs variations. Si on parle généralement de microsatellites, des dénominations telles que SSR («Simple Sequence Repeats»), STR («Short Tandem Repeats»), TR («tandem repeats») ou VNTR («Variable Number of Tandem Repeats»), peuvent encore être trouvées dans la littérature (Tautz, 1993). L'origine du nom est liée à l'appartenance des

microsatellites à la famille des séquences répétées en tandem, qui comprend aussi les satellites et les minisatellites. Les satellites ont été nommés ainsi parce qu'ils apparaissaient comme des bandes satellites lors des centrifugations en gradient de chlorure de césium d'ADN total. Par dérivation sont apparus les noms minisatellites (Jeffreys *et al.*, 1985) et microsatellites (Litt & Luty, 1989).

L'existence des microsatellites dans les génomes est connue depuis une vingtaine d'années, mais ils ont été longtemps considérés comme des séquences de peu d'intérêt. Ce n'est qu'à la fin des années 1980 qu'on a réalisé qu'il s'agissait probablement d'un des plus puissants marqueurs mendéliens qu'on ait mis en évidence jusqu'à présent. L'utilisation de ces marqueurs pour générer des empreintes génétiques apparaît vers 1985, mais la première véritable étude de structure de populations utilisant les microsatellites est celle de Edwards *et al.* (1992) sur des populations humaines.

Leur fonction biologique reste floue (Epplen *et al.*, 1993). Il se pourrait qu'ils jouent le rôle d'éléments codant et de régulation (Kunzler *et al.*, 1995 ; Kashi *et al.*, 1997). Comme séquences régulatrices, les microsatellites sont rencontrés de façon ubiquiste en amont des régions du promoteur des séquences codantes. Ils ont également été trouvés dans de nombreuses régions codant pour des protéines. Certaines études récentes montrent que les variations de longueur de certains microsatellites ont des effets phénotypiques sur la physiologie et le développement de certains organismes (king & Soller, 1998). En outre, certaines séquences de type (CA)*n* pourraient constituer des points chauds de recombinaison lors de la transcription ou intervenir dans la conformation de l'ADN (organisation de la chromatine). Les séquences de type (CTG)*n* sont quelques fois associées à des maladies génétiques humaines (Nordheim & Rich, 1983 ; Pardue *et al.*, 1987 ; Brahmachari *et al.*, 1995). La densité des microsatellites est plus importante dans les régions non codantes que dans les régions codantes (rapport de 1/15 chez *S. cerevisiae* (Field & Wills, 1998)).

II.2.2.2- Propriétés

Quelques propriétés importantes des microsatellites comme marqueurs pour les études de génétique de populations méritent d'être relevées :

i-Mise en évidence aisée des allèles

L'obtention du génotype de chaque individu se fait de manière directe grâce à la désignation d'amorces spécifiques dans les régions flanquantes de chaque locus, à l'amplification par PCR, à une électrophorèse sur acrylamide suivie de la révélation (Queller *et al.*, 1993).

Encadré 2

Les marqueurs microsatellites

(d'après Viard, 1996; Jarne & Lagoda, 1996)

Définition:

Séquences d'ADN répétées en tandem d'un motif de deux à six paires de bases.

5' GCCACGATGGTTTCG.....GTGTGTGTGTGT.....CGTCCAAGTACGATC 3'

séquence flanquante

6 répétitions

séquence flanquante

Les amorces de PCR correspondant à des séquences uniques (soulignées) sont déterminées dans les régions adjacentes aux répétitions (séquences flanquantes).

Structure:

Les marqueurs microsatellites peuvent être classés en trois familles selon leur structure interne.

Purs (parfaits)	GTGTGTGTGTGTGTGTGTGTGT
Composés	GTGTGTGTGTCACACACACA
Interrompus	GTGTGTTTGTGTTTGTTTGT

Base de la variabilité:

Les séquences microsatellites présentent un polymorphisme de longueur, dû au nombre variable de répétitions qui les composent. Chaque allèle correspond à un nombre de répétitions (en théorie) et à une taille d'amplifiat (en pratique).

Dans l'exemple ci-dessous, les séquences flanquantes font 80 paires de base et le motif répété est un dinucléotide (2 pb):

Nombre de répétitions	Fragment amplifié	Taille du fragment
12	$\rho\rho\rho\rho\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\rho\rho\rho\rho$	$12 \times 2 + 80 = 104$ pb
11	$\rho\rho\rho\rho\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\rho\rho\rho\rho$	$11 \times 2 + 80 = 102$ pb
10	$\rho\rho\rho\rho\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\rho\rho\rho\rho$	$10 \times 2 + 80 = 100$ pb
9	$\rho\rho\rho\rho\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\textcircled{R}\rho\rho\rho\rho$	$9 \times 2 + 80 = 98$ pb

ρ = séquences flanquantes \textcircled{R} = motif répété (2pb)

ii-Polymorphisme élevé

Le taux de mutation des microsatellites est généralement très élevé chez la plupart des espèces étudiées. Il est de l'ordre de 10^{-2} par locus/réplication chez *E. Coli* (Levinson & Gutman, 1987), 10^{-3} chez l'homme (Weber & Wong, 1993) et entre 10^{-3} et 10^{-4} chez la souris (Dallas, 1992). Cette dernière valeur est près de 1000 fois supérieure à celles connues pour les isoenzymes (Ayala, 1976). Ce taux de mutation est fonction de la longueur de la séquence puisque les locus microsatellites ayant un nombre de motifs répétés élevé ont un taux de mutation généralement plus élevé que les locus ayant un faible nombre de motifs répétés (Ellegren, 2000). Ces forts taux de mutation rendent les microsatellites très polymorphes et donc très résolutifs pour les études de génétique des populations. Les mécanismes moléculaires à l'origine de cet important polymorphisme sont classés dans deux types de modèle. Le premier modèle est le SSM ou «Slipped Strand Mismatching» qui est caractérisé par le glissement d'un brin d'ADN sur l'autre lors de la réplication ou en dehors de celle-ci (Levinson & Gutman, 1987 ; Tautz & Schötterer, 1994). Ceci survient à cause du dérapage de l'ADN polymérase (Slippage) dû à des problèmes de translocation de l'enzyme sur la matrice à recopier. Suivant le sens du décalage induit par ce glissement, il y a insertion ou délétion d'une unité de répétition au niveau du locus répété. Le second modèle concerne la recombinaison entre molécules d'ADN (Smith, 1976 ; Jeffreys *et al.*, 1994). La recombinaison interchromosomique peut potentiellement altérer la taille des microsatellites de deux manières notamment par les crossing-over inégaux ou par la conversion génique (processus dans lequel l'appariement des nucléotides entre deux gènes suffisamment homologues s'accompagne de l'excision de tout ou une partie de la séquence des nucléotides de l'autre gène) (Ellegren, 2000).

iii- Conservation interspécifique des locus microsatellite

La conservation de certains locus microsatellites d'une espèce à une autre a été démontrée dans des groupes aussi variés que les cétacés (Schlötterer *et al.*, 1991), les oiseaux (Primmer *et al.*, 1996a), les primates (Deka *et al.*, 1994), les tortues marines (FritzSimmons *et al.*, 1995) ou les mammifères en général (Stallings, 1995) ce qui peut permettre l'utilisation d'amorces hétérologues pour des études comparatives de la variabilité génétique entre espèces proches.

iv-Modèles de mutation

Deux modèles simples, développés dans les années 60 et 70 afin d'interpréter les données obtenues à partir de marqueurs protéiques (isoenzymes), ont d'abord été considérés pour les microsatellites : le modèle IAM (Infinite Allele Model) (Kimura & Crow, 1964) et le

modèle SMM (Stepwise Mutation Model) (Kimura & Ohta, 1978). Dans le modèle IAM, chaque mutation va générer un allèle différent de ceux existants dans la population. Le nouvel allèle peut avoir n'importe quelle taille. En revanche, les possibilités sont plus restreintes dans le modèle SMM où une mutation diminue ou augmente, avec la même probabilité, le nombre de répétitions d'une unité. Afin de prendre en compte des limites éventuelles sur la taille des variants alléliques, un autre modèle a été considéré pour les microsatellites. Il s'agit du modèle à k-allèles (KAM ou «K Allèle Model», Crow & Kimura, 1970), qui présuppose l'existence de k états alléliques possibles avec une probabilité constante pour chaque allèle [$\mu/(k-1)$] de muter vers un des k-1 états alléliques disponibles. De nombreuses études ont montré que si les mutations au niveau des locus microsatellites correspondaient majoritairement à des changements d'un motif répété, une proportion non négligeable de mutations correspondait à un changement du nombre de motifs répétés supérieurs à un. En tenant compte de cette caractéristique, Di Rienzo *et al.* (1994) ont proposé le modèle à deux phases (TPM ou Two-Phase Model). Dans ce modèle, une proportion p de mutations suit un SMM et une proportion 1-p se traduit par des ajouts ou des délétions de plusieurs motifs. Au contraire de l'IAM, les modèles SMM et TPM tiennent donc compte de la relation entre allèles en fonction de leur taille : la taille d'un allèle mutant est liée à la taille de l'allèle qui a muté, et deux allèles de même taille n (répétitions) peuvent avoir été générés par des processus différents.

Ces modèles ont généré le développement de nouvelles méthodes mathématiques d'analyse des données spécifiques aux locus microsatellites qui tiennent compte de leur fort taux de mutation et de l'inter-dépendance de la taille des allèles (Wang *et al.*, 2001).

II.2.2.3- Contraintes associées aux microsatellites

Des contraintes inhérentes aux microsatellites sont courantes et peuvent se manifester à plusieurs niveaux et induire des interprétations erronées en terme de flux géniques et de structure de populations. Parmi ces contraintes, on peut citer :

- Les contraintes agissant sur la taille des allèles

Les répétitions d'un allèle microsatellite ne peuvent pas s'étendre infiniment et par conséquent le nombre d'allèles à un locus est limité. Cette observation a été confirmée aussi bien par des données expérimentales (Bowcock *et al.*, 1994 ; Deka *et al.*, 1994 ; Garza *et al.*, 1995) que par des simulations sur ordinateur (Nauta & Weissing, 1996 ; Lehmann *et al.*, 1996). Etant donné le fort taux de mutation observé chez les microsatellites, ces contraintes pourraient introduire un biais dans l'estimation des flux de gènes en surestimant l'identité des

allèles entre différentes populations. Cette situation pourrait entraîner les locus microsatellites à être peu résolutifs en ce qui concerne la structure des populations de grande taille ou la reconstruction des phylogénies d'espèces ayant divergées depuis longtemps.

- Le phénomène d'homoplasie

Il est possible que certains allèles microsatellites identiques par leur taille ne le soient pas par ascendance car résultant de processus évolutifs différents (encadré 3). Ces excès de similarité entre allèles peuvent se manifester par des différences (insertion, substitution, délétion) dans les séquences flanquantes ou au niveau des interruptions dans le cas des microsatellites interrompus, ce qui ne sera révélé que par le séquençage des différents allèles. Toutefois, ce phénomène s'avère être plus commun entre espèces éloignées qu'entre espèces phylogénétiquement proches (Estoup *et al.*, 1995b). Les microsatellites peuvent donc s'avérer peu recommandables pour l'étude de phylogénies interspécifiques (Garza *et al.*, 1995 ; Doyle *et al.*, 1998)

- L'hétérogénéité du taux de mutation

De nombreuses observations montrent que les locus microsatellites ne présentent pas tous les mêmes taux de mutation. On a ainsi pu mettre en évidence des différences de taux de mutation entre espèces pour le même locus (Rubinsztein *et al.*, 1995) et en fonction de la nature du locus (pur, composé ou interrompu) et du motif (di-, tri-, ou tétranucléotide), des différences s'observent au sein d'une même espèce (Chung *et al.*, 1993 ; Pépin *et al.*, 1995 ; Brinkmann *et al.*, 1998). Ce taux de mutation varie aussi en fonction de la taille des allèles, les plus longs étant plus exposés à la mutation (Wierdl *et al.*, 1997 ; Schlötterer *et al.*, 1998). Toutes ces données mettent en avant le fait que le mécanisme moléculaire d'évolution des locus microsatellites est encore mal compris et ce fait doit être pris en compte lors des différentes études utilisant ces marqueurs (Kruglyak *et al.*, 1998 ; Colson & Golstein, 1999).

- La présence d'allèles nuls (non détectables sur le gel)

L'existence des allèles nuls constitue l'une des contraintes les plus importantes et la plus répandue pour les microsatellites. Ces allèles nuls sont dus à des mutations ponctuelles qui rendent impossible l'appariement des amorces des locus concernés durant la PCR (Callen *et al.*, 1993 ; Pemberton *et al.*, 1995). Des individus hétérozygotes apparaissent comme homozygotes et les individus homozygotes ne donnent aucun amplifiat. Leur mise en évidence se fait par séquençage des zones d'appariement des amorces (Callen *et al.*, 1993), par des confrontations de génotypes parent-descendants (Paetkau & Strobeck, 1995) ou par la

Encadré 3

Exemples de mise en évidence de phénomènes d'homoplasie

Variations dans la nature et la place des interruptions (Estoup *et al.*, 1995b)

Allèles observés au locus A113 chez différentes sous-espèces d'*Apis mellifera* (*A. m.*).

Espèce	Séquence de l'allèle (5' → 3')	Taille
<i>A. m. mellifera</i>	(TC) ₂ C(TC) ₂ TT(TC) ₄ TT(TC) ₂ TT(TC) ₉ GTTTCG(TC) ₂	214 pb
<i>A. m. ligustica</i>	(TC) ₂ C(TC) ₂ TT(TC) ₄ TT(TC) ₂ TT(TC) ₉ GTTTCG(TC) ₂	214 pb
<i>A. m. scutellata</i>	(TC) ₂ C(TC) ₂ TT(TC) ₅ TT(TC) ₁₁ GTTTCG(TC) ₂	214 pb
<i>A. m. capensis</i>	(TC) ₂ C(TC) ₂ TT(TC) ₅ TTTCTT(TC) ₃ TT(TC) ₁₀ GTTTCG(TC) ₂	214 pb

Ces allèles sont tous identiques par état (même taille du fragment amplifié) mais l'analyse de leur composition montre des différences substantielles révélant qu'ils ne sont pas identiques par ascendance. La position et la nature des interruptions peuvent être très informatifs dans ce cas.

Nombre variable de répétitions au niveau d'un locus composé (Garza & Freimer, 1996)

Allèles observés au locus Mfd79 chez le chimpanzé *Pan troglodytes*:

	Séquence de l'allèle (5' → 3')	Taille
Allèle 1	(CA) ₁ (TA) ₂₀ (CA) ₁₂	172 pb
Allèle 2	(CA) ₁ (TA) ₂₁ (CA) ₁₁	172 pb

Le nombre de répétitions des différents motifs constituant ce locus sont variables et leur différences peuvent s'annuler, induisant de l'homoplasie.

Variations dans les séquences flanquantes (Grimaldi & Crouau-Roy, 1997)

Allèles observés au locus MIB chez l'homme:

	Séquence de l'allèle (5' → 3')	Taille
Allèle 1	...GTTATCC(CA) ₂₃ TAAA - - - - - CATAAG...	350 pb
Allèle 2	...GTTATCC(CA) ₂₀ TAAACATAAACATAAG...	350 pb

Des délétions ou insertions dans les séquences flanquantes peuvent également permettre la mise en évidence de phénomènes d'homoplasie, y compris dans le cas de locus purs.

désignation de nouvelles amorces («alternative primers») qui évitent la mutation (Lehmann *et al.*, 1996). La présence de ce type d'allèles peut induire une sous-estimation de l'hétérozygotie. Une estimation de la fréquence des allèles nuls peut être faite sous l'hypothèse de proportions génotypiques conformes à celles attendues sous l'équilibre d'Hardy-Weinberg (Brookfield, 1996). Chakraborty *et al.* (1992) estiment la fréquence de l'allèle nul à :

$$r = (H_e - H_o) / (H_e + H_o)$$

H_e étant la fréquence attendue d'hétérozygotes et H_o celle observée.

La fréquence de ces allèles nuls au niveau de certains locus, a été estimée à environ 15% par Jarne & Lagoda (1996), et à peu près 25% des locus microsatellites seraient concernés. On a aussi mis en évidence l'existence de tels allèles avec les isoenzymes (Foltz, 1986). Dans ce cas l'allèle nul ne possède pas l'activité enzymatique permettant sa révélation.

Lors des études de structuration génétique des populations utilisant les microsatellites, toutes les contraintes ou limites mentionnées ci-dessus doivent minutieusement être prises en compte afin de réduire tous les biais éventuels qu'elles pourraient induire. Toutefois il convient de noter que la plupart de ces contraintes sont spécifiques de locus et ne sauraient toucher tous les locus en même temps et avec la même influence. Par contre, les forces qui régissent la structure des populations telles que la dérive, la sélection et la migration, s'exercent de manière uniforme sur tout le génome. Ceci impose de prêter une attention particulière aux informations fournies par chaque locus et ensuite à celles issues de l'ensemble des locus.

II.2.2.4- Applications

Grâce à leur ubiquité, leur abondance sur le génome et leur fort taux de mutation qui leur confère un niveau de polymorphisme élevé, les microsatellites sont aujourd'hui couramment utilisés dans de nombreux domaines d'études :

- La biologie et la génétique des populations (Bossart & Pashley-Powell, 1998),
- La cartographie du génome (Bell & Ecker, 1994 ; Zheng *et al.*, 1996) ainsi que son application dans l'analyse des traits quantitatifs (QTL, Gorman *et al.*, 1997),
- La génétique humaine, en particulier comme marqueurs des maladies génétiques (Sutherlands & Richards, 1995), et de plus en plus pour l'étude de certaines phylogénies et de l'histoire de l'humanité (Bowcock *et al.*, 1994 ; Goldstein *et al.*, 1995b).

Nous présentons maintenant en bref de façon non exhaustive, d'autres marqueurs génétique souvent utilisés pour des études de génétique de populations.

II.2.3- Autres marqueurs ou techniques utilisés en génétique des populations

En dehors des microsatellites, plusieurs marqueurs techniques sont également utilisés en génétique des populations. Nous pouvons citer entre autres les isoenzymes, le polymorphisme des sites de restriction (RFLP), les RAPD et le SSCP.

Les isoenzymes sont les seuls marqueurs d'utilisation courante s'appuyant sur le polymorphisme protéique. Son principe est basé sur la séparation des allèles d'un même gène par électrophorèse dans un support de type gel (Amidon, acétate de cellulose) et après une coloration spécifique. Il s'agit donc d'un marqueur codominant. Les génotypes individuels obtenus permettent de déterminer la distance génétique de Nei (Nei, 1987) entre les populations. Ce marqueur présente l'inconvénient d'être très souvent peu polymorphe.

La technique de RFLP (Restriction Fragment Length Polymorphism) permet l'analyse du polymorphisme des sites de restriction dans des fragments d'ADN pour lesquels on dispose des sondes. Elle nécessite la coupure de l'ADN par des enzymes de restriction, une électrophorèse, un transfert de l'ADN sur membrane par «Southern Blot», une hybridation et une révélation. Des formes plus élaborées font appel à la PCR. Ici le polymorphisme est étudié sur un fragment préalablement amplifié par PCR. On parle alors de PCR-RFLP.

La technique de RAPD est basée sur le polymorphisme de courtes amorces (autour de 10 paires de bases) pour des amplifications aléatoires de séquences cibles dans un génome donné.

La technique de SSCP (Single-Strand conformation polymorphism) quant à elle permet de différencier des séquences d'ADN ne différant que par une seule paire de bases. En effet, les formes simple-brin de l'ADN prennent des conformations secondaires dépendant de leur séquence, et migreront en conséquence dans un support approprié. Il est ainsi possible de détecter des variations d'une paire de bases ou plus.

Les techniques de séquençage complète de certains gènes sont de plus en plus utilisées en génétique des populations. Dans ce cadre, c'est surtout l'ADN mitochondrial (ADNmt) qui est utilisé en raison de quelques-unes de ses propriétés remarquables (petite taille, facilité d'extraction, transmission uniparentale sans recombinaison, variabilité importante). L'ADNmt permet en effet de révéler des structurations géographiques des populations là où d'autres marqueurs génétiques s'avèrent inefficaces.

II.3- Espèce et spéciation

La notion d'espèce a toujours fait l'objet de beaucoup de débats au sein de la communauté scientifique aussi bien pour l'origine, la définition que pour la description. La définition classique de l'espèce biologique est celle de Mayr (1942) : «les espèces sont des groupes de populations réellement ou potentiellement capables de se croiser et qui sont reproductiblement isolés des autres

groupes ayant les mêmes propriétés». Celle-ci revient à considérer l'espèce comme un ensemble de gènes qui se transmettent d'une génération à l'autre, protégé des pools géniques des autres espèces par l'isolement reproductif. Ainsi cette notion d'espèce ne s'arrête pas sur des critères morphologiques puisque des spécimens morphologiquement indifférenciables peuvent être reproductivement isolés et appartenir à des espèces distinctes que l'on appelle «espèces jumelles». L'ensemble des espèces jumelles résultant du démembrement d'une espèce morphologique est appelé complexe d'espèces. La spéciation quant à elle est la formation à partir de l'ensemble génétique fermé qu'est l'espèce, de deux ensembles n'ayant plus aucun échange génique, c'est à dire de deux espèces.

Les phénomènes de spéciation surviennent sous l'impulsion de mécanismes qui, bloquant les flux géniques entre deux communautés, donnent à celles-ci le statut d'espèce. Parmi ces mécanismes, on peut citer :

- Les mécanismes d'isolement prézygotique qui regroupent l'ensemble des obstacles empêchant l'accouplement ou la fécondation entre deux espèces. Dans cette catégorie, nous pouvons mentionner l'isolement écologique (les partenaires possibles ne se rencontrent pas), l'isolement sexuel (les partenaires se rencontrent mais ne s'accouplent pas), l'isolement mécanique (l'accouplement a lieu mais le sperme n'est pas transféré) ou l'isolement gamétique (le sperme est transféré mais la fécondation n'a pas lieu).

- Les mécanismes d'isolement post-zygotique qui agissent après l'accouplement et la fécondation pour interrompre le développement ultérieur de l'individu en formation. Ces facteurs induisent la mortalité, la faible vigueur, la stérilité ou la dépression des hybrides.

On distingue en général trois formes de spéciation :

II.3.1- La spéciation géographique ou allopatrique

D'après ce modèle, une nouvelle espèce se développe lorsqu'une population isolée géographiquement de l'espèce parentale, a durant cette période, acquis les caractères susceptibles d'assurer son isolement reproductif avec l'espèce parentale même quand les barrières extérieures qui les séparaient ont disparu (Mayr, 1942). Ici la spéciation apparaît donc comme la résultante de la différenciation génétique. Ce modèle qui est le plus commun dans la plupart de groupes sexués, comprend deux types : la spéciation géographique classique (vicariante) et la spéciation à effet fondateur (peripatrique).

Dans le modèle de spéciation vicariante, les populations de l'espèce sont séparées en deux sous-ensembles importants qui se différencient sous l'effet des conditions sélectives différentes qu'ils subissent. Cette différenciation génétique peut rendre impossible les croisements entre ces sous-ensembles de telle sorte que deux espèces soient ainsi créées.

Par contre, dans le modèle à effet fondateur, c'est seulement une petite partie de la population qui est séparée de la principale. Cette population isolée subit une dérive génétique et réorganise son pool génétique. Cette différenciation génétique avec la population d'origine peut rendre impossible les croisements entre individus de ces deux pool génétiques créant ainsi deux espèces.

II.3.2- La spéciation sympatrique

La spéciation sympatrique est l'apparition au sein d'une même population de deux ensembles qui n'arrivent plus à se croiser, autrement dit de deux espèces. Cette spéciation survient grâce aux mutations géniques et au phénomène de polyploïdie (doublement du nombre de chromosomes). Ce type de spéciation est sujet à beaucoup de controverses.

II.3.3-Spéciation par remaniements chromosomiques

Les remaniements chromosomiques spontanés peuvent se fixer dans une petite population quelque part dans l'aire de répartition de l'espèce et induire un isolement postzygotique : c'est la spéciation dite stasipatrique ou parapatrique. Ce mode de spéciation est retrouvé dans le processus de spéciation de certaines espèces comme la drosophile. Certaines inversions chromosomiques y sont impliquées (Solignac *et al.*, 1995).

CHAPITRE III : Le complexe *Anopheles gambiae*

III.1- Historique

Le complexe *Anopheles gambiae* a d'abord été considéré comme étant une seule espèce (figure 2). Les divers croisements faits au laboratoire de manière réciproque entre les différentes espèces, ont produit en F1 des femelles fertiles, mais des mâles stériles avec deux exceptions : *An. quadriannulatus* femelle x *An. gambiae* mâle et *An. quadriannulatus* femelle x *An. bwambae* mâle qui ont donné des femelles et les mâles fertiles en F1. Actuellement sept espèces sont formellement reconnues au sein de ce complexe :

- Anopheles gambiae* sensu stricto Giles, 1902
- An. arabiensis* Patton, 1904
- An. bwambae* White, 1985
- An. melas* Theobald, 1903
- An. merus* Doenitz, 1902
- An. quadriannulatus A* Theobald, 1911
- An. quadriannulatus B*, (Hunt *et al.*, 1998)

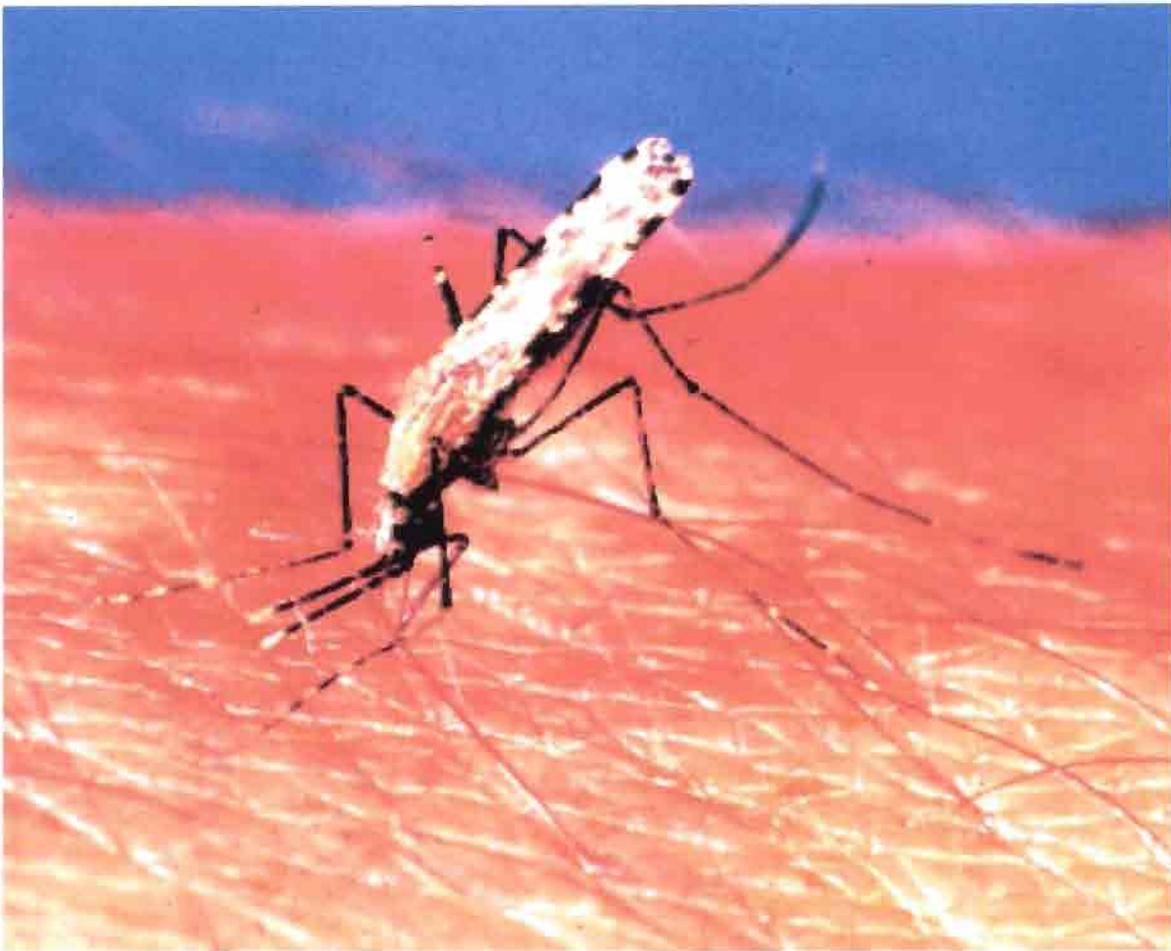


Figure 2 : Femelle d'*An. gambiae s.l.* prenant un repas de sang sur homme

Source TDR/OMS

La découverte d'espèces distinctes au sein du complexe *An. gambiae* a commencé à partir des études sur la transmission des gènes de résistance aux insecticides. Les croisements entre certaines lignées d'*An. gambiae s.l.* avaient donné une descendance mâle stérile (Davidson, 1962 ; 1964 ; Davidson & Jackson, 1962 ; Davidson & Hunt, 1973). Les travaux de cytogénétique ont confirmé que ces espèces étaient génétiquement distinctes (Coluzzi, 1966 ; Coluzzi & Sabatini, 1967, 1968, 1969). Ils ont montré que ces espèces étaient chacune caractérisées par des inversions fixées et stables même dans les zones de sympatrie. Les inversions chromosomiques sont mises en évidence par coloration des chromosomes géants isolés des ovaires de femelles semi-gravidés ou des glandes salivaires des larves de stade IV. (Green, 1972). Les études à l'aide d'isoenzymes (Bullini & Coluzzi, 1978 ; Mahon *et al.*, 1976 ; Miles, 1978) et les différences au niveau des profils obtenus en chromatographie en phase gazeuse des hydrocarbures cuticulaires (Milligan *et al.*, 1993) ont aussi confirmé la forte divergence génétique entre ces différentes espèces. Grâce aux progrès des techniques de biologie moléculaire, des méthodes de diagnostic à base de sondes d'ADN ont été développées (Collins *et al.*, 1987, 1988). Ceci a abouti à la mise au point d'un test diagnostique des

différents membres du complexe *An. gambiae* basé sur l'amplification par PCR des gènes codant pour les ARN ribosomiaux (Scott *et al.*, 1993). Cette méthode présente l'avantage de permettre l'identification rapide des spécimens collectés sur le terrain, quels que soient le sexe et le stade de développement. Son efficacité a été démontré (Fontenille *et al.*, 1993) et il est maintenant utilisé en routine dans plusieurs laboratoires.

Plusieurs méthodes ont montré que *An. gambiae s.l.* est bien constitué de plusieurs espèces, mais il subsiste la possibilité d'un flux de gènes (introgression) interspécifique occasionnel à travers des femelles hybrides qui sont fertiles (Besansky *et al.*, 1997 ; Thewell *et al.*, 2000 ; Touré *et al.*, 1998).

III.2- Biologie et distribution géographique des membres du complexe

Cette distribution géographique est illustrée par la figure 3.

An. gambiae et *An. arabiensis* sont les espèces les plus anthropophiles du complexe *An. gambiae* et ayant la plus large distribution géographique sur le continent africain. Malgré l'existence de différences au niveau de leur comportement et de leur capacité vectorielle, ces espèces sont toutes deux fortement impliquées dans la transmission du paludisme en Afrique. *An. gambiae* et *An. arabiensis* partagent une vaste zone de recouvrement, mais elles n'ont pas la même aire de répartition (figure 3). Dans les zones de sympatrie, on observe une variation saisonnière de leur prévalence, à cause d'une différence de comportement vis à vis des conditions climatiques. *An. arabiensis* est prédominant en saison sèche. *An. gambiae* a une préférence pour les zones de forêt et de savane humide alors qu'*An. arabiensis* est plus adapté aux zones de savane sèche et aride et aux steppes. Des études indiquent aussi qu'*An. arabiensis* serait plus zoophile et exophile qu'*An. gambiae* (Fontenille *et al.*, 1997). Leurs larves se développent dans les collections d'eau claire peu profondes, ensoleillées et souvent dépourvues de végétation. Ce sont en général des empreintes de pas, des sillons creusés par des roues de voitures, des fossés, des flaques ou encore des rizières avec le système d'irrigation. Les espèces *An. quadriannulatus A* et *B* sont reconnues comme étant essentiellement zoophiles et de ce fait n'ont aucun intérêt médical. Elles sont rencontrées en sympatrie avec *An. arabiensis* et moins avec *An. gambiae*. Les larves de ces espèces se développent dans l'eau douce. *An. quadriannulatus A* n'est rencontrée que dans quelques pays notamment l'Éthiopie, l'Afrique du Sud, Zimbabwe, Mozambique et Zanzibar.

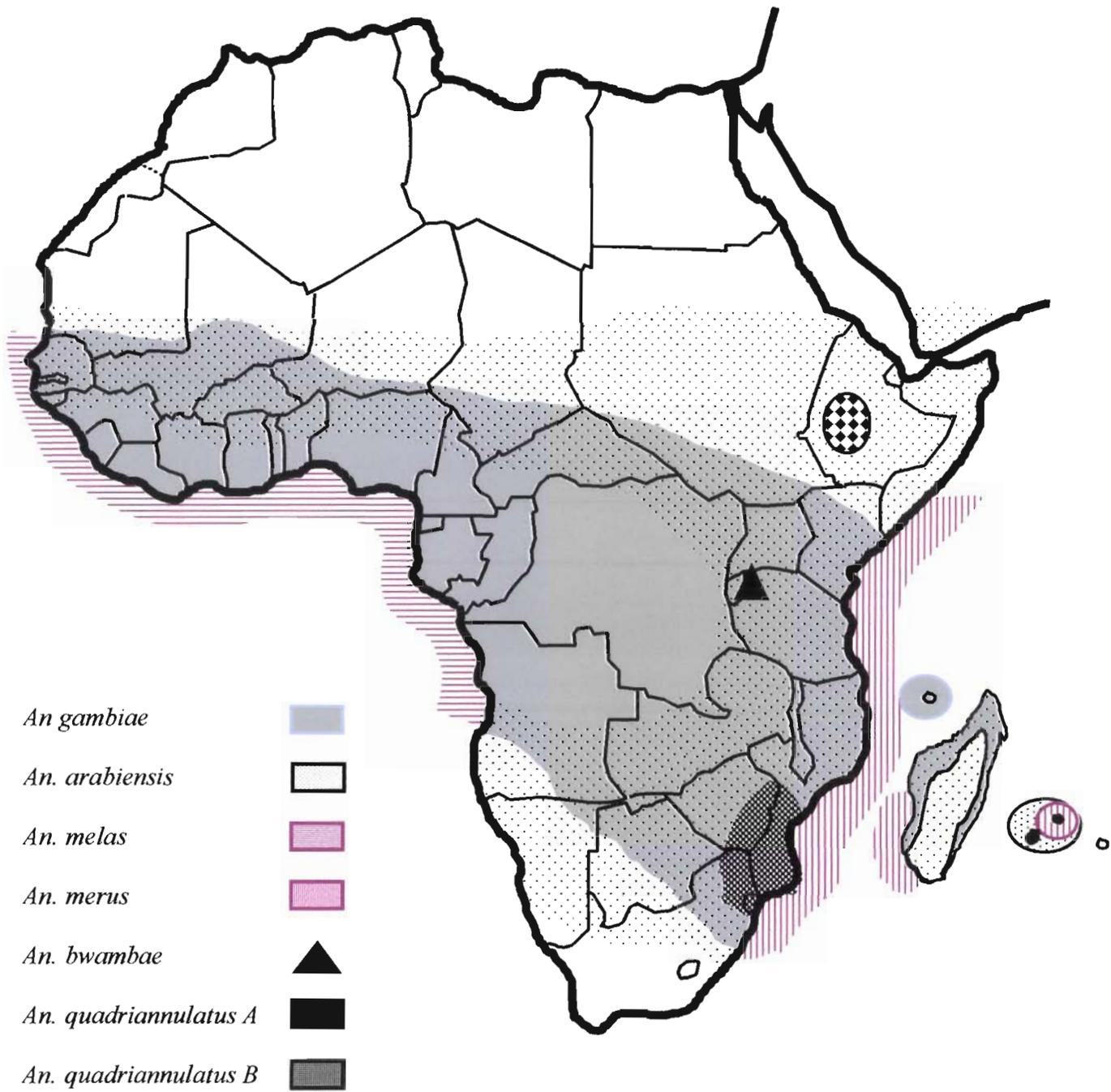


Figure 3 : Répartition des espèces du complexe *Anopheles gambiae* en Afrique.

D'après Danis & Mouchet, 1991

An. quadriannulatus A a été trouvé complètement exophile dans certaines régions alors qu'en zone d'altitude elle tendait à être plus endophile. *An. quadriannulatus B* n'a été signalé qu'en Ethiopie. Ces deux espèces ne sont différenciables qu'au travers des expériences de croisement car ils produisent en F1 des hybrides avec les mâles stériles (Hunt *et al.*, 1998).

An. melas et *An. merus* sont toutes deux des espèces associées aux eaux salées des côtes Ouest et Est africaines respectivement. Ces espèces peuvent piquer l'homme en l'absence d'hôtes alternatifs. Toutefois, leur capacité vectorielle pour la transmission des *Plasmodium* ou des filaires de Bancroft est nettement plus faible que celles d'*An. gambiae* et *An. arabiensis*. Elles peuvent néanmoins jouer un rôle local dans la transmission du paludisme. Ces espèces sont aussi bien zoophiles et qu'exophiles.

L'espèce *An. bwambe* n'a été trouvée que dans la région de la forêt de Semliki en Ouganda où elle est associée aux sources thermales dans lesquelles se développent ses larves (Davidson & White, 1972). Son exophilie a été signalée ainsi que sa zoophile, mais elle pique parfois l'homme. Elle a été trouvée infestée de *Plasmodium* ce qui suppose qu'elle pourrait contribuer localement à la transmission du paludisme (White, 1985).

An. melas, *An. merus* et *An. bwambe* sont complètement allopatriques les unes des autres mais, elles sont parfois rencontrées en association avec *An. gambiae* et *An. arabiensis* au moins au stade adulte.

III.3- Structure génétique et spéciation au sein du complexe *An. gambiae*

III.3.1- Cytologie d'*An. gambiae s.l.*

Les différences trouvées entre les espèces du complexe *An. gambiae* au niveau des chromosomes polytènes, consistent essentiellement en des changements de bandes de séquences dues à des inversions paracentriques. Celles-ci ayant été déterminées et classifiées à partir des séquences standard d'*An. quadriannulatus*. Certaines inversions sont fixées et sont utilisées comme marqueurs d'espèces, d'autres sont polymorphes. Le bras 2R étant celui qui possède le plus d'inversions. Il a été démontré que le polymorphisme du deuxième bras chromosomique était associé étroitement à l'adaptation aux différentes conditions éco-climatiques telles que l'aridité (Coluzzi *et al.*, 1979). Une relation entre le comportement des anophèles et leurs caryotypes a aussi été mis en évidence. Ainsi chez *An. gambiae*, les spécimens se reposant à l'intérieur des maisons n'ont pas les mêmes caryotypes que ceux s'abritant à l'extérieur. Ceci est autant vrai chez *An. gambiae* que chez *An. arabiensis*. L'on observe une hétérogénéité dans la distribution des inversions au sein des différentes espèces. *An. merus* ne présente aucun polymorphisme alors que *An. bwambe*, *An. quadriannulatus* et *An. melas* en ont juste un peu. *An. gambiae* et *An. arabiensis* présentent un fort polymorphisme des inversions chromosomiques.

L'existence de plusieurs taxons chez *An. gambiae* a été mise en évidence sur la base des déséquilibres de Hardy–Weinberg observés pour certains caryotypes du chromosome 2. Ce déséquilibre étant dû à un déficit en hétérozygotes pour ces caryotypes. Ces travaux ont été menés au Mali ainsi que dans les pays voisins par Coluzzi *et al* (1985) et Touré *et al* (1983, 1994). Cinq formes chromosomiques ont été définies : Savanna, Bamako, Mopti, Bissau et Forest. La fréquence de ces formes varie en fonction des faciès écologiques et les saisons.

La forme Savanna est dépendante de la pluie pour ses gîtes larvaires et par conséquent est rare en saison sèche. Cette forme est répandue à travers l'Afrique. En Afrique de l'Est, elle est caractérisée par un polymorphisme des inversions 2Rb et 2La, alors qu'en Afrique de l'Ouest, elle présente des fréquences élevées des arrangements 2Rbcu et 2Rcu.

La forme Bamako est étroitement associée aux bassins de rivières. Ses gîtes larvaires sont les bords de rivières. Elle est trouvée uniquement au Mali et au Nord de la Guinée (Conakry). La forme Bamako est caractérisée par des arrangements fixés : 2Rjcu et 2La, alors que l'inversion 2Rb est polymorphe.

La forme Mopti est étroitement associée aux plaines inondées et champs irrigués. Pendant la saison sèche, elle est souvent dans de nombreuses régions, la seule forme d'*An. gambiae* et par conséquent, est plus adaptée à l'aridité. La littérature signale sa présence au Mali, Guinée Conakry, Côte d'Ivoire, Burkina Faso et au Nigeria. Cette forme est caractérisée par un polymorphisme de l'arrangement 2Rbc/u/+ et l'inversion 2La est presque fixée.

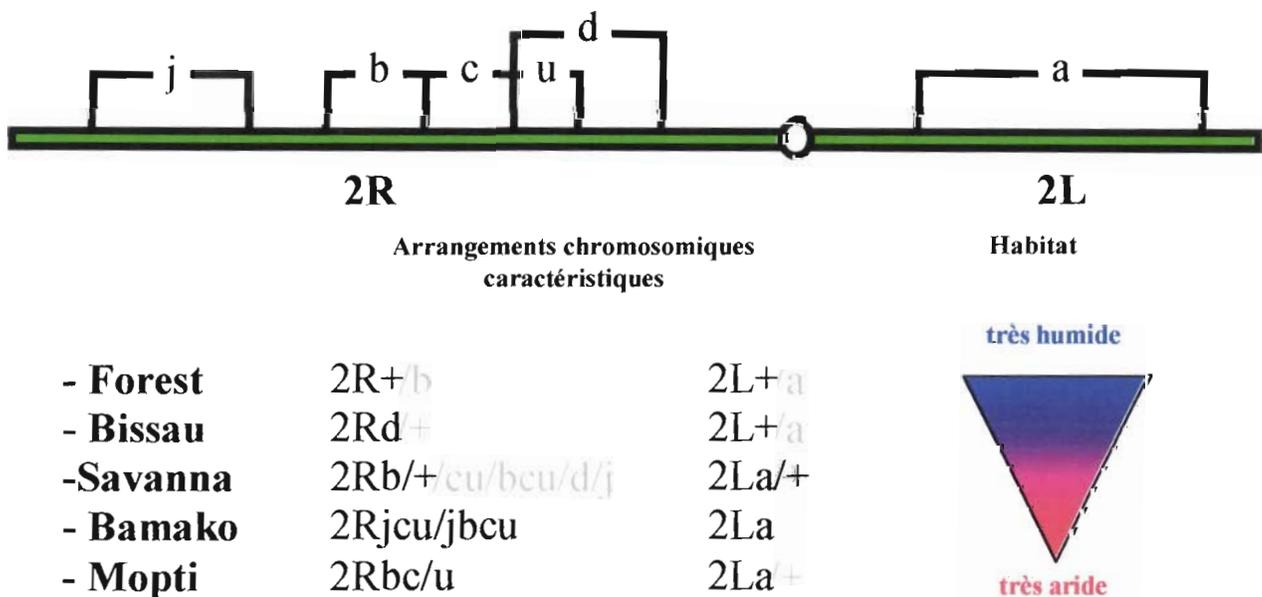


Figure 4 : Inversions et gradient climatique caractérisant les cinq formes chromosomiques d'*An. gambiae*

La forme Forest est caractérisée par l'arrangement standard sur tout le chromosome 2 (2R+/+ ; 2L+/+) avec peu d'inversions à de faibles fréquences (2La, 2Rb, 2Rd). Elle se trouve dans les zones de forêt et de savane humide.

La forme Bissau est caractérisée par l'arrangement 2Rd. Elle se trouve dans les zones de riziculture irriguées des côtes de la Gambie et de la Casamance au Sénégal.

Ces différentes formes chromosomiques sont partiellement ou totalement isolées dans la nature. Les études effectuées par Milligan *et al* (1993) ont suggérée l'existence probable d'un mécanisme d'isolement pré-zygotique entre certaines de ces formes. En effet, ces auteurs ont trouvé que les profils d'hydrocarbones cuticulaires des formes chromosomiques étaient plus différenciés lorsque les populations de formes différentes étaient sympatriques mais étaient par contre plus homogènes lorsqu'elles étaient allopatriques. Les hydrocarbones cuticulaires sont en effet reconnus comme jouant un rôle important dans les processus de reconnaissance précédant le croisement entre partenaires sexuels chez de nombreuses espèces d'insectes (Higgie *et al.*, 2000). Très peu d'hybrides de ces formes chromosomiques sont rencontrés dans la nature. Par contre, aucun isolement reproductif post-copulatoire entre ces formes n'a été trouvé car l'on obtient des hybrides entre différentes formes au laboratoire et parfois avec une fécondité plus élevée (Di Deco *et al.*, 1980). Toutes ces données laissent supposer l'existence d'un mécanisme de spéciation chez *An. gambiae*.

An. arabiensis présente également un polymorphisme au niveau des inversions mais, les différents arrangements sont généralement en équilibre de Hardy-Weinberg suggérant la panmixie à l'intérieur de cette espèce (Mekuria *et al.*, 1982 ; Coluzzi *et al.*, 1985). Toutefois, une corrélation a été mise en évidence entre la fréquence de certaines inversions et le comportement trophique de cette espèce (zoo- ou anthropophile, endo- ou exophile) ou l'environnement dans lequel il évolue (climat plus ou moins sec, saison de pluies ou saison sèche) (Petarca & Beier, 1992 ; Petarca *et al.*, 1986, 1987 ; Mnzava *et al.*, 1995). Toutefois, ces observations d'hétérogénéité génétique sont moins évidentes que chez *An. gambiae*.

L'étroite association entre les formes chromosomiques et différents niveaux d'aridité, suggère que les inversions qui les caractérisent, représentent une adaptation aux conditions environnementales du milieu où vit *An. gambiae*. Une association similaire a été aussi mise en évidence chez la drosophile et a permis de constater que les inversions représentent des complexes de gènes co-adaptés ou super gènes (Dobzhansky, 1970 pp. 143-146). Ainsi les inversions apparaissent comme étant de larges unités de recombinaison impliquant un ensemble de gènes co-adaptés, conférant aux individus des capacités différentes d'adaptation à des différents milieux ou

III.3.2- Etudes biochimiques à base d'isoenzymes

Des méthodes de diagnostic à l'aide des marqueurs isoenzymatiques ont été développées afin d'identifier les différentes espèces du complexe *An. gambiae* (Mahon *et al.*, 1976). Elles consistent essentiellement à la recherche d'allèles spécifiques à chaque espèce du complexe. Un système enzymatique (aspartate aminotransférase) a dans ce cadre été identifié chez *An. quadriannulatus* qui la différencie aisément d'*An. arabiensis* (Mpofu *et al.*, 1993). Les isoenzymes ont permis de mettre en évidence des taux de différenciation significatifs (bien qu'étant assez modérés), entre des populations d'*An. gambiae* d'Afrique de l'Ouest et celles de l'Est (Miles *et al.*, 1978 ; Lehmann *et al.*, 1996a). Des études visant à estimer le niveau de différenciation génétique entre les formes chromosomiques Bamako, Savanna et Mopti ont été menées en utilisant 27 systèmes enzymatiques (Bullini & Coluzzi, 1978 ; Bullini, 1984 ; Cianchi *et al.* 1983). Celles-ci n'ont pas permis la mise en évidence d'une différenciation significative entre ces différentes formes chromosomiques d'*An. gambiae*, qu'elles soient sympatriques ou non.

III.3.3- Les études de génétique moléculaire

La plupart des études moléculaires menées jusqu'à présent se sont concentrées sur *An. gambiae* et *An. arabiensis*, les espèces les plus importantes du complexe. Les techniques de biologie moléculaire ont permis des avancées considérables dans la compréhension de la phylogénie, de la structuration génétique au sein du complexe *An. gambiae*.

Les marqueurs RAPD (Random Amplified Polymorphic DNA) ont été largement utilisés comme outils diagnostic pour différencier les membres du complexe *An. gambiae*. Ils ont aussi été utilisés pour la construction des cartes génétiques (Dimopoulos *et al.*, 1996). Les RAPD ont servi de marqueurs pour localiser sur le génome de *An. gambiae s.s.*, des traits particuliers tels que les breakpoints (points de rupture sur les chromosomes) des inversions (Favia *et al.*, 1994).

La structure et la séquence des unités répétitives de l'ADN ribosomal (ADNr) ont été comparées entre les différentes espèces du complexe (McLain *et al.*, 1989). Ces unités sont composées de régions codants pour la sous-unité 28S, 5,8S et 18S des ARN ribosomiaux, de régions d'espaces transcrits internes («Internal transcribed spacer ; ITS») et les espaces intergéniques (IGS pour Intergenic Spacer). Les variations dans les IGS ont été utilisées pour développer une méthode diagnostic permettant de différencier les membres du complexe *An. gambiae* (Scott *et al.*, 1993). Néanmoins cette technique ne permet pas de différencier *An. quadriannulatus A* d'*An. quadriannulatus B* car ils ne présentent aucune différence au niveau des séquences de l'ADNr. Seule les méthodes de croisement permettent leur séparation (Hunt *et al.*, 1998).

L'analyse de l'ADN mitochondrial a révélé de très faibles divergences tant intra qu'interspécifiques entre les membres du complexe *An. gambiae*. Cette faible divergence a poussé certains auteurs à l'expliquer par un phénomène d'introgession (passage de gènes d'une espèce à

une autre) entre espèces (Besansky *et al.*, 1997). Ce type d'étude a particulièrement été mené entre *An. gambiae* et *An. arabiensis*. L'analyse de séquences partielles des gènes du cytochrome b, ND1 et ND5, n'a révélé aucune différence fixée entre espèces. Les phénomènes d'introgession ont aussi été signalés avec l'ADNr et l'ADNmt d'*An. bwambae* (Thelwell *et al.*, 2000). De tout ceci, il résulte que les phénomènes d'introgession, bien qu'étant des événements rares, surviennent plus fréquemment entre espèces du complexe *An. gambiae* qu'on ne le supposerait.

Au sein d'*An. gambiae*, les études moléculaires ont permis de développer des méthodes pour identifier les formes moléculaires de cette espèce. Dès 1994, Favia *et al.* trouvaient des marqueurs RAPD différenciant certaines formes chromosomiques. Une méthode diagnostic de PCR-RFLP basée sur l'analyse des séquences d'IGS de l'ADN ribosomal a été mise au point par Favia *et al.* (1997). Elle permet, au Mali et au Burkina Faso, de différencier les formes chromosomiques Mopti de Savanna et Bamako, mais ne sépare pas la forme Bamako de la forme Savanna. Cette méthode a été améliorée par Favia *et al.* (2001) en supprimant l'étape de la RFLP pour permettre une identification directe par PCR. Il a été convenu d'appeler «forme moléculaire M», la forme correspondant à la forme chromosomique Mopti et d'appeler «forme moléculaire S», la forme moléculaire correspondant aux formes chromosomiques Bamako et Savanna (Della Torre *et al.*, 2001). Toutefois, cette correspondance n'existe qu'au sein de deux pays à savoir le Mali et le Burkina Faso, puisque hors de ces pays, cette correspondance ne tient plus. En plus, certaines formes chromosomiques comme Forest et Savanna comprennent les deux formes moléculaires M et S (Della Torre *et al.*, 2001). Il est à noter que le marqueur moléculaire distinguant la forme chromosomique Mopti de Bamako et Savanna est situé sur le chromosome X, indépendamment du chromosome 2 où sont situées les inversions spécifiques à chaque forme chromosomique. Un autre élément appuyant l'isolement reproductif entre les formes moléculaires M et S vient des travaux de Chandre *et al.* (1999a et b), qui ont montré qu'en Côte d'Ivoire, le gène *kdr* (Knock down resistance) n'était présent que chez la forme S et non chez la forme moléculaire M quand ces formes sont sympatriques. En plus, aucun hybride M/S n'a été trouvé dans ce pays dans la nature. De ce fait, le statut taxonomique de ces deux formes moléculaires est sujet à d'intenses débats. S'agit-il de deux espèces ou de deux unités d'une même espèce?

III.3.4- Apport des microsatellites

Pour ce qui concerne le complexe *An. gambiae*, les locus microsatellites ont été utilisés pour comparer les populations sympatriques d'*An. arabiensis* et d'*An. gambiae* (Kamau *et al.*, 1998b). Cette étude a révélé un niveau de flux génique limité entre ces espèces jumelles.

La structure génétique au sein d'*An. gambiae* a aussi été étudiée à l'aide des microsatellites. La première étude fut celle conduite par Lanzaro *et al.* (1995) en utilisant 11 locus et qui montra le fort polymorphisme des microsatellites chez *An. gambiae* suggérant que ce marqueur génétique

pouvait être très utile pour les études de génétique de populations chez les anophèles. Divers types d'études ont été menés à l'instar des études macrogéographiques sur plus de 6000 km entre les populations du Sénégal et du Kenya (Lehmann *et al.*, 1996a). Ce travail qui a utilisé 5 locus microsattellites a indiqué une faible différenciation entre ces populations malgré la grande distance géographique entre elles ($F_{st} = 0,016$). Il est à noter que ces populations appartenaient toutes à la forme chromosomique Savanna. Au niveau microgéographique, les études menées n'ont indiqué aucune différenciation entre populations prises au niveau des maisons d'un même village (Lehmann *et al.*, 1997). Le diamètre associé à un dème (sous-population) a été estimé à 50 km. Des études ont également été menées pour estimer les flux de gènes entre les différentes formes chromosomiques Mopti, Savanna et Bamako au Mali (Lanzaro *et al.* 1998). Vingt et un locus microsattellites situés autant à l'extérieur qu'à l'intérieur des inversions chromosomiques ont été utilisés par les auteurs et seuls ceux situés à l'intérieur des inversions ont indiqué une différenciation génétique entre ces formes. Par contre, une homogénéité entre les formes chromosomiques était trouvée au niveau des locus situés hors des inversions. Ceci a amené les auteurs à suggérer que la divergence génétique observée entre ces formes chromosomiques serait plutôt due à un phénomène de sélection associé aux inversions chromosomiques (nous y reviendrons plus en détail dans la deuxième partie de ce travail).

Il ressort de ces travaux sur *An. gambiae*, que la structure génétique des populations au sein de cette espèce est sujet à controverse et notamment la question du statut taxonomique des différentes formes chromosomiques et moléculaires mises en évidence jusqu'à ce jour. C'est dans la perspective de répondre à ces questions que nous avons entrepris l'étude de la structure génétique d'*An. gambiae* au Cameroun, pays vierge de toute étude sur ce domaine jusqu'à présent. Les différents facteurs (tels que les inversions chromosomiques) ayant pu introduire des erreurs d'interprétations dans les données des précédents travaux sur la structure génétique d'*An. gambiae*, peuvent être évités au Cameroun. En effet, on y trouve de façon prédominante dans sa partie sud, la forme chromosomique Forest qui n'a pas ou alors très peu d'inversions.

III.4- Lutte anti-vectorielle

La lutte contre le paludisme constitue un énorme challenge pour la communauté scientifique. En effet, de nombreux obstacles jalonnent cette initiative à savoir entre autres 1) la résistance aux divers antimalariques de *Plasmodium falciparum*, principal agent du paludisme, 2) la difficulté à mettre au point un vaccin efficace contre les différents stades du parasite. De ce fait, la lutte antivectorielle demeure une méthode majeure de prévention de masse qui soit applicable dans les zones d'endémie palustre. Elle vise principalement à réduire les contacts hommes-moustiques et/ou empêcher la prolifération des différents vecteurs. De nombreuses techniques de lutte

antivectorielle ont été mises au point. Mais pour répondre à des critères d'efficacité et de sélectivité, la lutte doit tenir compte de la bio-écologie et du comportement des espèces cibles ainsi que du contexte épidémiologique de la zone de lutte (OMS, 1998). On peut distinguer les types de lutte suivants :

III.4.1- Lutte antilarvaire

Cette forme de lutte s'attaque uniquement au stade larvaire des anophèles. Elle est surtout préconisée lorsque le nombre de gîtes est limité et facilement repérable et accessible. On utilise pour cela des moyens mécaniques tels que l'aménagement de l'environnement et le drainage des zones marécageuses (Mouchet *et al.*, 1991) afin de permettre une limitation des gîtes larvaires. On utilise aussi des traitements larvicides à l'aide d'insecticides chimiques comme le téméphos ou d'agents biologiques tels que les pathogènes bactériens comme *Bacillus thuringensis*, des prédateurs tels que les poissons larvivores (*Gambusia*) et des parasites comme les champignons, les nématodes ou les microsporidies. Il faut signaler qu'à cause de la variété des gîtes des anophèles (marécages, cours d'eau etc...), ces méthodes dites biologiques sont moins efficaces que d'autres moyens de contrôle comme la lutte imagocide (Danis & Mouchet, 1991) sauf dans certains cas particuliers (zones urbaines et périurbaines).

III.4.2- Lutte imagocide

Avec la mise au point des insecticides chimiques dont le DDT à la fin des années 40, la lutte contre le stade adulte des moustiques et particulièrement des anophèles a connu une véritable révolution. Toutefois, l'apparition de résistances aux insecticides notamment au DDT chez les anophèles, impose des contraintes qu'il faut prendre en compte dans la gestion de ce moyen de contrôle. Cette lutte se fait par des pulvérisations intra-domiciliaires d'insecticides rémanents particulièrement bien adaptés pour lutter contre les vecteurs du paludisme ayant un comportement endophile. Les pyréthriinoïdes ont progressivement pris le pas sur le DDT comme principaux composés insecticides utilisés dans cette lutte. On peut citer entre autres la perméthrine, la deltaméthrine et la lambda-cyhalothrine.

Une autre étape de la lutte se fait au travers de la protection individuelle et collective. En effet, les moustiquaires par la barrière qu'elles constituent entre l'anophèle et l'homme, apportent une certaine protection contre les piqûres. Leur utilisation a connu un regain d'intérêt grâce à leur imprégnation à l'aide de pyréthriinoïdes (Darriet *et al.*, 1984). Les essais en cases expérimentales ont montré l'efficacité des moustiquaires imprégnées de pyréthriinoïdes (Carnevale *et al.*, 1988 ; Darriet, 1991). Cette efficacité se situait à plusieurs niveaux 1) Elles entraînent souvent une diminution du nombre de moustiques qui entrent à l'intérieur de la maison, 2) elles réduisent le taux de femelles gorgées même lorsque les moustiquaires sont trouées ou mal bordées ; 3) elles

Darriet, 1991). Cette efficacité se situait à plusieurs niveaux 1) Elles entraînent souvent une diminution du nombre de moustiques qui entrent à l'intérieur de la maison, 2) elles réduisent le taux de femelles gorgées même lorsque les moustiquaires sont trouées ou mal bordées ; 3) elles augmentent la proportion de femelles qui s'échappent des maisons. Les moustiquaires imprégnées entraînent ainsi une réduction importante du contact homme-vecteur et ont un effet délétère sur la survie et la fécondité des vecteurs anthropophiles comme *An. gambiae* et *An. arabiensis*. A l'échelle de la communauté, l'utilisation des moustiquaires constitue un moyen efficace de prévention contre le paludisme (Robert *et al.*, 1989a).

III.4.3- Voie nouvelle de lutte par la génétique

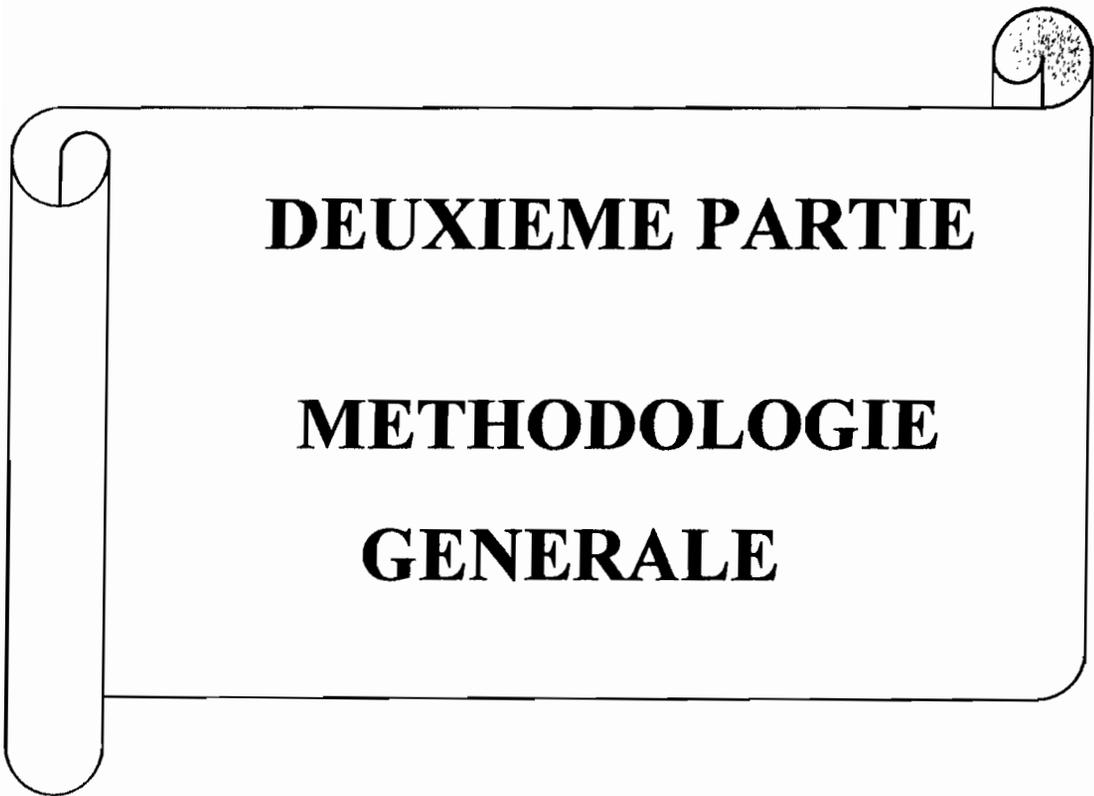
Les progrès enregistrés en biologie moléculaire ont permis le développement d'une nouvelle discipline, l'entomologie moléculaire. Celle-ci s'intéresse à l'étude du génome des insectes dans le but de mieux comprendre son organisation et de pouvoir aussi trouver de nouveaux moyens de lutte contre ceux-ci. C'est dans cette perspective que s'inscrit la manipulation génétique des vecteurs pour en altérer la capacité vectorielle (Crampton *et al.*, 1990 ; Collins & James, 1996). Cette technique suppose l'introduction chez les populations naturelles d'anophèles par les moyens de la transgénèse, de gènes conférant un caractère réfractaire à une infection par *Plasmodium*. Des efforts considérables ont été faits pour identifier les mécanismes par lesquels les moustiques sont réfractaires au parasite (Zheng *et al.*, 1997 ; Feldmann *et al.*, 1998 ; Vernick *et al.*, 1995). Deux espèces d'anophèles, *An. stephensi* (Catterucia *et al.*, 2000) et *An. gambiae* (Grossman *et al.*, 2001), ont ainsi déjà été transformées au laboratoire. Toutefois, les études visant à introduire ces gènes dans les populations naturelles des anophèles bien que n'étant pas encore très avancées (Kidwell & Ribeiro, 1998), sont l'objet de beaucoup d'efforts de recherche (Boëte & Koella, 2003). Le succès d'un programme de contrôle du paludisme basé sur la modification génétique des populations requiert une connaissance approfondie de la structure génétique des populations vectorielles et des différentes barrières au flux de gènes. Cette connaissance sera primordiale pour la détermination de l'échelle spatiale et temporelle nécessaire pour l'introduction de tels gènes. De telles études sont aussi importantes pour une meilleure gestion des moyens de lutte disponibles car elles apportent des informations permettant d'évaluer la capacité de diffusion des gènes de résistance aux insecticides ou celles des gènes conférant une plus grande compétence vectorielle chez certains sous-populations d'une anophèles. Notre travail s'inscrit pleinement dans cette perspective.

III.5- Présentation de l'étude

Les études de génétique des populations ont été menées jusqu'à présent au niveau des populations de vecteurs du paludisme en Afrique de l'Ouest et de l'Est, mais pas du tout en Afrique Centrale. Pourtant cette région présente de très grandes particularités écologiques (végétation, faune, relief, pluviométrie) susceptibles d'induire de nombreuses variabilités génétiques chez ces populations. Dans le but de répondre aux nombreuses questions que l'on se pose à ce jour sur la caractérisation et la structuration génétique des populations d'*An. gambiae* et d'évaluer les implications éventuelles sur le contrôle de cet important vecteur, nous avons entrepris une étude sur les deux volets suivants :

Nous avons dans la première partie de ce travail entrepris la caractérisation génétique des populations d'*An. gambiae* sur les plans moléculaire, cytogénétique et écologique. Nous avons également procédé à l'étude de la distribution géographique des différentes formes moléculaires M et S d'*An. gambiae* au Cameroun avec une évaluation du rôle épidémiologique de chacune de ces formes dans la transmission du paludisme.

Nous avons dans la deuxième partie, étudié la structuration génétique des populations d'*An. gambiae* au Cameroun en fonction des formes moléculaires M et S. Dans le souci d'éviter l'influence observée des inversions chromosomiques sur le profil de différenciation génétique entre populations d'*An. gambiae*, nous avons pour cette étude, ciblé les populations de forme chromosomique Forest qui présentent l'arrangement standard sans inversion sur le chromosome 2. Nous avons utilisé comme marqueurs génétiques les locus ADN pour déterminer le degré de divergence génétique entre les formes moléculaires d'*An. gambiae*.

A decorative border resembling a scroll, with a vertical strip on the left side and a circular element at the top right corner containing a tree-like pattern.

DEUXIEME PARTIE

METHODOLOGIE

GENERALE

1- Description des sites choisis pour l'étude de la structuration génétique

1.1- Simbock

La localité de Simbock est un village situé à la périphérie de la ville de Yaoundé (3°50'N, 11°30'E) dans l'arrondissement de Yaoundé 6^{ème}. Ce village compte environ 300 habitants vivants pour la plupart dans des habitations de style traditionnel faites de terre battue et de toits en tôle ondulée. Avec la poussée de l'urbanisation, on observe de plus en plus des constructions modernes. La plupart des habitants travaillent à Yaoundé ou s'occupent par des travaux champêtres. Le village est traversé par la rivière Mefou qui crée une zone marécageuse permanente. La variété de l'écosystème de cette localité est responsable de la présence dans cette localité, d'une diversité de vecteurs du paludisme dont *An. moucheti*, *An. nili*, *An. funestus* et *An. gambiae*. En outre, la présence de plusieurs étangs de pisciculture augmente le nombre de gîtes larvaires induisant ainsi une transmission élevée dans ce village (environ 300 piqûres infectantes/an) (Antonio-Nkondjio *et al.*, 2002). Le climat est de type équatorial avec deux saisons sèches (décembre-février et juillet-août) et deux saisons pluvieuses (mars-juin et septembre-novembre). La pluviométrie moyenne annuelle est environ de 1500 mm avec une température minimale et maximale allant de 18-25°C en juillet à 20-29°C en mars. Cette localité est située à environ 750m d'altitude.

1.2- Mbalmayo

La ville de Mbalmayo (3°30 N et 11°26' E) est située dans le département du Nyong et So'o, Province du Centre. Elle est traversée par le Nyong qui est un fleuve à débit lent et qui peut être recouvert d'une abondante végétation. Anciennement lieu important de vente du cacao, Mbalmayo est aujourd'hui un centre urbain ayant plusieurs activités dont la principale est l'industrie du bois. La ville se situe dans une région forestière dégradée. Le climat est semblable à celui de Yaoundé et de Simbock. La pluviométrie moyenne annuelle se situe autour de 1615 mm de pluies. Les enquêtes ont été menées dans le quartier de la gare qui abrite la vieille gare ferroviaire sur les rives du Nyong. Cette localité est située à environ 640 m d'altitude. Le climat est semblable à celui de Simbock.

1.3- Dschang

La localité de Dschang (5°25'N ; 10°10'E) est située dans la Province de l'Ouest et est le chef lieu du département de la Ménoua. Cette ville est située à environ 1400 m d'altitude. Le climat y est généralement très doux avec des températures modérées autour de 18-20°C. La pluviométrie moyenne annuelle se situe autour de 1900 mm de pluies. La végétation est de type forêt d'altitude. L'activité principale de la population est l'agriculture. On note également la présence d'une Université dans la ville avec une population estudiantine importante. Le relief ici est assez accidenté

avec des collines et des vallées qui offrent des zones marécageuses créant des gîtes pour moustiques.

1.4- Santchou

La localité de Santchou (5°17'N ; 9°58'E) est située dans le département de la Ménoua Province de l'Ouest et chef lieu de l'arrondissement de Santchou. Elle est séparée seulement de 20 km (à vol d'oiseau) de la ville de Dschang mais se trouve à une altitude nettement plus faible (700m). La différence avec Dschang est assez frappante sur le plan de la végétation, du climat qui est plus chaud et proche de celui des zones littorale. La pluviométrie moyenne annuelle est de 2200 mm. La température moyenne est d'environ de 25°C. Le relief est assez plat et offre des opportunités aux flaques d'eau de stagner surtout en saison pluvieuse ce qui constitue de bons gîtes pour les moustiques.

1.5- Nkoteng

La localité de Nkoteng (4°30'N, 12°03'E) située à 590m est une bourgade du département de la Haute-Sanaga et caractérisée par l'existence de vastes champs de canne à sucre. Ces plantations sont la propriété de la SOSUCAM (Société Sucrière du Cameroun). L'on se trouve ici dans une zone de transition forêt-savane caractérisée par une végétation de savane arbustive et un climat relativement chaud avec des températures moyennes annuelles autour de 26°C. La pluviométrie moyenne annuelle est de 1550 mm. Les habitants pour la grande majorité sont employés dans les plantations de canne à sucre.

2- Echantillonnage des spécimens

L'échantillonnage des moustiques s'est fait lors d'enquêtes de terrain dans les zones d'études ciblées. Le choix des méthodes a été orienté en fonction des objectifs des différentes études.

2.1- Captures sur volontaires

Cette méthode a été utilisée pour étudier le contact entre l'homme et le vecteur par estimation des paramètres entomologiques tels que le taux de piqûre, le taux d'endo- ou d'exophagie du vecteur, le cycle d'agressivité nocturne. Les anophèles piquant leurs différents hôtes essentiellement pendant la nuit, les captures ont été organisées de 19h à 6h du matin sur des jeunes hommes volontaires recrutés dans les différents sites de capture. Ceux-ci sont positionnés à

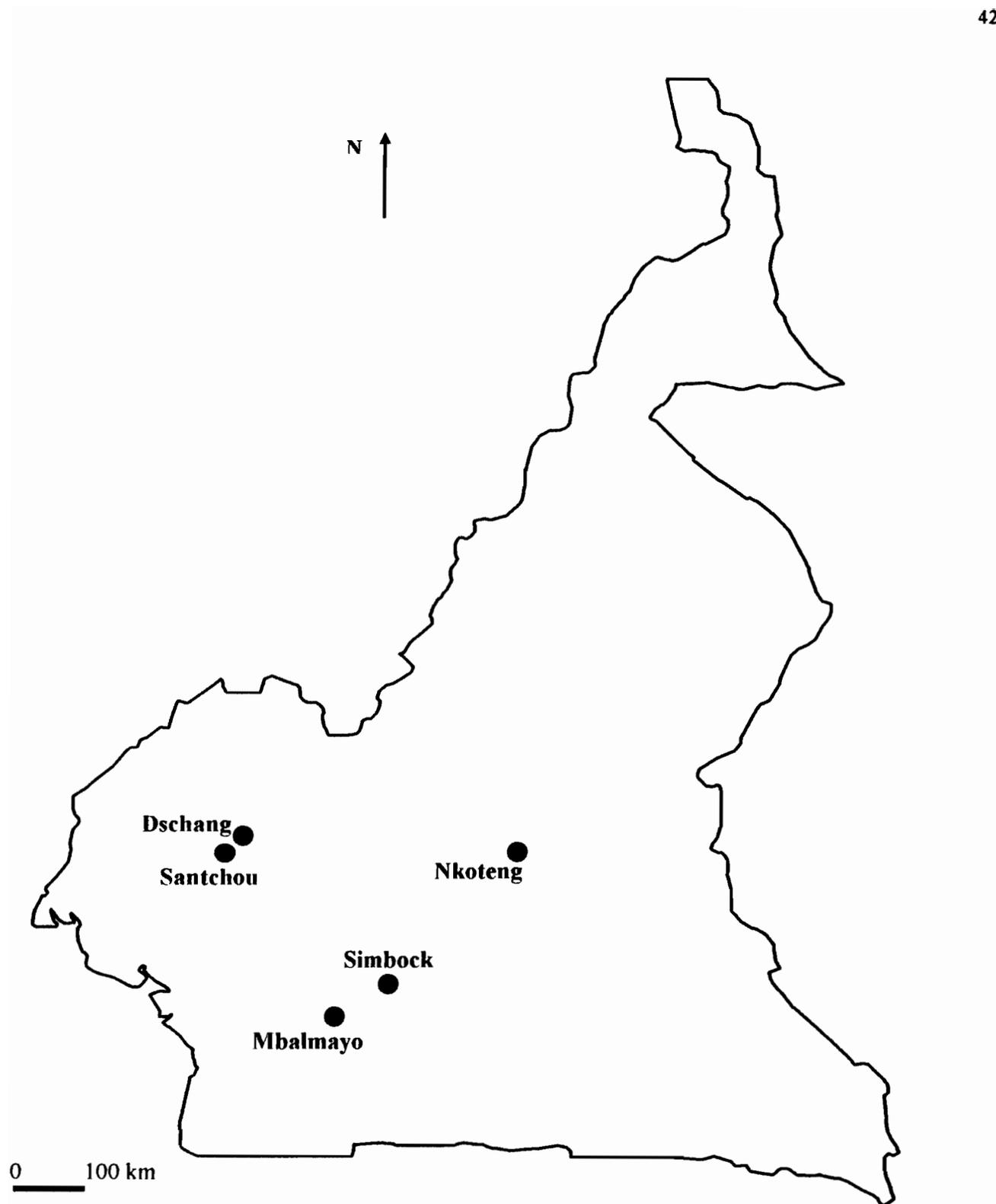


Figure 5 : Carte du Cameroun avec les différentes localités où les échantillons ont été collectés pour l'étude de la structure génétique d'*An. gambiae*

l'intérieur des cases et à l'extérieur, de préférence dans un endroit sombre. Une fois installés, ils relèvent leur pantalon pour permettre aux moustiques de se poser sur eux. A des intervalles de temps réguliers, ils éclairent leur jambes grâce à une torche électrique pour voir si un moustique s'est posé sur eux. Quand c'est le cas, ils le capturent à l'aide d'un tube à hémolyse qu'ils bouchent ensuite avec du coton. Ces tubes sont déposés ensuite dans des sacs correspondant à chaque tranche horaire (19-20h, 20-21h...). Les captureurs sont munis de réveils pour respecter les horaires. Les sacs sont collectés toutes les 2 heures et stockés au frais avant l'identification des moustiques.

2.2- Capture par pulvérisation intradomiciliaire d'insecticide

Cette méthode est utilisée afin de déterminer la proportion d'individus endophiles dans la population vectorielle par la collecte des spécimens se reposant à l'intérieur des cases. Elle permet aussi d'étudier les préférences trophiques des vecteurs ou encore d'obtenir des individus à l'état semi-gravide dont le caryotype sera déterminé par lecture des chromosomes polytènes. Cette méthode est réalisée entre 15h 30 et 18h car c'est l'intervalle de temps pendant lequel les spécimens semi-gravides d'*An. gambiae s.l.* au repos à l'intérieur des maisons peuvent être collectés (Service, 1993). Dans les maisons choisies, la surface interne du sol des chambres à coucher est recouverte de draps blancs afin de pouvoir récupérer tous les moustiques qui tomberont à la suite de la pulvérisation. Celle-ci se fait après fermeture de toutes les ouvertures et à l'aide d'un insecticide à base de pyréthre ou de pyréthrinoïdes. On laisse en moyenne 10 minutes après la pulvérisation pour permettre à l'insecticide d'agir. Les moustiques sont ensuite collectés sur les draps à l'aide de pinces fines et déposés dans des tubes à hémolyse pour identification.

D'autres méthodes ont été utilisées de temps à autre (piège OBET, pièges lumineux, capture par aspirateur) mais les deux indiquées ci-dessus ont été les principales lors de notre travail.

3- Identification morphologique

Les moustiques sont identifiés sous la loupe binoculaire à l'aide des clés de détermination taxonomique basées sur des critères morphologiques (Gillies & De Meillon, 1968 ; Gillies & Coetzee, 1987 ; Hervy *et al.*, 1998). Tous les anophèles sont stockés individuellement dans des tubes «eppendorf» de 1.5 ml contenant du silicagel (dessiccateur) et ensuite conservés au congélateur à -20°C.

Les spécimens semi-gravides obtenus à la suite de pulvérisations intradomiciliaires sont disséqués et les ovaires sont conservés dans du carnoy (1/3 d'acide acétique glacial + 2/3 d'éthanol absolu). Le repas de sang des spécimens gorgés est récupéré sur du papier Whattman pour l'analyse des préférences trophiques par ELISA selon Beier *et al.* (1988).

Les ovaires d'environ cinquante spécimens (capturés sur hommes) par journée d'enquête sont disséqués pour la détermination du taux de parturité.

4- Caractérisation moléculaire des spécimens

La caractérisation moléculaire des spécimens des différentes populations s'est imposée du fait de la complexité du système vectoriel, notamment le complexe *An. gambiae* au sein duquel les espèces sont morphologiquement indifférenciables. Cette caractérisation moléculaire a été réalisée pour identifier l'espèce ou la forme moléculaire à laquelle appartenait chaque spécimen. Pour cela, nous avons procédé à l'extraction de l'ADN des spécimens et procéder à des amplifications de cet ADN par PCR.

4.1- Extraction d'ADN des spécimens (Cornel & Collins, 1996)

Cette extraction est le fruit de plusieurs opérations qui se déroulent de la manière suivante:

Les moustiques préalablement conservés à -20°C , sont sortis sur la paillasse et en général 2 à 3 pattes sont récupérées pour être déposées dans un microtube de 1,5 ml. Ce tube stérile, contenant 50 μl de tampon de broyage dont la composition est présentée en annexe 4. Les pattes sont ensuite entièrement broyées à l'aide d'un piston de broyage qui est ensuite rincé avec 50 μl du tampon de broyage.

Le broyat de moustique est incubé pendant 30 minutes à 65°C dans un incubateur à sec, ceci pour dénaturer les nucléases. 13 μl d'acétate de potassium 8M (1M final) sont ajoutés à la solution qui est aussitôt bien homogénéisée grâce à un Vortex. Le nouveau mélange est incubé dans la glace pendant 30 minutes pour précipiter les débris cellulaires et les protéines dénaturées ou insolubles.

Une centrifugation des échantillons est effectuée à 14.000 tours/minutes pendant 15 minutes et à 4°C . Après cette centrifugation, le surnageant de chaque tube contenant l'ADN solubilisé est transféré dans un nouveau tube stérile de 1,5 ml.

200 μl d'éthanol absolu à froid sont ajoutés dans chaque tube, et après mélange, le tout est laissé sur la paillasse pendant au moins 5 minutes. A cette étape, la solution d'ADN peut être conservée pendant longtemps à -80°C mais, il y demeure encore des nucléases susceptibles de dégrader l'ADN.

Une nouvelle centrifugation des échantillons est réalisée à 14.000tours/minutes pendant 20 minutes pour précipiter l'ADN qui forme au fond du tube un culot peu visible. Après avoir délicatement enlevé l'éthanol, le culot est rincé avec 200 μl d'éthanol à 70%. Le culot est séché grâce à un sécheur sous vide (speed-vac).

L'ADN extrait est reconstitué dans 100 μl d'une solution de tampon TE (Tris/EDTA) ou dans de l'eau distillée stérile et ensuite laissé au réfrigérateur à 4°C pour au moins 12 heures pour la resuspension du culot. Cette solution d'ADN est ensuite conservée à -20°C .

4.2- Principe de la PCR

La PCR (Polymerase Chain Reaction) est une technique permettant d'amplifier *in vitro* des séquences d'ADN par répétition de réactions d'élongations en présence d'une ADN polymérase et d'amorces nucléotidiques spécifiques.

Le principe de l'amplification *in vitro* (figure 6) repose sur la répétition de trois processus :

i) La dénaturation des deux brins d'ADN à température élevée (environ 94°C) afin d'obtenir des molécules d'ADN monocaténaire,

ii) L'hybridation d'oligonucléotides (amorces) complémentaires d'une séquence de l'ADN monocaténaire cible (la température est alors ramenée à une valeur comprise entre 40°C et 65°C afin de permettre une bonne fixation des amorces),

iii) La réaction d'élongation par un ADN polymérase thermostable (la Taq polymérase) à partir des oligonucléotides, réalisée à la température optimale de 72°C.

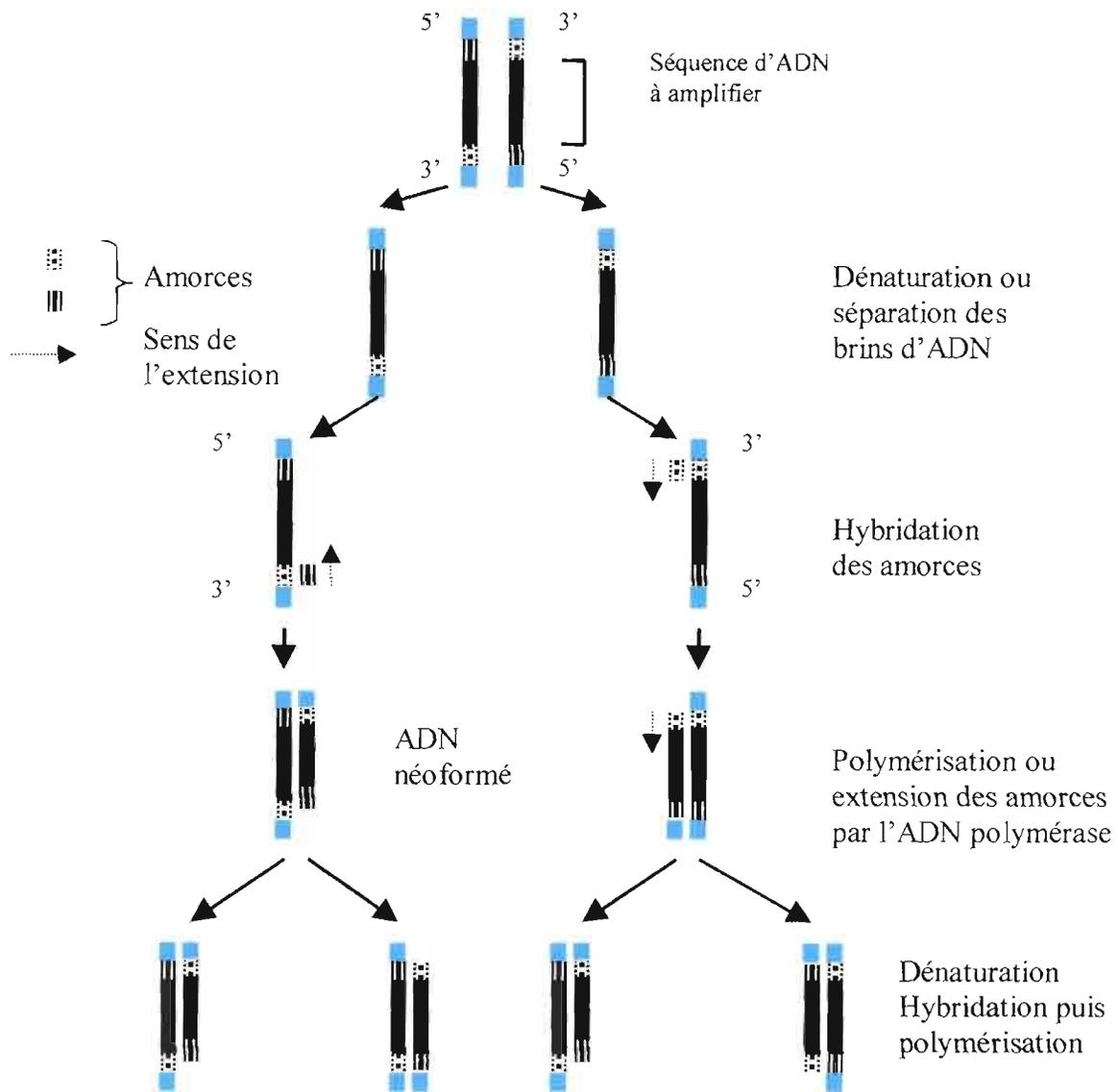


Figure 6 : Schéma du principe de la PCR

Les produits de ce premier cycle sont ensuite dénaturés par la chaleur. Les amorces s'hybrident à nouveau avec les brins d'ADN provenant du premier cycle d'amplification, chaque brin servant de matrice à la polymérase. A chaque cycle, le nombre de copies du fragment d'ADN est doublé : 2^n molécules sont ainsi obtenues après n cycles, soit par exemple 1 048 576 molécules après vingt cycles.

4.3- Identification des membres du complexe *An. gambiae s.l.*

L'identification des espèces du complexe *An. gambiae* se fait sur la base d'une PCR ciblant l'ADN ribosomal, mise au point par Scott *et al.* (1993). La réaction lors de cette PCR fait intervenir cinq amorces dont une dite «Universelle» (UN) et encore appelée amorce «sens» (Forward), désignée sur une séquence du gène codant pour la sous-unité 28S de l'ARN ribosomal qui est très conservée dans tout l'ordre des diptères. L'hybridation ayant lieu du côté 5' de cette séquence. Les quatre autres amorces appelées «antisens» (Reverse) sont spécifiques d'espèces et sont désignées dans les séquences polymorphes des espaces intergéniques (IGS) de l'ADNr. Les membres du complexe *An. gambiae* possèdent un ADNr constitué d'une collection d'unités répétées à plus de 500 copies (figure 7). Les fragments générés après la PCR sont de différentes tailles et permettent une identification aisée de chaque espèce.

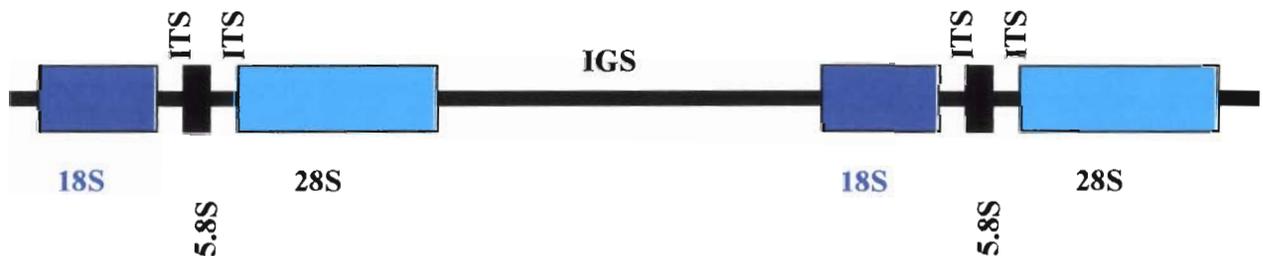


Figure 7 : Structure de l'ADNr des diptères

4.3.1- Déroulement de la PCR complexe *An. gambiae*

La PCR a été effectuée dans un volume final de 12,5 μ l de milieu réactionnel. Celui-ci était constitué de divers réactifs dont le tampon Taq 10X contenant 15mM de $MgCl_2$, une solution de dNTPs (constituée d'un mélange équimolaire de dATP, dCTP, de dTTP de dGTP), de 0.5 unité de Taq polymérase, d'un cocktail d'amorces (voir tableau 1) dont une dite «universelle» (UN) et 3 autres spécifiques de chaque espèce (*An. gambiae* (AG), *An. arabiensis* (AR) et *An. melas* (ML)), d'eau stérile, d'ADN extrait ou parfois quelques pattes introduites directement dans le mélange réactionnel. Une présentation plus détaillée de ce protocole est donnée à l'annexe 4. La réaction se déroule grâce à un thermocycler (Perkin Elmer 9700) pouvant prendre 96 échantillons.

Tableau 1: Amorces utilisées pour la PCR complexe *An. gambiae*

Amorces	Séquence des amorces	Taille des bandes
Universelle (UN)	GTG TGC CCC TTC CTC GAT GT	
<i>An. gambiae</i> (GA)	CTG GTT TGG TCG GCA CGT TT	390 pb
<i>An. arabiensis</i> (AR)	AAG TGT CCT TCT CCA TCC TA	315 pb
<i>An. melas</i> (ME)	TGA CCA ACC CAC TCC CTT GA	464/466 pb
<i>An. quadriannulatus</i> (QD)	CAG ACC AAG ATG GTT AGT AT	153 pb

4.3.2- Electrophorèse des échantillons issus de la PCR

La méthode d'électrophorèse a été utilisée pour révéler les amplifiats issus de la PCR. Un gel d'agarose à 1,5% est préparé en pesant 1,5 g de poudre d'agarose dans 100 ml de tampon TBE (Tris Borate EDTA) à 1 M. La solution est portée à ébullition dans un four à micro-ondes pour environ 2 minutes puis 5µl de bromure d'éthidium (BET) y est introduit. Après refroidissement, la solution est coulée dans un moule d'électrophorèse avec un peigne pour former les puits.

Après la polymérisation du gel, les échantillons sont déposés dans les différents puits après avoir été mélangés à une solution de bleu de bromophénol/xylène cyanol (marqueurs de migration). Cette solution contient du glycérol qui augmente la densité du mélange de dépôt. Un marqueur de poids moléculaire (100 paires de bases) est introduit dans le gel pour une estimation de la taille des différentes bandes, ainsi que des témoins de chaque espèce, plus un témoin négatif. L'électrophorèse se déroule sous une tension de 120 volts à 100 mA pendant environ 40 minutes.

Le gel à la fin de l'électrophorèse est porté sous la lumière ultraviolette pour la détection des différentes bandes ayant été amplifiées. L'espèce de chaque échantillon est déterminée en fonction des témoins d'espèces et de l'échelle de taille (voir figure 8).

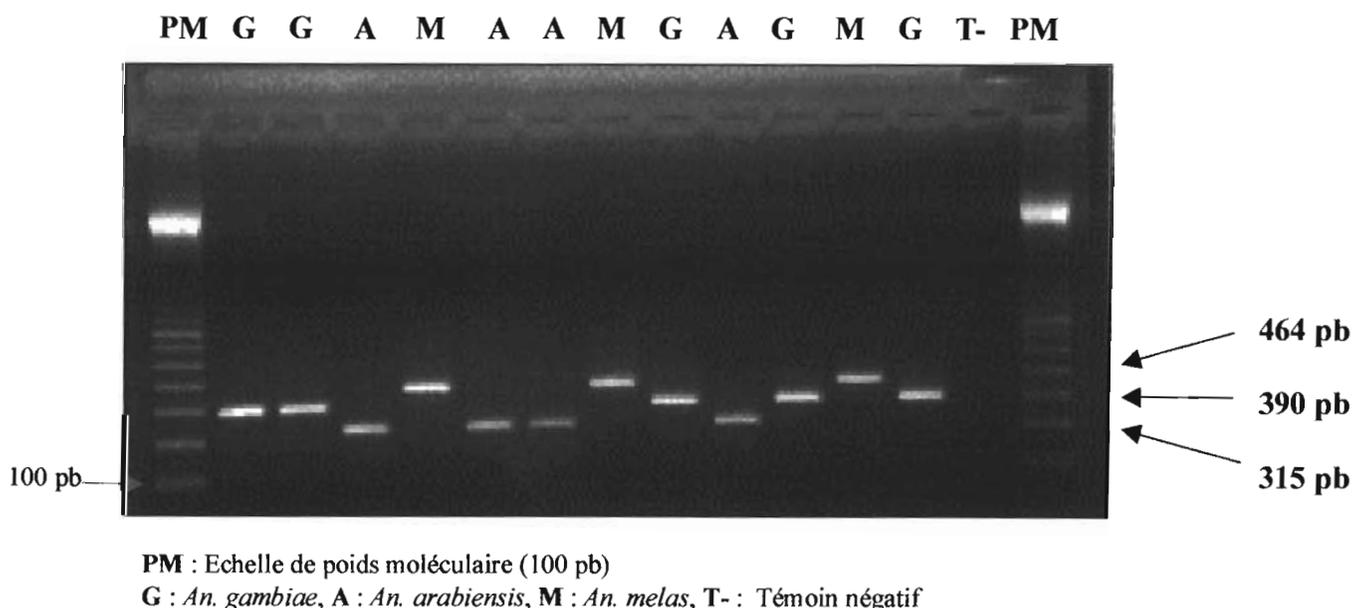


Figure 8 : Profil de migration sur gel d'agarose des spécimens du complexe *An. gambiae* après PCR d'identification selon Scott *et al.* (1993).

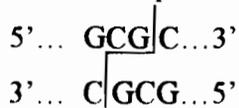
4.4- Détermination de la forme moléculaire des spécimens d'*An. gambiae* s.s.

Les spécimens identifiés comme *An. gambiae* s.s. ont été testés par la PCR pour la détermination de leur forme moléculaire. Pour ce faire, nous avons utilisé deux méthodes basées sur la PCR, l'une comportant une étape de digestion enzymatique du fragment amplifié par RFLP et l'autre se faisant de façon directe. Ces deux méthodes ont été développées par Favia *et al.*, (1997 ; 2001).

4.4.1- Méthode de PCR-RFLP

Cette méthode se fait en deux étapes dont une PCR pour amplifier la séquence cible et une deuxième pour digérer le fragment amplifié au niveau des sites de restriction de l'endonucléase HhaI. La PCR se fait à l'aide de deux amorces décrites par Paskewitz & Collins (1990), A0 désignée dans la séquence 28S et Ag1.3 désignée dans la région IGS. La PCR se déroule selon le même principe que celui décrit précédemment. Le protocole plus détaillé est donné à l'annexe 3. L'amplification fournit un fragment de 1.3 kb.

Une fois que l'amplification de la bande de 1.3kb est confirmée par migration des amplifiats sur un gel d'agarose à 2,5%, la digestion enzymatique est réalisée en utilisant l'enzyme de restriction HhaI (*Haemophilus haemolyticus*). Cette endonucléase coupe le fragment au niveau des sites de restriction présentés ci-dessous.



La digestion enzymatique se fait après la préparation d'un premier mélange de 7µl contenant 1,25 unités d'enzyme, le tampon d'enzyme et de l'eau distillée. La restriction est engagée après l'addition de 13µl d'ADN issu de l'amplification par PCR, dans chaque tube. La durée de la digestion est de 3 heures à 37°C. Les bandes sont visualisées à la fin de la digestion sur un gel d'agarose à 2% sous UV. Les spécimens appartenant à la forme moléculaire M présentent 3 bandes à 100, 300 et 800 pb alors que les spécimens de forme S présentent 4 bandes dont 2 situées entre 100 et 200 pb, une à 300 pb et une autre à 600 pb.

4.4.2- Méthode par PCR uniquement

Cette PCR a été mise au point après l'analyse d'une séquence de 2,3 kb d'un segment d'ADN situé dans la région des IGS (voir figure 9) chez des spécimens de 3 formes chromosomiques (Mopti, Savanna et Bamako). Cette analyse a révélé que la forme Mopti différait de Savanna et de Bamako par un total de dix substitutions nucléotidiques alors que ces deux dernières formes étaient identiques (Favia *et al.*, 2001). Au Mali, la forme moléculaire M d'*An. gambiae* correspond au profil de la séquence obtenue pour Mopti, alors que la forme moléculaire S correspond au profil de Savanna et Bamako. Des amorces spécifiques de chaque forme moléculaire

correspond au profil de Savanna et Bamako. Des amorces spécifiques de chaque formes moléculaires ont été mises au point sur la base de cette variabilité pour pouvoir identifier chacune des deux formes moléculaires M et S (Tableau 2).

Tableau 2 : Amorces utilisées dans la PCR pour identification des formes moléculaires d'*An. gambiae*

Amorces	Séquence des amorces 5' à 3'	Taille des bandes
R5 ou A0	GCC AAT CCG AGC TGA TAG CGC	
R3 ou Agd1	CGA ATT CTA GGG AGC TCC AG	
Mopint (Forme M)	GCC CCT TCC TCG ATG GCA T	727 pb
B/Sint (Forme S)	ACC AAG ATG GTT CGT TGC	475 pb

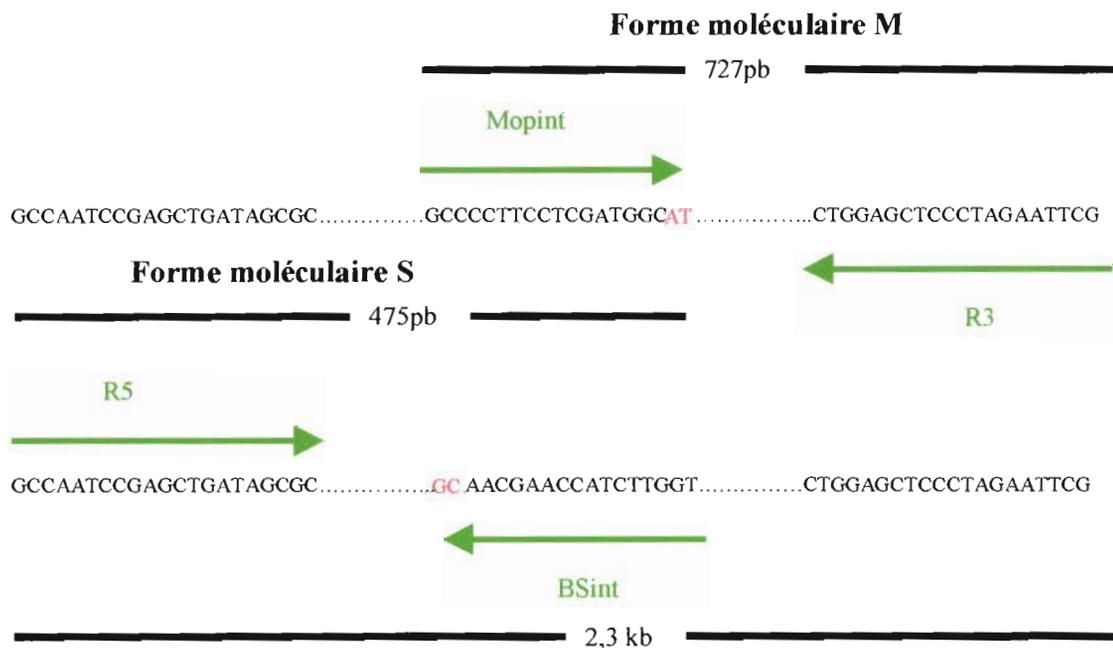


Figure 9 : Localisation des amorces utilisées pour la détermination des formes moléculaires d'*An. gambiae* sur un fragment de 2,3 kb dans la région des IGS de l'ADNr d'après Favia *et al.* (2001)

Le déroulement de la PCR est présenté plus en détail à l'annexe 3.

Les produits d'amplification sont visualisés sur un gel d'agarose à 1,5% sous UV après une migration d'environ 30 minutes. Les spécimens de forme moléculaire S présentent une bande caractéristique à 475pb alors que la bande diagnostic de la forme M se situe à 727pb (voir figure 10).

PM M M M M S M S S M M M M M M M M M M S

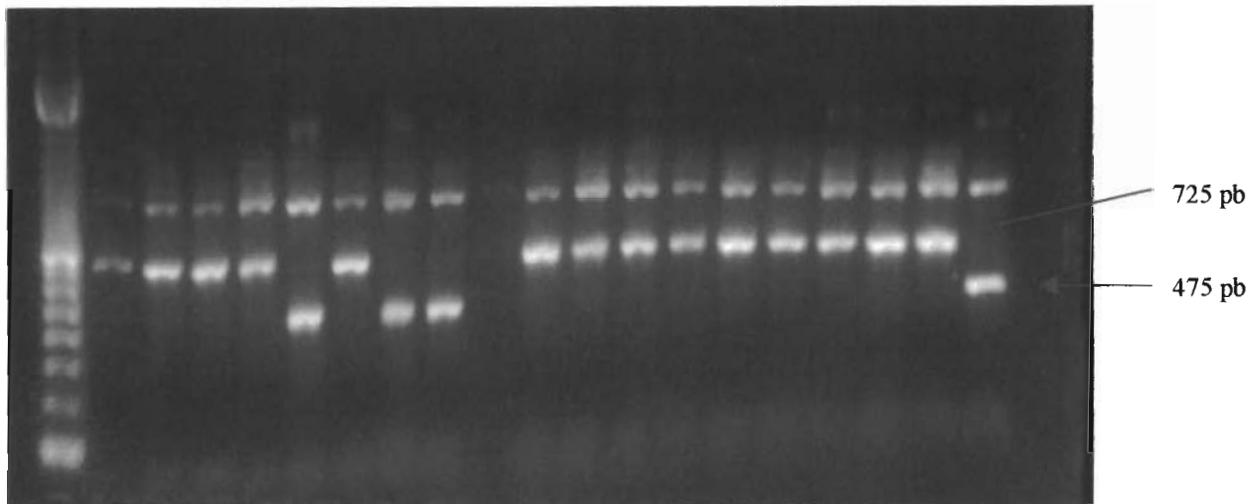


Figure 10 : Profil de migration sur gel d'agarose des spécimens de formes moléculaires M et S d'*An. gambiae*

5- Détermination de la forme chromosomique

Le caryotype des spécimens (arrangement chromosomique des inversions) d'*An. gambiae* a été déterminé par la lecture des chromosomes polytènes (figure 11) contenus dans les ovaires à l'état semi-gravide. La préparation des chromosomes s'est faite sur des lames selon Coluzzi & Sabatini (1968) et Hunt (1973). Les inversions paracentriques ont été déterminées après l'examen des chromosomes au microscope à contraste de phase (x160 à x400) selon Coluzzi *et al.* (1979) avec en référence, la carte de chromosomes polytènes standard du complexe *An. gambiae*. La récolte des ovaires et leur conditionnement ont été effectués par nos soins sur le terrain. Ces déterminations ont été réalisées en collaboration avec le Dr Vincenzo Petrarca et le Dr Alessandra Della Torrè du laboratoire de l'«Istituto di Parassitologia Universita 'la Sapienza'» à Rome en Italie où les préparations et la lecture des chromosomes ont eu lieu. La nomenclature des inversions suit avec quelques modifications mineures, un système largement utilisé chez la drosophile par Wasserman (1963) et par Carson *et al.* (1967). Toute inversion recensée est désignée par une lettre minuscule indépendamment des autres bras chromosomiques à l'exemple de 2Rb, 2Rbcd, 2La, etc (Coluzzi *et al.*, 1979). Le symbole d'hétérozygotie (a/+) signifie que l'inversion «a» est polymorphe et le signe «+» est utilisé pour indiquer l'arrangement chromosomique sans inversion.

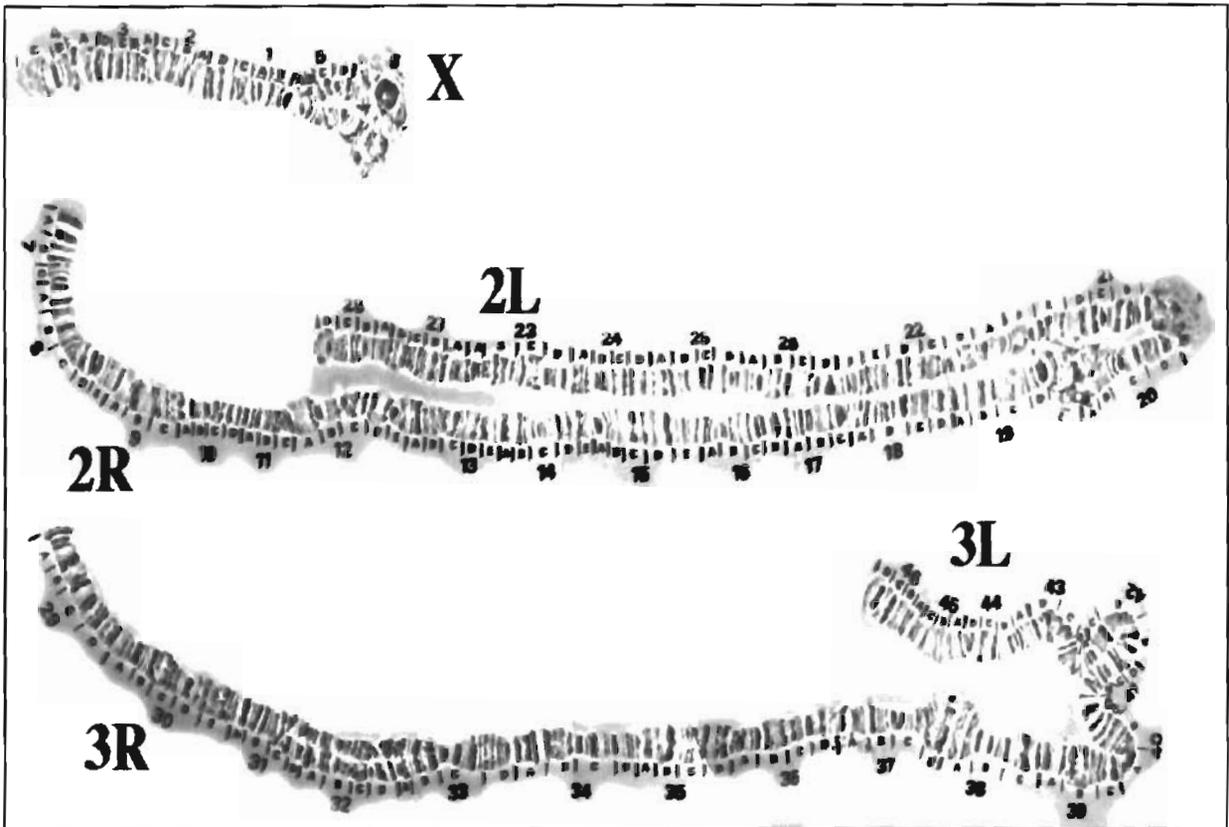


Figure 11 : Chromosomes polytènes d'*An. gambiae*

6- Détermination des paramètres entomologiques de la transmission

Les techniques immuno-enzymatiques (ELISA) ont été utilisées pour déterminer le taux d'infection des moustiques et leur préférence trophique. Ces deux paramètres ont été utilisés pour obtenir d'autres paramètres entomologiques notamment le taux d'inoculation entomologique.

6.1- Détermination de la préférence trophique par ELISA repas de sang

6.1.1- Principe

La détermination de la source des repas sanguins est basée sur des caractères sérologiques du plasma (recherche d'IgG). Les anticorps spécifiques d'hôtes potentiels de moustiques (homme, bœuf, cheval, porc...) sont mis en contact avec le sang que ces moustiques ont absorbé. Ces anticorps sont marqués par une enzyme (la peroxydase) et en présence du substrat de celle-ci, une réaction colorée révèle la présence des IgG recherchés (figure 12). La technique utilisée ici a été développée par Beier *et al.* (1988).

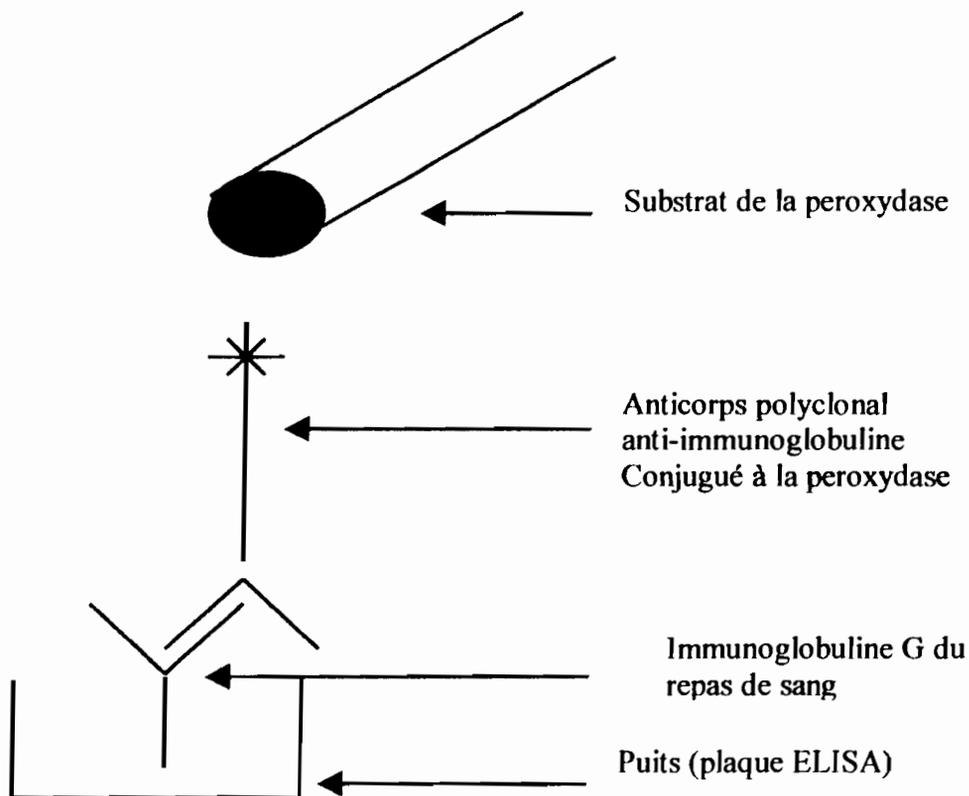


Figure 12 : Schéma du principe de l'ELISA repas de sang

6.1.2- Déroulement de la manipulation

Les rondelles de papier Whatmann (sur lesquelles les repas de sang des moustiques après sectionnement de leurs abdomens ont été préalablement déposés sur le terrain), ont été découpées et déposées dans des tubes à hémolyse contenant 0,8 ml de tampon PBS. La solution est bien homogénéisée et laissée au repos pour au moins 1h. Les plaques sont sensibilisées par le dépôt de 50 μ l des solutions issues de l'homogénéisation précédente, dans les différents puits d'une plaque en raison d'un échantillon par colonne. Deux colonnes sont réservées aux témoins dont une pour le contrôle négatif constitué par 50 μ l de tampon PBS dans chaque puits et l'autre pour les contrôles positifs constitués de 50 μ l de sérum homologue dilué au 1/100 dans chaque puits. Cette étape permet aux IgG de se fixer sur la paroi des puits. Les plaques sont recouvertes et incubées durant 3 heures à température ambiante ou toute la nuit à +4°C.

Les anticorps spécifiques d'espèces et marqués à la peroxydase sont ensuite préparés avec les sérums hétérologues pour limiter les réactions croisées. Pour cela, différentes dilutions sont effectuées. Les plaques sont lavées deux fois à l'aide du PBS/Tween 20. 50 μ l de solutions d'anticorps de chaque espèce sont déposés dans les puits de la colonne correspondante puis incubés pendant 1 heure. Les anticorps anti-IgG se fixent sur les IgG correspondants. 4 lavages successifs au PBS/Tween 20 s'en suivent afin d'éliminer les anticorps anti-IgG non fixés. 100 μ l du substrat de la peroxydase (composé d'ortho-tolidine, de N,N-diméthyl formamide, de tampon citrate et de

H₂O₂) sont ensuite distribués dans chaque puits. Les plaques sont incubées à l'obscurité pendant 30 minutes puis les puits positifs donnent une coloration bleue. La réaction est arrêtée en ajoutant 50µl d'acide sulfurique 4N dans chaque puits. Les densités optiques sont déterminées après lecture au spectrophotomètre à 450 nm.

6.2- Estimation du taux d'infestation par ELISA

6.2.1- Principe

Cette technique consiste à détecter l'antigène circumsporozoïtique (CSP) qui est la principale protéine de la surface des sporozoïtes (l'un des stades du *Plasmodium* chez les moustiques), grâce à des anticorps monoclonaux. Il s'agit d'un ELISA «sandwich». L'antigène circumsporozoïtique lorsqu'il est présent, est fixé par un anticorps monoclonal de capture anti-circumsporozoïte fixé sur les parois de la plaque (Wirtz *et al.*, 1987). Le complexe antigène-anticorps formé est révélé par un anticorps monoclonal anti-CS conjugué à une peroxydase. La technique se fait en deux étapes (figure 13). Dans la première, un ELISA «screen» est procédé pour déterminer les moustiques infectants sans tenir compte de l'espèce plasmodiale. Par contre, dans la deuxième étape dite ELISA «monospécifique», seuls les échantillons ayant été positifs lors de l'ELISA «screen» sont testés pour déterminer l'espèce plasmodiale responsable de l'infection. Cette procédure de sélection des échantillons positifs pour l'étape de détermination de l'espèce plasmodiale entraîne un gain en temps et en produits de laboratoire puisque le nombre d'échantillons se trouve fortement réduit.

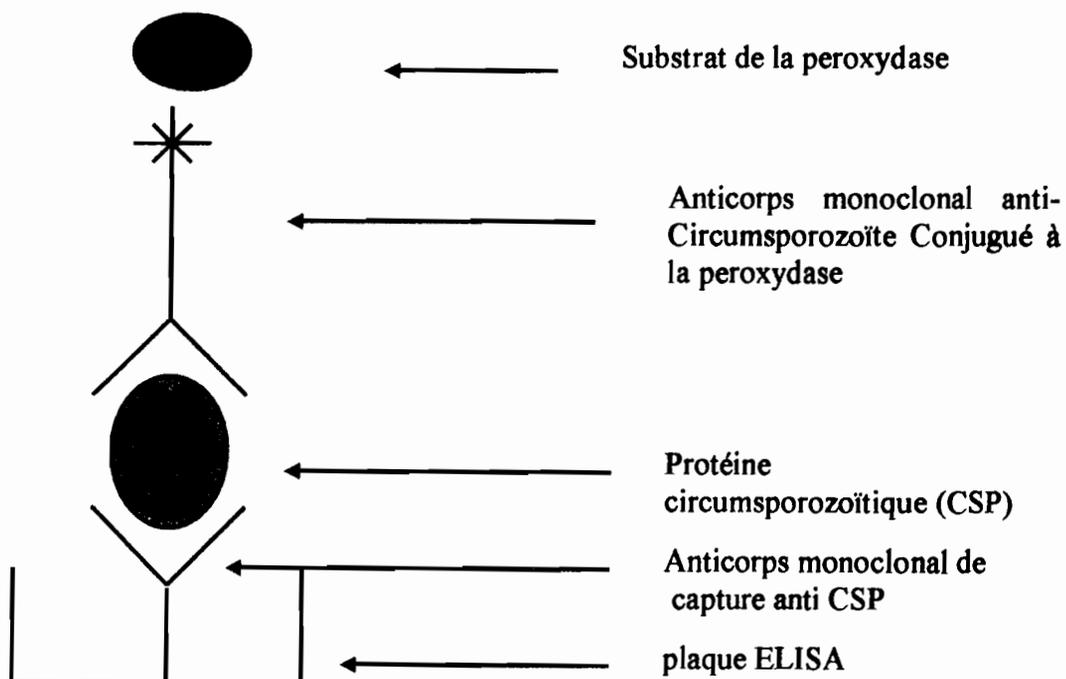


Figure 13 : Principe de l'ELISA Circumsporozoïtique (CSP)

6.2.2- Déroulement de la manipulation

La tête et le thorax de chaque moustique sont découpés et déposés dans des tubes numérotés de 1.5ml. Pour faciliter le broyage, une solution de NP 40 est ajoutée dans chaque tube (il s'agit d'un détergent qui facilite la destruction des cellules par hydrolyse des protéines membranaires). Le broyage se fait dans 2 fois 190µl de tampon de broyage (BB pour «Blocking Buffer»). Les broyats sont conservés à -20°C. 150µl d'un mélange des 3 anticorps monoclonaux pour chaque espèce plasmodiale (*P. falciparum*, *P. malariae* et *P. ovale*) sont introduits dans chaque puits. La plaque est incubée toute la nuit sur la paillasse. Le lendemain, la plaque est vidée sans lavage puis 200µl de tampon de broyage sont versés dans chaque puits, ceci pour saturer les sites de fixation des anticorps de capture. La plaque est incubée ensuite pendant 1h et vidée après sans lavage. 50µl de broyat de chaque moustique sont introduits dans les puits correspondants. Dans la colonne des contrôles négatifs, c'est 50µl de tampon BB qui sont déposés alors que pour les puits des témoins positifs, c'est 50µl de protéine CSP (2pg/µl pour *P. falciparum*, 5ng/µl pour *P. malariae*) de chaque espèce plasmodiale qui sont déposés. La plaque est incubée pour 2h pendant lesquelles l'antigène CS se fixe sur les anticorps de capture anti-CS. Les plaques sont ensuite lavées 2 fois au PBS/Tween 20. Une solution d'anticorps monoclonaux conjugués à la peroxydase est préparée. Elle est constituée des anticorps de toutes les 3 espèces plasmodiales ainsi que du tampon BB. Un volume de 150µl de ce mélange est déposé dans chaque puits et après une heure d'incubation, les plaques sont lavées 4 fois au PBS/Tween 20. Le substrat de la peroxydase qui est un mélange de tampon citrate, d'eau oxygénée, d'ortho-tolidine, de N, N diméthylformamide est déposé à raison de 100µl par puits. Les plaques sont incubées pendant 30 minutes à l'obscurité. La réaction est arrêtée après ajout de 50µl d'acide sulfurique 4N. La densité optique de chaque échantillon est déterminée par lecture de la plaque au spectrophotomètre à 450 et 620nm. Les échantillons sont considérés positifs quand la valeur de leur D.O. est supérieure à 2 fois la moyenne de la D.O. des témoins négatifs.

La technique ELISA dite monospécifique, utilisée pour la détermination de l'espèce plasmodiale responsable de l'infection, est semblable à la précédente. La différence se situe au fait que pour le monospécifique chaque anticorps monoclonal correspondant à chaque espèce est testé individuellement contrairement à l'ELISA «screen» où ils sont tous mélangés.

7- Détermination du génotype des spécimens aux différents locus ADN microsatellites

7.1- Amplification des différents locus

L'amplification des allèles aux différents locus sélectionnés (voir le deuxième chapitre des résultats de ce travail pour plus de détails sur la démarche de sélection des locus microsatellites que

nous avons utilisés) s'est faite par PCR dans un volume final de 12,5µl. Le protocole complet de cette PCR est présenté en détail à l'annexe 3.

7.2- Détermination des génotypes

7.2.1- Préparation des gels de polyacrylamide

L'électrophorèse des échantillons amplifiés par la PCR a été réalisée sur un gel de polyacrylamide vertical à 15%. A cet pourcentage ce gel permettait une bonne séparation de bandes correspondantes à des allèles dont la taille diffère de 2 paires de bases.

La préparation de deux gels à 15% se fait en prélevant les solutions suivantes :

18,75 ml de solution d'acrylamide/bis acrylamide à 40%

26,25 ml d'eau distillée

5 ml de tampon Tris-Borate-EDTA (TBE) 10X pH 8,9

Cette solution est dégazée à l'aide d'une pompe sous vide pour éviter que les molécules d'oxygène ne nuisent à la polymérisation par formation d'oxydes avec les radicaux libres. Les réactifs suivants y sont ajoutés :

225µl de persulfate d'ammonium à 10% qui fournit les radicaux pour la réaction,

100µl de N,N,N',N' Tetra-méthyl-éthylène-diamine (TEMED) qui catalyse la réaction de polymérisation.

Le gel est ensuite coulé entre deux plaques de verre montées sur un support de coulage. Un peigne de 24 puits y est déposé. La polymérisation se déroule pendant environ 40 minutes.

7.2.2- Migration

Les gels polymérisés sont fixés sur un support et introduits dans la cuve d'électrophorèse. Le tampon de migration est le TBE 1X (pour la composition, voir annexe 5). Les échantillons sont déposés dans les différents puits à l'aide d'une solution constituée de bleu de bromophénol et de xylène cyanol. Chaque gel comptait 22 échantillons dont un marqueur de poids moléculaire de 100 pb, 3 témoins servant d'échelle de taille. La migration se déroule à 550 volts/595 mA/300W et de durée variable en fonction de la taille des allèles amplifiés (2 à 3h 30').

7.2.3- Révélation des génotypes sur gel de polyacrylamide

Le génotype des différents échantillons est révélé grâce à une coloration des gels au nitrate d'argent selon le protocole de Sanguinetti *et al.* (1994).

La révélation se fait grâce à 3 solutions dont les compositions sont les suivantes :

-Solution de fixation

Pour deux gels,

-Solution de fixation

Pour deux gels,

94,5 ml d'éthanol à 95%

4,5 ml d'acide acétique glacial

Eau distillée quantité suffisante pour (q.s.p) 900 ml

-Solution de coloration au nitrate d'argent (AgNO₃) 0,2%

0,8 g d'AgNO₃

400 ml de solution de fixation

-Solution de révélation 3% NaOH, 0,5% formaldéhyde à 37%

12 g de NaOH

5,4 ml de formaldéhyde à 37%

Eau distillée q.s.p 400ml

La révélation proprement dite commence par une incubation des gels dans la solution de fixation pendant 3 minutes. Les gels sont ensuite baignés dans la solution de coloration pour environ 5 minutes à l'obscurité sur agitateur. Après cette étape, les gels sont lavés une fois pendant 20 s et une fois pendant 1 minute pour éliminer l'excès d'AgNO₃ qui ne s'est pas fixé sur les fragments d'ADN du gel. Les gels sont ensuite incubés dans la solution de révélation jusqu'à apparition des bandes. Une fois que les bandes sont bien visibles, les gels sont à nouveau baignés dans la solution de fixation pour 2 minutes et ensuite ils sont introduits dans de l'eau distillée. La détermination des génotypes des différents spécimens se fait en déposant le gel sur un transilluminateur qui grâce à sa luminosité vive, rend plus visibles les différentes bandes d'ADN du gel. A l'aide du poids moléculaire mais surtout du témoin servant d'échelle de taille pour le locus, les allèles sont identifiés et numérotés par ordre croissant du bas vers le haut du gel. Les individus homozygotes sont ceux présentant une bande et les hétérozygotes sont ceux présentant deux bandes.

Les gels sont ensuite déposés sur papier buvard et séchés grâce à un sécheur de gel et une pompe sous vide. Une fois séchés, les gels sont conservés dans des chemises et peuvent aisément être consultés pour toute vérification ultérieure des génotypes.

7.2.4- Construction des échelles de taille pour chaque locus

Pour une bonne détermination des génotypes des spécimens de chaque population et pour une comparaison entre les populations, il a fallu construire des échelles de taille servant de référence pour chaque locus microsatellite.

Douze spécimens de terrain ont servi à la construction de ces échelles de taille. Ils ont été amplifiés pour chaque locus et révélés sur gel de polyacrylamide coloré au nitrate d'argent. Toutes les bandes représentant les différents allèles du locus ont été découpées du gel et introduites dans

des tubes individuels stériles de 1,5 ml contenant 400 µl de tampon Tris-EDTA (TE). Les tubes ont été agités et incubés durant toute la nuit à 37°C sur un bain sec. Le jour suivant, chaque tube a été agité et centrifugé pendant 3 minutes à 1000 rpm. Les surnageants ont été ensuite récupérés et déposés dans des tubes stériles de 2 ml dans lesquels 100 µl d'acétate d'ammonium (7 M) et 1 ml d'éthanol absolu ont été introduits. Le contenu de chaque tube a été bien mélangé et porté ensuite à -70°C pendant 15 minutes. A la suite de cette étape, les tubes ont été centrifugés à 14000 rpm pendant 5 minutes et le surnageant a été éliminé. Le culot de chaque tube a été lavé avec 500 µl d'éthanol à 70% par une centrifugation pendant 5 minutes et ensuite séché. L'ADN de chaque tube a été re-solubilisé dans 50 µl de tampon TE à 56°C pendant toute la nuit. Les solutions d'allèles obtenues pour chaque locus sont mélangées et diluées pour fournir la meilleure amplification afin de constituer l'échelle d'allèles du locus.

8- Analyse des données

Les différentes données collectées ont été analysées pour répondre à des questions précises. Plusieurs programmes informatiques ont été utilisés à cet effet. Les données générées par les locus microsatellites ont été analysées par le logiciel GENEPOP 3.2 (Raymond & Rousset, 1995a), mais aussi à l'aide de ARLEQUIN v1.1 (Schneider *et al.*, 1997).

Nous décrivons maintenant brièvement les principaux tests que nous avons utilisés dans notre étude.

8.1- Test de conformité à l'équilibre d'Hardy-Weinberg

Les données de fréquences alléliques et génotypiques obtenues de chaque population ont été analysées pour vérifier leur conformité à l'équilibre de Hardy-Weinberg. Nous avons utilisé les tests exacts fournis par le logiciel GENEPOP 3.2 (Raymond & Rousset, 1995a). L'hypothèse nulle H_0 est l'union aléatoire des gamètes. Les fréquences génotypiques observées sont comparées à celles estimées à partir des fréquences alléliques sous l'hypothèse de Hardy-Weinberg par un test exact de Fisher (Haldane, 1954). Lorsque le nombre d'allèles par locus est supérieur à 4, la méthode des chaînes de Markov (Guo & Thompson, 1992), est utilisée pour estimer sans biais la probabilité associée à l'hypothèse nulle H_0 . Pour chaque test, plusieurs indices sont fournis pour chaque locus. Nous citons la valeur de la probabilité («P-value»), la valeur de l'erreur standard et deux estimations de Fis dont la première est celle de Weir & Cockerham (1984) et la deuxième est calculée selon Robertson & Hill (1984). Un test global est aussi effectué sur l'ensemble des locus selon la méthode de Fisher, supposant l'indépendance des locus. Une valeur significative avec une probabilité $P < 0,05$ indique que H_0 n'est pas respectée au moins une fois sur l'ensemble des données. Les tests exacts sont plus appropriés que le test classique du χ^2 pour les locus polymorphiques tels que les microsatellites car ils présentent généralement un grand nombre

d'allèles dont certains sont représentés en très faibles fréquences (Chakraborty & Zhong, 1994 ; Rousset & Raymond, 1995b).

Le test séquentiel de Bonferroni (Holm, 1979) a été réalisé pour corriger la valeur de la probabilité P en prenant en compte le nombre de tests effectués simultanément.

8.2- Déséquilibre de liaison entre différents paires de locus

Ce test évalue l'indépendance entre des locus. L'hypothèse nulle est ainsi formulée : «les génotypes à un locus sont indépendants des génotypes à un autre locus». Un test de probabilité (ou test exact de Fisher) est réalisé à l'aide des tables de contingence en utilisant la chaîne de Markov (Guo & Thompson, 1992).

8.3- Test de différenciation entre populations

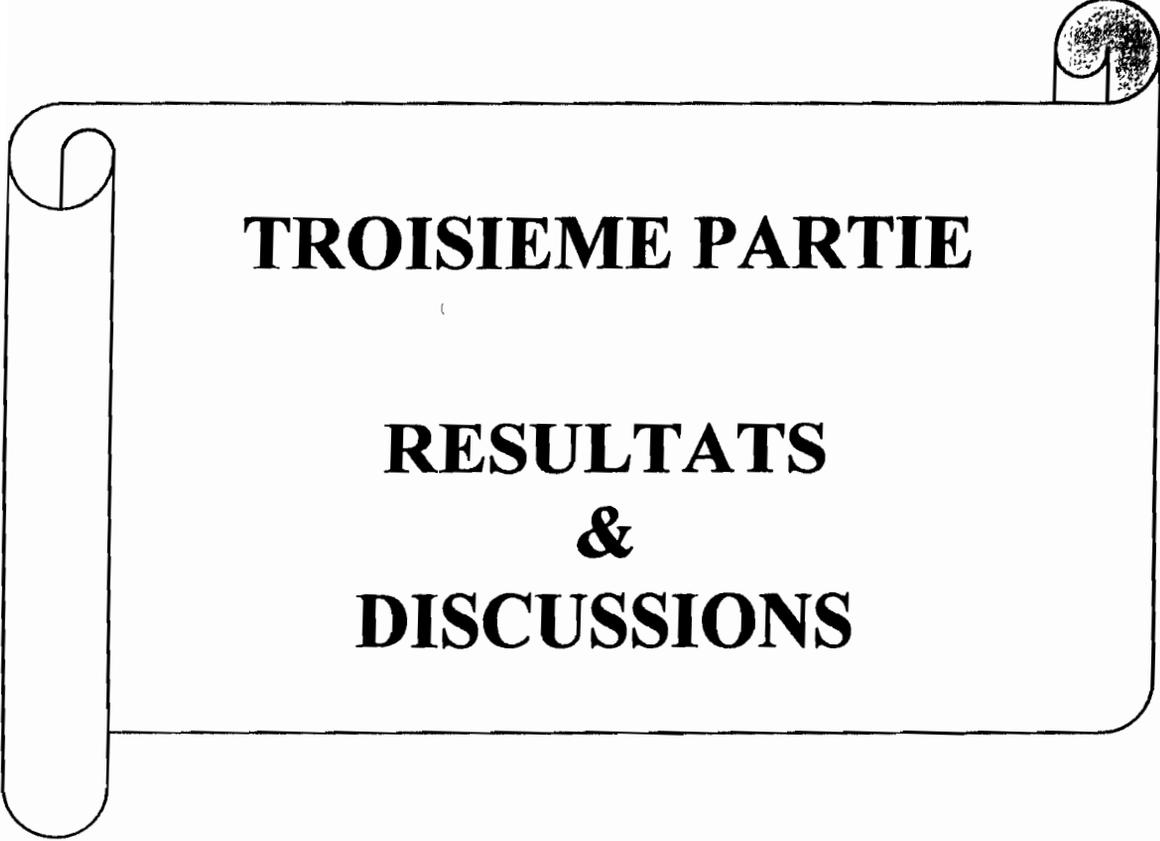
Un test de différenciation génique basé sur la distribution des allèles dans les différentes populations est réalisé. L'hypothèse nulle est la suivante : «la distribution allélique est identique entre les populations». Pour chaque locus, ce test est réalisé à l'aide de tables de contingence. Une valeur de probabilité P (test exact de Fisher) est estimée pour chaque locus et pour toutes les paires de populations. L'homogénéité de la distribution des fréquences alléliques est vérifiée à l'aide du test exact de Fisher utilisant les chaînes de Markov (Raymond & Rousset, 1995a et b).

Un autre test qui a été réalisé est celui de la différenciation génotypique, basé sur la distribution des génotypes à travers toutes les populations. L'hypothèse nulle H_0 testée est ainsi formulée : «la distribution des génotypes est identique entre les populations». Le test G de Goudet *et al* (1996) est utilisé pour vérifier l'homogénéité de la distribution des génotypes entre populations.

8.4- Les statistiques F

La différenciation génétique entre les populations a été estimée grâce aux statistiques F (Wright, 1978). Les indices de fixation F_{is} , F_{st} et F_{it} ont été estimés selon Weir & Cockerham (1984) soit sur l'ensemble des populations, soit pour chaque paire de populations, pour chaque locus indépendamment, ainsi que sur l'ensemble des locus. La valeur de F_{st} est testée à l'aide du test exact d'homogénéité de la distribution des génotypes entre les populations indiqué ci-dessus (Goudet *et al.*, 1996).

La variation de la distribution des fréquences alléliques entre les populations peut être mesurée en calculant la valeur de F_{st} qui est la variance entre les fréquences alléliques standardisées par la fréquence allélique moyenne parmi les populations. Les valeurs de F_{st} sont comprises entre 0 et 1, les plus élevées indiquant une plus forte différenciation entre les populations.



TROISIEME PARTIE

RESULTATS & DISCUSSIONS

CHAPITRE I

Caractérisation moléculaire et
cytologique des formes moléculaires
M et S d'*An. gambiae* au Cameroun et
leur implication dans la transmission
du paludisme

I.1- Rappel de la problématique

Les études antérieures conduites sur *An. gambiae* indiquent une forte complexité génétique au sein de cette espèce sur toute son aire de répartition géographique. L'existence de cinq formes chromosomiques et de deux formes moléculaires au sein de cette espèce contribue à rendre difficile la détermination des facteurs responsables de cette complexité génétique. La subdivision d'*An. gambiae* en cinq formes chromosomiques à savoir Forest, Savanna, Mopti, Bamako et Bissau a fait l'objet de nombreuses études visant à déterminer leur répartition géographique, leur degré d'isolement génétique voire leur rôle épidémiologique (Coluzzi *et al.*, 1985 ; Touré *et al.*, 1998). Ces études ont été pour la plupart conduites en Afrique de l'Ouest. Par contre, la subdivision d'*An. gambiae* en deux formes moléculaires est plus récente et les études spécifiques à ce sujet sont moindres. En effet, leur distribution géographique ainsi que les facteurs qui la déterminent ne sont pas encore bien établis tout comme le rôle épidémiologique de chacune des deux formes moléculaires dans la transmission du paludisme. La différence que présentent les formes M et S au niveau de la sensibilité aux insecticides notamment aux pyréthrinoïdes (le gène *kdr* étant plus présent chez la forme S que chez la forme M (Chandre *et al.*, 1999)), impose une bonne connaissance de la distribution géographique de chacune de ces formes moléculaires afin de mieux envisager des mesures de contrôle efficaces contre elles.

Pour une plus grande compréhension des facteurs déterminant la complexité génétique chez *An. gambiae*, de l'implication épidémiologique de cette complexité génétique, nous avons entrepris une étude de la distribution géographique des formes moléculaires M et S et celles des formes chromosomiques d'*An. gambiae* au Cameroun. Nous avons également évalué l'existence d'une éventuelle variation saisonnière des formes M et S. Enfin, nous avons entrepris d'estimer la contribution de chacune des formes moléculaires M et S à la transmission du paludisme.

I.2- Résultats

1.2.1- Distribution géographique des formes moléculaires M et S au Cameroun

Grâce aux techniques de PCR mises au point par Favia *et al.* (1997, 2001), nous avons entrepris une étude de la distribution géographique des formes moléculaires M et S d'*An. gambiae* au Cameroun. Nous avons aussi recherché quels étaient les facteurs qui influençaient cette distribution et enfin nous avons comparé les deux techniques de PCR utilisées dans le but d'en retenir une seule pour les prochaines études.

Notre étude s'est effectuée dans le but d'avoir une carte de la distribution des deux formes moléculaires M et S dans la plus grande partie du Cameroun. Les échantillons ont été collectés de 1997 à 2002 lors de multiples enquêtes entomologiques sur le terrain. Les échantillons testés proviennent des différentes régions du pays (19 localités) où *An. gambiae* est présent. Les résultats

obtenus sont présentés dans le tableau 3. Ces résultats indiquent que la forme M est dominante dans 7 localités alors que la forme S l'est dans 12 localités. Les deux formes ont été retrouvées en sympatrie dans 14 des 19 localités. Toutefois, l'une des deux formes a toujours prédominé sur l'autre lorsqu'elles étaient trouvées ensemble dans une localité. Nous n'avons pas observé de localité où elles étaient présentes à des proportions égales. Les proportions de la forme prédominante dans chaque localité se situent autour de 90% et celles de la minoritaire à 10%. Ceci pourrait laisser entrevoir un phénomène d'exclusion entre les formes moléculaires M et S. Sur un total de 1521 spécimens identifiés, nous n'avons pas trouvé de spécimens hybrides M/S. Ce constat indique qu'il existe une barrière reproductive entre les formes moléculaires M et S. Ceci est mis en évidence dans les localités de sympatrie où ces deux formes coexistent. C'est le cas de la localité de Simbock où nous avons identifié 720 spécimens avec près de 92% qui sont de forme M. Dans cette localité, si les populations des deux formes moléculaires étaient en équilibre de Hardy-Weinberg, c'est à dire si les croisements étaient libres (panmixie), on se serait attendu à observer au moins 108 spécimens hybrides. En effet, si p est la fréquence de l'allèle M et q celle de l'allèle S, nous nous attendions à avoir sous l'hypothèse de Hardy-Weinberg, une proportion d'individus M correspondant à p^2 , une proportion d'individus S correspondant à q^2 et celle d'hybrides M/S correspondant à $2pq$. Dans ces conditions, nous aurions dû avoir 609 spécimens de forme M, 5 spécimens de forme S et 108 spécimens hybrides M/S. La déviation de l'équilibre de Hardy-Weinberg est significative ($P < 0,01$) due à un fort déficit en hétérozygotes. Ceci confirme l'isolement reproductif entre les formes moléculaires S et M. Il faut signaler que dans cette localité de sympatrie, ces deux formes sont de la forme chromosomique Forest ce qui conforte l'opinion selon laquelle l'isolement reproductif au sein d'*An. gambiae* ne dépend pas des formes chromosomiques.

La distribution des formes moléculaires M et S à travers le Cameroun semble avoir un lien avec les conditions bioclimatiques. En effet, il apparaît de notre étude une tendance à retrouver l'une ou l'autre forme en fonction soit de la végétation, soit du climat, soit de la nature des gîtes larvaires disponibles. (Tableau 4 et Figure 14). La carte de distribution de ces deux formes d'*An. gambiae* indique que la forme M est prédominante dans les localités situées plus au sud où la végétation est plus de type forêt humide avec une pluviométrie plus importante. C'est le cas de Simbock, de Mbalmayo ou de Douala où la forme M est seule présente ou alors très dominante. Il faut aussi signaler que dans ces localités, la plupart des spécimens dont la lecture des chromosomes a été effectuée, ont présenté l'arrangement standard $2R^{+/+}; /2L^{+/+}$ sans ou avec très peu d'inversions. Par contre, nous avons observé que dans les localités de forêt plus dégradée ou de savane avec une pluviométrie de niveau modéré (1000-1300mm de précipitation par an), la forme S était presque toujours prédominante ou seule présente. Dans ces localités, les études de

cytogénétique ont révélé que les spécimens de forme chromosomique Forest n'avaient pas que l'arrangement standard sans inversions comme dans les localités où M prédomine. Ici, les inversions sur le bras droit du chromosome 2 sont assez fréquentes en particulier 2Rb. Néanmoins, ce profil de distribution en fonction des facteurs éco-climatiques n'est pas homogène car nous avons observé une prédominance de la forme S dans la localité de Nyabessan située en pleine zone forestière du Sud.

Les spécimens de trois localités de la province de l'Adamaoua (Banyo, Bankim et Tibati) ont également été étudiées. Dans ces localités, deux autres formes chromosomiques ont été retrouvées à savoir Mopti et Savanna. A Bankim et Banyo (la forme chromosomique Savanna étant seule présente dans cette dernière localité), nous n'avons observé que la forme moléculaire S alors qu'à Tibati (où les deux formes chromosomiques Mopti et Savanna coexistent), nous avons retrouvé les deux formes moléculaires M et S. Tous les spécimens de la forme chromosomique Mopti sont tous de la forme moléculaire M alors que les spécimens de la forme chromosomique Savanna appartiennent aux deux formes moléculaires M et S.

1.2.2- Comparaison des deux méthodes (PCR-RFLP et PCR directe) d'identification des formes moléculaires M et S.

Les deux méthodes ont été mises au point par les mêmes auteurs Favia *et al.* (1997 et 2001) et testées sur des spécimens de la forme chromosomique Savanna, Mopti et Bamako au Mali et au Burkina Faso. Nos objectifs étaient de voir si les résultats fournis par les deux méthodes étaient strictement identiques et quelle était celle susceptible d'être préconisée pour les prochaines études.

Pour cela, nous avons choisi 80 spécimens de la localité de Simbock de forme chromosomique Forest, pour lesquels nous avons identifié la forme moléculaire en utilisant les deux méthodes. Les résultats que nous avons obtenus indiquent une exacte correspondance entre ces deux méthodes d'identification impliquant la PCR. Nous avons obtenu 59 spécimens de forme moléculaire M et 21 de forme S. Ainsi, ces deux méthodes sont identiques pour ce qui est de leurs résultats. En examinant chacune d'elles, nous constatons que la méthode de PCR-RFLP se fait en deux étapes, une réservée à l'amplification par PCR de la séquence de 1300 pb qui, dans la seconde étape, est hydrolysée enzymatiquement au niveau des sites de restriction de l'endonucléase HhaI.

Tableau 3 : Distribution des formes moléculaires M et S d'*An. gambiae* au Cameroun du Nord au Sud

Localités	Coordonnées géographiques	Période de capture	Forme Chromosomique	Forme moléculaire		
				M	S	M/S
1- Banyo	6°45'N; 11°49'E	Août 2001	Savanna, Forest	0	61	0
2- Tibati	6°28'N; 12°37'E	Août 2001	Savanna, Mopti	36	63	0
3- Bankim	6°00'N; 11°40'E	Août 2001	Forest, Savanna	0	25	0
4- Dschang	5°25'N; 10°10'E	Mai 1999 à Août 2000	Forest & Savanna (1)	1	117	0
5- Santchou	5°17'N; 9°58'E	Juillet 2000	NF	1	55	0
6- Nkoteng	4°30'N; 12°03'E	Décembre 1998 à Février 2000	Forest & Savanna (2)	1	66	0
7- Ebehda	4°11'N; 11°17'E	Février et Avril 2001	Forest	1	20	0
8- Mbebe	4°10'N; 11°00'E	Février et Avril 2001	Forest	1	18	0
9- Obala	4°09'N; 11°33'E	Novembre 1998 à Mars 2000	Forest	1	51	0
10 -Buéa	4°09'N; 9°14'E	Novembre 1999	NF	7	2	0
11- Esse	4°05'N; 11°53'E	Mai 1997	NF	0	15	0
12- Douala	4°03'N; 9°44'E	Mai 2000	NF	31	1	0
13- Yaoundé	3°52'N; 11°31'E	Avril 1997	Forest (1)	19	3	0
OCEAC (Ydé)	Idem	Colonie	NF	26	0	0
14- Simbock	3°51'N; 11°30'E	Décembre 1998 à Mars 2002	Forest	661	59	0
15- Mfou	3°44'N; 11°38'E	Avril 1997 + Avril 2001	NF	6	9	0
16- Mbalmayo	3°31'N; 11°30'E	Juin 2000	NF	117	3	0
17- Olama	3°26'N; 11°17'E	Février à Décembre 2000	NF	9	6	0
18- Nyabessan	2°24'N; 10°24'E	Février 2001	NF	0	29	0
TOTAL				918	603	0

NF : Non fait

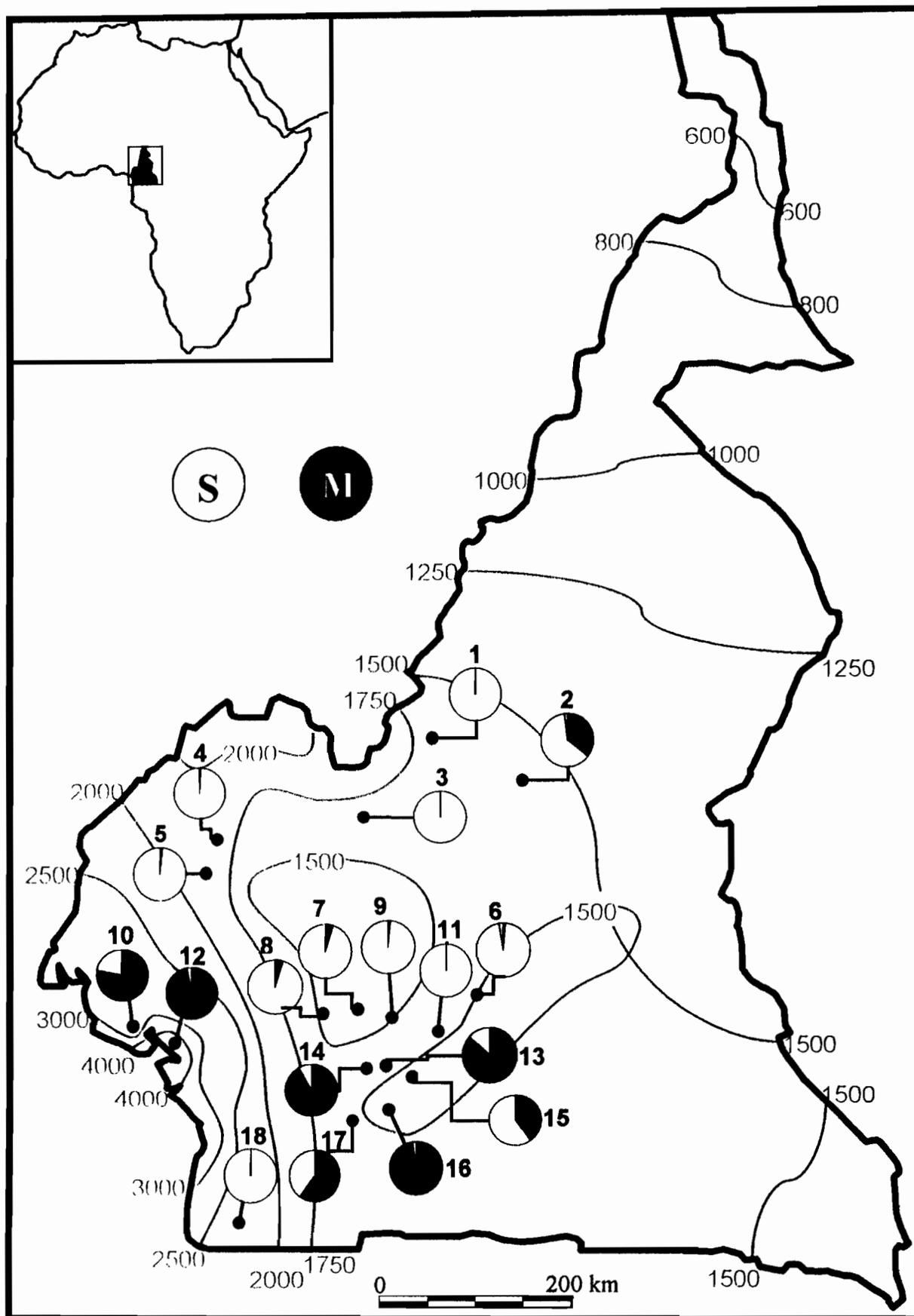


Figure 14 : Carte du Cameroun illustrant les différentes localités où la détermination des formes moléculaires M et S a été effectuée
 Les chiffres correspondent aux localités tel qu'indiqué au tableau 3. Les courbes d'isohyètes indiquent la pluviométrie des différentes régions.

Tableau 4 : Paramètres éco-climatiques des localités étudiées (en allant du Nord au Sud).

Localités	Type de végétation	Altitude en m	Humidité relative en %	Pluviométrie moyenne annuelle	T° Moyenne annuelle en °C	Type de climat
Banyo	Savane de l'Adamaoua	1110	65	1877	21	Climat tropical d'altitude de l'Adamaoua
Tibati	Savane de l'Adamaoua	874	65	1718	22	Climat tropical d'altitude de l'Adamaoua
Bankim	Savane de l'Adamaoua	700	65	1800	23	Climat tropical et équatorial de transition
Dschang	Forêt d'altitude	1400	75	1906	18	Climat tropical de montagnes de l'Ouest
Santchou	Forêt de la façade maritime	700	80	2200	25	Climat équatorial côtier Nord
Nkoteng	Savane périforestière	590	75	1550	26	Climat tropical et équatorial de transition
Ebebda	Forêt semi-décidue	550	75	1500	24	Climat tropical et équatorial de transition
Mbebe	Forêt semi-décidue	600	80	1550	24	Climat équatorial
Buéa	Forêt d'altitude	46	85	3200	23	Climat équatorial côtier Nord
Esse	Forêt semi-décidue	600	75	1500	24	Climat tropical et équatorial de transition
Douala	Forêt de la façade maritime	18	85	4108	26	Climat équatorial côtier Nord
Yaoundé	Forêt semi-décidue	760	75	1611	23	Climat équatorial
Simbock	Forêt semi-décidue	700	75	1611	23	Climat équatorial
Mfou	Forêt semi-décidue	600	80	1600	24	Climat équatorial
Mbalmayo	Forêt semi-décidue	641	80	1615	24	Climat équatorial
Olama	Forêt semi-décidue	628	80	1600	24	Climat équatorial
Obala	Forêt semi-décidue	610	75	1400	25	Climat tropical et équatorial de transition
Nyabessan	Forêt congolaise	407	80	1675	24	Climat équatorial

La méthode de PCR directe quant à elle se fait en une seule étape de PCR avec un cocktail d'amorces spécifiques de chacune des deux formes. Par conséquent, la PCR directe est plus rapide que celle impliquant la RFLP. En effet, la PCR-RFLP se fait durant environ une journée (PCR, migration et révélation sur gel d'agarose, hydrolyse enzymatique par HhaI et ensuite migration et révélation sur gel d'agarose) alors que la PCR directe se réalise sur une durée de 2h environ et permet ainsi le traitement de plusieurs spécimens en une journée. En plus, la méthode par PCR directe est plus économique car nécessitant moins de réactifs. Elle nécessite également moins d'expertise puisque l'étape de la RFLP est ôtée. Fort de tout ceci, nous préconisons désormais l'utilisation de la méthode de PCR directe pour les prochaines identifications de formes moléculaires chez *An. gambiae*.

I.2.3- Caractérisation chromosomique des populations d'*An. gambiae*

Pour une meilleure caractérisation d'*An. gambiae* au Cameroun, nous avons (en collaboration avec V. Petrarca et A. Della Torrè) déterminé les formes chromosomiques des spécimens rencontrés dans diverses localités et évalué la fréquence des différentes inversions chromosomiques (en fonction des conditions environnementales des localités). Très peu d'études ont été faites sur la caractérisation chromosomique d'*An. gambiae* en Afrique Centrale et donc au Cameroun où seul Robert *et al* (1993) et Fondjo (1996) ont entrepris ces déterminations mais seulement dans quelques localités notamment dans la partie septentrionale du pays. Les résultats présentés dans les tableaux 5 et 6 indiquent un polymorphisme des inversions chromosomiques chez *An. gambiae* en fonction des localités de notre étude. En effet, dans les localités situées plus au Sud du Cameroun dans la région de forêt semi-décidue (voir le tableau 4 pour plus de détails sur les paramètres géographiques de localités), nous observons peu ou pas d'inversions chromosomiques chez les différents spécimens traités. La forme chromosomique Forest est la seule rencontrée à l'exception de la localité de Nkoteng où un spécimen de la forme Savanna a été observé.

Ainsi les spécimens des localités de Simbock (8 spécimens) et de Mbébé (18 spécimens) ne présentent aucune inversion que ce soit pour le bras chromosomique 2R que pour le 2L (arrangement standard 2R+/+ ; 2L+/+).

Tableau 5 : Détermination des formes chromosomiques par localité

Localités	Nombre de spécimens	Forme chromosomique identifiée	Forme moléculaire identifiée
Mbébé	19	Forest	18S 1M
Ebebda	22	Forest	20S 1M
Obala	32	Forest	31S 1M
Simbock	8	Forest	7M 1S
Nkoteng	9	7 Forest 1 Savanna	9S Aucun M
Dschang	18	Forest	18S Aucun M
Tibati	30	22 Savanna 7 Mopti 1 Forest/Savanna	15 M 19 S
Bankim	28	17 Forest 1 Savanna 3 Forest/Savanna	24 S Aucun M
Banyo	49	33 Savanna 3 Forest 3 Mopti/Savanna 1 Forest/Savanna	49 S Aucun M

La lecture des chromosomes de certains spécimens n'a pas été possible à cause d'un mauvais conditionnement ou parce que leur stade n'était pas adéquat. De même, la forme moléculaire de certains spécimens n'a pas été déterminée à cause d'échec de PCR.

2Rb/b-2L+/+

Banyo	----(+/+)	7,5	-(+/+)	2,5	40	2R+/+-2La/+
	-1---(b/+)	10	1(a/+)	25		2Rb/+ -2La/a
	-211-(bcd/b)	20	2(a/a)	72,5		2Rb/+ -2La/a
	-2---(b/b)	50				2Rb/b-2La/+
	-21--(bc/b)	7,5				2Rb/b-2La/a
	-111-(bcd/+)	5				2Rbc/b-2L+/+
						2Rbc/b-2La/a
				2Rbcd/b-2La/a		
				2Rbcd/-2La/+		
				2Rbcd/+ -2La/a		

Sur le bras 2R, la configuration chromosomique est codée selon un protocole standard avec la séquence des chiffres correspondant aux inversions j, b, c, d et u toutes définies selon l'arrangement standard. Le signe '-' traduit l'état homozygote standard, '1' traduit l'hétérozygote et '2' signifie la présence de l'inversion à l'état homozygote. Ainsi, la configuration '-111-' correspond à un spécimen porteur des inversions b, c et d à l'état hétérozygote alors que la configuration '-211-' correspond à un spécimen homozygote pour l'inversion b et hétérozygote pour les inversions c et d. Les lettres f, s et n représentent de nouvelles inversions observées dans notre étude.

Sur le bras 2L, seule une inversion «a» est observée et les configurations ont été notées comme pour le bras 2R

Les spécimens de la localité d'Ebebda ne présentent également aucune inversion sur le bras 2R mais on y observe l'inversion «a» sur le bras 2L à une fréquence de 9%.

Les spécimens d'Obala, localité située à la lisière de la forêt semi-décidue et proche de la savane périforestière, présentent un polymorphisme autant sur le bras 2R que sur le bras 2L. L'inversion 2Rb est rencontrée à une fréquence de 8,82% et l'inversion 2La est observée à une fréquence de 8,9%. Cette inversion 2La y a même été rencontrée à l'état homozygote. Coluzzi *et al.* (1979) ont montré que la fréquence de l'inversion 2La augmente avec l'aridité de l'environnement.

La localité de Dschang située dans la région des hautes montagnes de l'Ouest Cameroun, caractérisée par la forêt d'altitude, présente l'inversion 2Rb à une fréquence de 25% et l'arrangement standard pour le bras 2L à 100%. Néanmoins nous devons signaler la faiblesse de la taille de notre effectif (seulement 4 spécimens). En revanche, la fréquence élevée de l'inversion 2Rb dans cette localité est confortée par les études antérieures (Robert *et al.*, 1993) qui ont même signalé la présence de la forme chromosomique Savanna dans cette région.

Sur les 8 spécimens de la localité de Nkoteng située en zone de transition forêt-savane (savane périforestière), un appartient à la forme chromosomique Savanna (2Rbcd/+). L'inversion 2Rb y est observée à une fréquence de 12,5% alors que l'arrangement standard 2R+/+ est observé à une fréquence de 75%. Le bras chromosomique 2L présente l'inversion

2La à une fréquence de 18,75 %, ce qui est dans notre étude, la fréquence la plus élevée pour cette inversion dans les populations de la région du grand Sud du Cameroun. Cette observation est en accord avec les caractéristiques du milieu écologique de cette localité de savane péri-forestière qui est plus sèche et chaude que d'autres localités du Sud.

La localité de Bankim est située à la lisière de la savane de l'Adamaoua dont elle est séparée par la falaise du plateau de l'Adamaoua. La forme chromosomique Forest y est dominante. En effet, sur 22 spécimens caryotypés, 18 sont de la forme Forest, 1 de la forme Savanna et 3 sont des hybrides Forest/Savanna. Ces hybrides possèdent les inversions 2Rb ou 2La à l'état homozygote ou les deux à l'état hétérozygote. L'inversion 2Rb y est observée à une fréquence de 17% et l'inversion 2La à 19%. Une nouvelle inversion que nous appellerons 'n' a été observée chez un spécimen de cette localité. Nous ne savons pas à quoi cette inversion est associée mais comme les autres nous pensons qu'elle contribue à l'adaptation du moustique aux conditions de son écosystème. Les caryotypes observés dans cette localité confirment bien la position intermédiaire de cette population d'*An. gambiae* entre les populations de forêt et celles de savane.

La population de la localité de Banyo est largement dominée par les spécimens de forme chromosomique Savanna (33 spécimens). Nous y avons également observé 3 spécimens de forme Forest, 3 hybrides Mopti/ Savanna et un hybride Forest/Savanna. Ceci résulte d'un fort polymorphisme des inversions sur le bras 2R où des inversions comme 'c', 'd' apparaissent alors qu'elles étaient absentes dans les populations où la forme Forest étaient dominante. Les fréquences des différentes inversions y sont élevées à l'exemple de 2Rb qui a une fréquence de 85%. Les inversions 'c' et 'd' y sont présentes aux fréquences respectives de 16,25% et 12,5% respectivement. L'arrangement standard 2R+/+ n'y est observée qu'à une faible fréquence (7,5%). Le bras 2L présente quant à lui l'inversion 2La à une fréquence de 85%. Ce polymorphisme correspond bien à celui observé généralement dans d'autres études sur des populations d'*An. gambiae* des régions de savane soudanaise africaines (Touré *et al.*, 1998). La présence de la forme Forest dans cette localité montre bien que nous avons encore à faire à une population de savane humide.

La localité de Tibati offre un profil de résultats différent de celui des autres localités étudiées jusqu'à présent. En effet, bien que la forme Savanna y soit dominante (23 spécimens sur les 32), nous y observons 7 spécimens de forme Mopti caractérisés par l'arrangement 2Rbc ou 2Rbc/u. La présence de cette forme chromosomique est associée par Coluzzi *et al.* (1979) à l'existence d'un milieu irrigué comme la riziculture ou des fleuves, ce qui est quelque peu le cas à Tibati où l'on trouve un grand lac d'environ 70km². En outre, sur les 22

spécimens classés comme étant de la forme Savanna, 11 présentent l'arrangement standard 2R+/+ ; 2L+/+ et peuvent aussi être considérés comme étant de la forme Forest. Tous ceci fait de la localité de Tibati un lieu de sympatrie de plusieurs formes chromosomiques. Une nouvelle inversion que nous appellerons 'f' y a été observée chez 4 spécimens. Comme à Banyo, l'inversion 'd' y est observée et on y trouve particulièrement l'arrangement 2Rcu/+ qui est très typique des populations de savane notamment au Mali (Touré *et al.*, 1998). Les fréquences des inversions rencontrées sont de 22% pour 2Rb, 15% pour 2Rbc et 42% pour l'inversion 2La.

I.2.3.1- Test de Hardy-Weinberg pour les populations de l'Adamaoua

Les populations du Sud Cameroun que nous avons étudiées étaient de faible taille pour quelques-unes ou alors étaient composées de spécimens présentant l'arrangement standard 2R+/+ ; 2L+/+ pour la majorité. De ce fait, nous n'y avons pas effectué le test de Hardy-Weinberg pour voir si elles constituaient des unités panmictiques car seule la forme chromosomique Forest y est présente. Pour cette raison, nous ne nous sommes intéressés qu'aux trois populations de la province de l'Adamaoua où un plus grand polymorphisme des inversions a été observé, ainsi que de nouvelles formes chromosomiques et même des hybrides. Les tests de Hardy-Weinberg ont été réalisés grâce au logiciel Genepop (Raymond & Rousset, 1995), en considérant une inversion chromosomique comme étant un locus.

i) Bankim

Les deux inversions observées ont été 2Rb et 2La. Ces deux inversions sont toutes deux en équilibre de Hardy-Weinberg dans cette population. En effet, l'hypothèse d'un équilibre de Hardy-Weinberg au niveau de l'inversion 2La est retenue avec une probabilité de $P = 0,12$, tout comme au niveau de l'inversion 2Rb avec une valeur de $P = 0,2$. Ces résultats indiquent que la population de Bankim est génétiquement homogène. Le fait que seule la forme moléculaire S y ait été trouvée conforte cette homogénéité observée.

ii) Banyo

Les tests de Hardy-Weinberg ont été réalisés avec les inversions 2Rb, 2Rd et 2La. Nous observons un équilibre pour ce qui est des inversions 2Rd et 2La. Par contre, le système 2Rb présente une déviation par rapport à l'équilibre, notamment lorsque 2Rb/+ est comparé à 2Rb/b et 2Rbc. La valeur du χ^2 est de 6,8 avec un nombre de degré de liberté df égal à 1 et la valeur de $P=0,009$. Néanmoins, cette population est génétiquement homogène dans son ensemble. Le fait que seule la forme moléculaire S y soit présente conforte cet avis.

iii) Tibati

Les inversions considérées pour le test sont 2La et les inversions du système 2Rb. Pour l'inversion 2La, nous observons une déviation par rapport à l'équilibre de Hardy-Weinberg à cause d'un déficit en hétérozygotes ($\chi^2=6,8$; $P = 0,009$). Pour le système 2Rb en mettant de côté l'arrangement 2Rcu/+, on observe un déficit significatif en hétérozygotes (2Rb/b contre 2Rbc/+ et 2Rb/+). La valeur du χ^2 est de 5,9 et celle de $P = 0,02$. Ces déficits en hétérozygotes sont un indicateur de ce que la population d'*An. gambiae* de Tibati est composée de deux entités bien différentes qui y coexistent. Ceci est conforté par la présence dans cette localité des deux formes moléculaires M et S.

I.2.4- Variations saisonnières des formes moléculaires M et S dans la localité de sympatrie de Simbock

La distribution des formes moléculaires M et S telle qu'observée dans notre étude, présente bien des particularités. Dans chaque localité étudiée, l'une des deux formes est toujours largement prédominante (environ 90% pour la forme prédominante et 10% pour l'autre forme). Dans plusieurs régions, une seule forme est présente. Une pareille observation amène à se poser la question de savoir quels sont les facteurs qui régissent la distribution des formes moléculaires M et S. Nous pensons que l'exclusion compétitive et les facteurs éco-climatiques sont les principaux éléments influençant cette répartition. Le profil de distribution des deux formes moléculaires montre que la forme M est prédominante dans les zones plus humides avec une végétation plus dense alors que la forme S se retrouve dans les zones plus sèches avec une végétation plus dégradée. Si cette hypothèse est correcte, on peut s'attendre à une variation au niveau de la proportion de l'une ou l'autre des formes moléculaires M et S dans une zone de sympatrie en fonction de la saison et de la variation annuelle de paramètres éco-climatiques. Ainsi en saison pluvieuse, la fréquence de M pourrait être plus élevée.

Nous avons conduit une étude sur deux ans dans la localité de Simbock où les formes M et S sont sympatriques avec une prédominance de la forme M afin d'évaluer une éventuelle variation saisonnière entre M et S en fonction des paramètres éco-climatiques.

Tableau 7 : Nombre d'individus des formes M et S identifiés mensuellement à Simbock d'Octobre 1998 à Mars 2001

	1998		1999										2000						2001				
	O	D	J	F	M	A	M	J	J	A	S	O	N	J	M	A	M	J	S	N	J	F	M
M	15	30	20	17	16	24	31	8	6	10	7	24	57	13	23	2	44	19	13	11	10	12	15
S	0	8	1	0	0	1	1	3	1	0	2	7	6	5	0	2	2	0	0	2	0	1	3

Les lettres sont les premières des noms des mois d'une année

Nous avons procédé à l'identification moléculaire des spécimens d'*An. gambiae* durant 23 mois d'Octobre 1998 à Mars 2001. Le tableau 7 présente les résultats obtenus. Les individus de forme M ont été présents pendant toute la période d'étude alors que ceux de forme S ont été absents au cours de 8 mois. Le nombre moyen de spécimens identifiés par mois est de 21 (4 et 63 spécimens étant le minimum et le maximum respectivement). Pour chaque mois, la proportion de M a toujours été plus élevée que celle de S. Une seule fois nous avons obtenu une proportion de 50% pour chacune des deux formes, pendant le mois d'avril 2000, mais il faut noter que l'effectif était très faible (4 spécimens seulement). Les figures 15 et 16 présentent les proportions des différentes formes moléculaires mensuellement et leur nombre en fonction de la pluviométrie. Nous avons évalué la présence éventuelle d'une variation saisonnière en considérant de façon générale que dans la localité de Simbock, nous avons dans un premier temps deux saisons climatiques dont une sèche et une pluvieuse. Dans un second cas, nous avons considéré 4 saisons climatiques dont 2 pluvieuses et 2 sèches. La grande saison pluvieuse (en terme de quantité de précipitation) allant de septembre à novembre et la petite d'avril à juin, tandis que la grande saison sèche va de décembre à mars et la petite de juillet à août. Nous avons évalué une éventuelle corrélation entre les différentes saisons et les proportions de chacune des deux formes moléculaires.

I.2.4.1- Hypothèse de deux saisons climatiques à Simbock

Le tableau 8 indique les proportions des deux formes M et S en fonction des deux saisons climatiques.

Tableau 8 : Distribution des deux formes M et S en fonction des saisons

	Saison pluvieuse	Saison sèche	Total
M	236	191	427
S	26	19	45
Total	262	210	472

Nous avons effectué le test de chi-deux et estimé la probabilité P pour une association entre la pluviométrie et les proportions des deux formes. Nous avons obtenu pour un nombre de degré de liberté égale à 1, une valeur de χ^2 égale à 0,11 et une probabilité P égale à 0,74. Nous n'avons pas mis en évidence de lien entre les périodes hydrométriques et les proportions des deux formes moléculaires M et S.

I.2.4.2- Hypothèse de 4 saisons climatiques à Simbock

Nous avons ensuite évalué l'influence de la pluviométrie sur les proportions de M et S en considérant la présence de quatre saisons à Simbock. Le tableau 9 indique les proportions de M et S selon les quatre saisons.

Tableau 9 : Distribution des formes M et S en fonction des 4 saisons climatiques à Simbock

	Grande Saison pluvieuse	Petite saison pluvieuse	Grande saison sèche	Petite saison sèche	Total
M	127	109	156	35	427
S	17	9	18	1	45
Total	144	118	174	36	472

Le test de chi-deux réalisé donne une valeur de χ^2 égale à 3,4 avec le nombre de degré de liberté égale à 3 et la probabilité associée P égale à 0,33. Une fois encore, nous constatons que les proportions de M et S ne sont pas significativement influencées par les saisons climatiques. En effet, la valeur de P obtenue dans cette seconde hypothèse confirme bien l'indépendance des deux paramètres à savoir la proportion des deux formes moléculaires et la pluviométrie.

Du moment où il apparaît que la pluviométrie n'influence pas significativement les proportions relatives des formes M et S même dans une localité de sympatrie, il est clair que d'autres facteurs sont impliqués. Ce pourrait notamment être le phénomène d'exclusion compétitive entre ces deux formes ou celui d'isolement écologique encore appelé «niche partitionning».

I.2.5- Implications des formes moléculaires M et S d'*An. gambiae* dans la transmission du paludisme

L'hétérogénéité génétique observée entre les formes moléculaires M et S d'*An. gambiae* caractérisée par la présence d'un isolement reproductif entre elles, en fait 2 entités indépendantes dont le rôle épidémiologique dans la transmission du paludisme doit être examiné. Nous avons entrepris d'évaluer ce rôle épidémiologique dans 4 sites, chacune des deux formes étant majoritaire ou seule présente dans 2 localités. Les deux localités où la forme M est prédominante sont Simbock et Mbalmayo alors que les deux localités de la forme S sont Nkoteng et Dschang. Nous avons retenu comme indices entomologiques pouvant nous

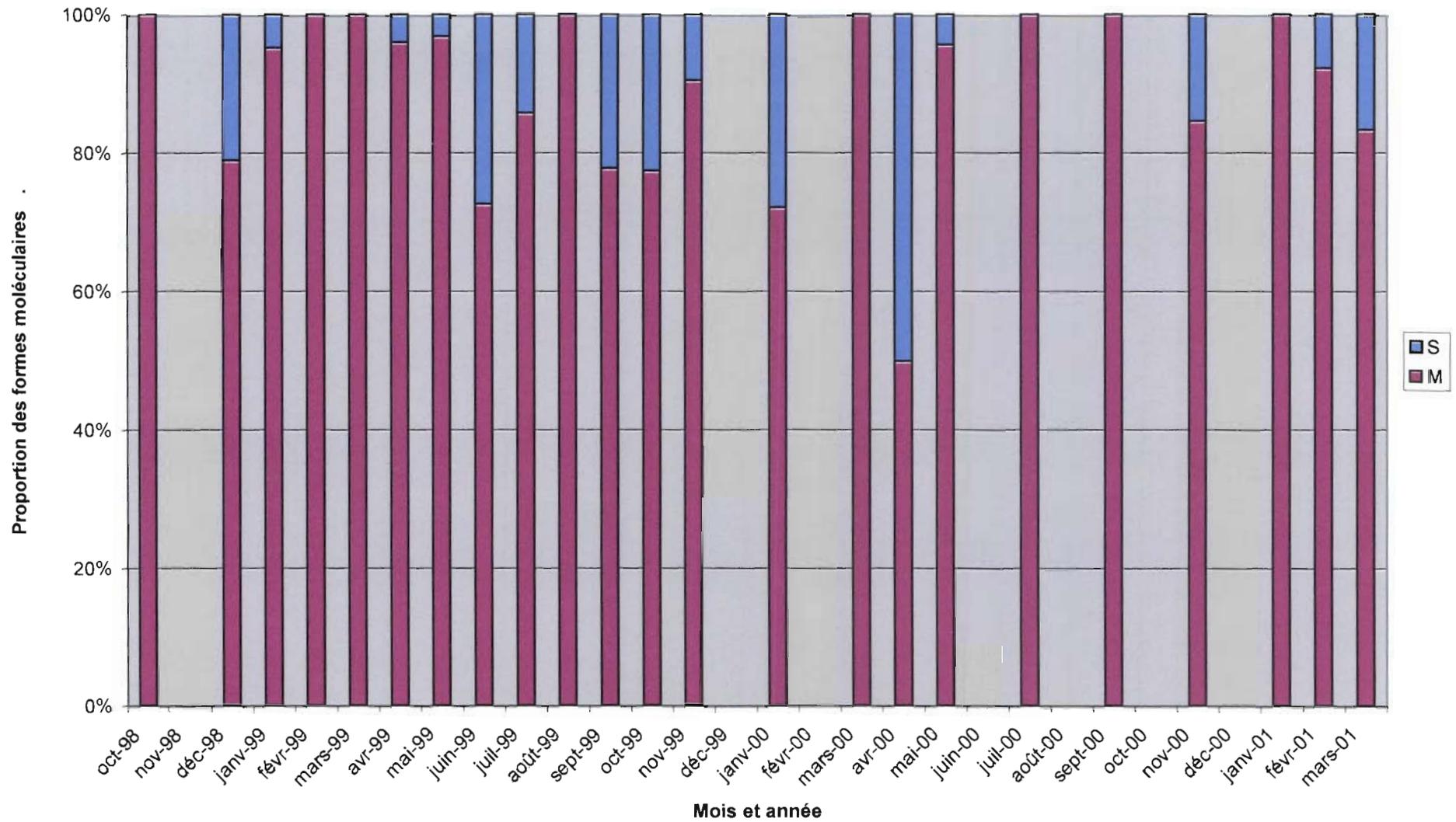


Figure 15 : Fréquences mensuelles des formes moléculaires M et S d'octobre 1998 à mars 2001 à Simbock

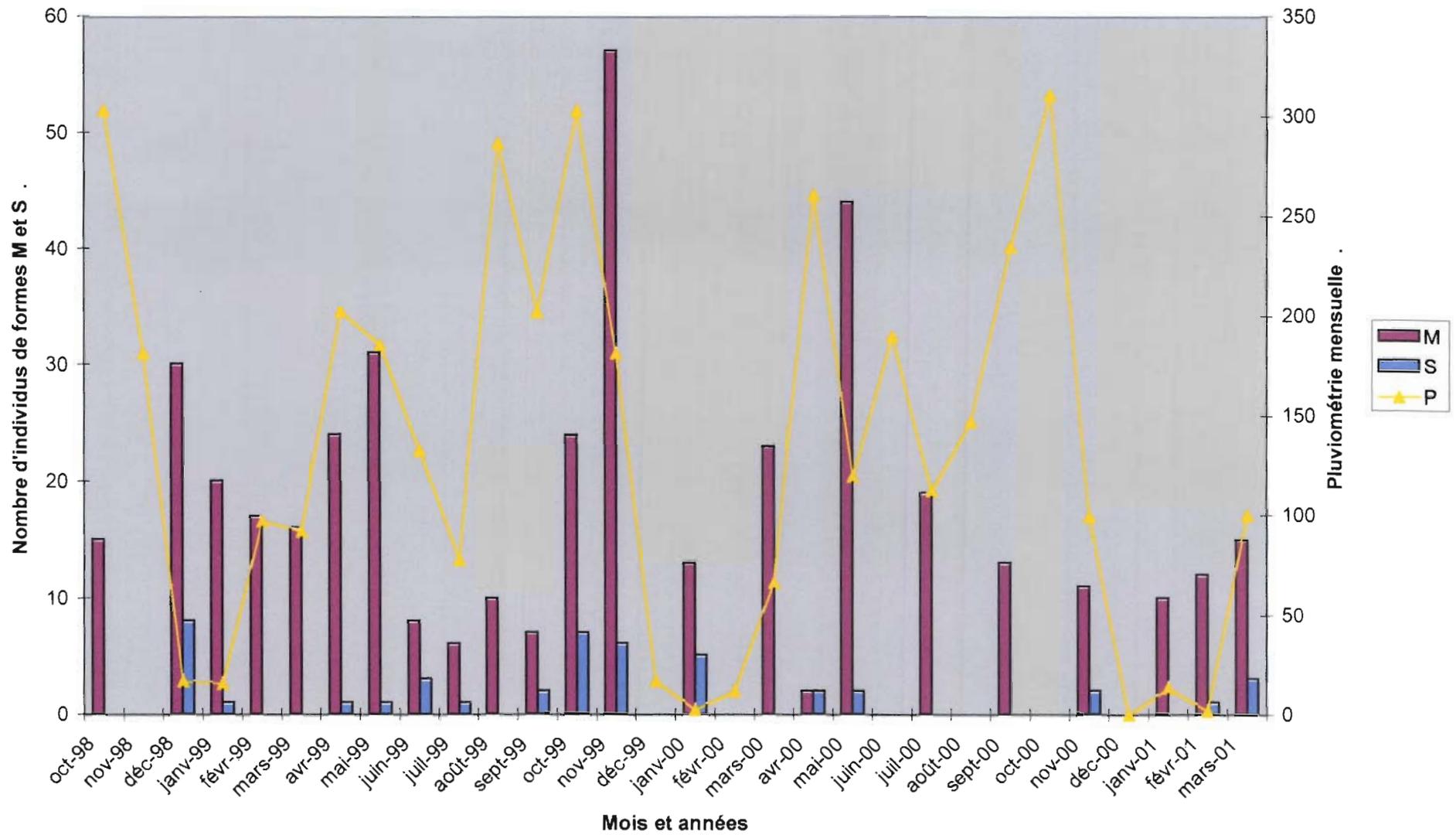


Figure 16 : Variation mensuelle des formes M et S à Simbock en fonction de la pluviométrie d'octobre 1998 à mars 2001

aider à évaluer ces rôles épidémiologiques, l'indice circumsporozoïtique, les préférences trophiques et le taux d'inoculation entomologique (TIE).

I.2.5.1- Indice circumsporozoïtique

Dans la localité de Simbock, nous avons évalué cet indice pour les deux formes bien que le nombre de spécimens de forme M soit nettement plus élevé. Après ELISA CSP, nous avons obtenu 26 spécimens infectés sur 502 pour la forme M (soit 5,2%). Le parasite responsable de l'infection ici étant dans 24 cas uniquement *P. falciparum*. Deux infections mixtes à *P. falciparum* et *P. malariae* ont été observées.

Pour la forme S, nous avons observé 3 infections toutes à *P. falciparum* sur 44 spécimens testés soit un indice circumsporozoïtique égale à 6,8%.

Dans la localité de Mbalmayo où la forme M prédomine à 97%, nous avons observé sur 424 spécimens d'*An. gambiae* capturés sur volontaires durant la nuit, 32 positifs à *P. falciparum* soit un indice circumsporozoïtique de 6,1%.

Sur 81 spécimens d'*An. gambiae* recensés à Nkoteng où la forme S est très majoritaire à 98%, 6 ont été positifs à *P. falciparum* soit un indice CS égale à 7,4%.

Dans la localité de Dschang où la forme S est très majoritaire à 99%, nous avons enregistré 18 spécimens positifs à l'ELISA CSP dont 17 à *P. falciparum* et 1 à *P. malariae* sur un total de 1069 spécimens d'*An. gambiae* capturés sur homme. L'indice circumsporozoïtique est de 1,7%, le plus bas enregistré dans notre étude.

I.2.5.2- Taux d'inoculation entomologique

Nous avons estimé ce paramètre au niveau des populations des deux formes pour évaluer leur degré d'implication dans la transmission du paludisme car il nous permet de déterminer le nombre de piqûres infectantes dans une échelle de temps donné.

A Simbock, sur les deux années de l'étude longitudinale, nous avons enregistré pour *An. gambiae* composé à grande majorité de forme M, un taux d'inoculation annuelle de 76,75 piqûres infectantes par homme et par an. Pour la localité de Mbalmayo, ce taux est de 94,5 piqûres infectantes par homme et par an.

Dans les localités où la forme S est prédominante, nous observons un taux de 11,1 piqûres infectantes par homme par an à Dschang et de 39,9 piqûres infectantes par homme par an à Nkoteng.

I.2.5.3- Préférence trophique

Nous avons cherché à voir quelles étaient les sources préférées de repas de sang pour chacune des deux formes moléculaires.

Dans la localité de Simbock, sur 19 spécimens d'*An. gambiae* de la forme M testés pour déterminer l'origine de leur repas de sang, 18 ont pris leur repas sur homme et 1 repas mixte pris sur homme et bœuf a été noté. Pour 6 repas de spécimens de forme S testés, tous ont été pris sur homme. A Mbalmayo, sur 10 repas de sang d'*An. gambiae* M testés, nous avons observé 10 repas pris sur homme.

Quant aux populations de forme S, sur 24 spécimens testés à Nkoteng, nous avons obtenu un pourcentage de 100% de repas pris sur homme. A Dschang, sur 25 repas de sang d'*An. gambiae* testés par ELISA, nous avons eu 24 d'origine humaine et 1 pris sur mouton.

L'ensemble des résultats est regroupé dans le tableau 10. Le caractère anthropophile des deux formes moléculaires M et S d'*An. gambiae* se dégage donc très fortement de cette étude. En effet, pour les spécimens de forme M un seul cas de repas mixte a été mis en évidence, le reste étant pris sur homme alors que pour les spécimens de forme S, un seul repas pris sur un autre hôte que l'homme (le mouton), a été signalé.

Tableau 10 : Récapitulatifs des valeurs des paramètres entomologiques des formes moléculaires M et S d'*An. gambiae* dans 4 localités du Cameroun.

Forme moléculaire	Localités	Agressivité (p/h/n)	Anthropophilie	IS	TIE annuel (pi/h/an)
M	Simbock M	3,9	18/19 et 1 mixte	5,2	74,75
	Mbalmayo	4,24	10/10	6,1	94,5
S	Nkoteng	0,41	24/24	7,4	11,1
	Dschang	6,43	24/25 et 1 mouton	1,7	39,9
	Simbock S	/	6/6	6,8	/

P/h/n : piqûres par homme et par nuit ; IS : Indice circumsporozoïtique ; Pi/h/n : Piqûres infectantes par homme et par an ; TIE : Taux d'inoculation entomologique

I.3- Discussion

L'étude que nous avons conduite, nous a permis d'établir une carte de la répartition géographique des formes M et S ainsi que celle des inversions chromosomiques chez *An. gambiae* au Cameroun. De cette étude, de nombreuses observations, hypothèses et conclusions peuvent être faites sur l'isolement reproductif entre les formes moléculaires M et S ainsi que leur impact dans la transmission du paludisme.

I.3.1- Facteurs influençant la distribution des formes moléculaires M et S

La répartition des formes moléculaires M et S au Cameroun a présenté des traits particuliers indiquant qu'elle serait sous l'influence de plusieurs facteurs. Dans toutes les localités, une des deux formes moléculaires a toujours été prédominante, laissant penser à l'existence d'un phénomène d'exclusion compétitive entre les formes moléculaires M et S. Nous avons aussi noté une certaine corrélation entre le faciès écologique des localités et le type de forme moléculaire qui y est rencontré, suggérant une influence des facteurs éco-climatiques sur la distribution des formes moléculaires M et S.

I.3.1.1- Exclusion compétitive

Ce phénomène est mis en exergue du fait de la nette prédominance d'une forme sur l'autre dans toutes les localités étudiées et par la présence dans certaines localités d'une seule forme moléculaire. Par ce phénomène d'exclusion, nous pensons que la présence d'une forme moléculaire d'*An. gambiae* dans une localité serait un facteur limitant pour l'établissement de l'autre forme. Ceci pourrait expliquer pourquoi dans deux localités proches comme Yaoundé et Obala, nous avons dans l'une la forme M prédominante et dans l'autre plutôt la forme S. D'autres auteurs travaillant sur cette thématique ont émis la même hypothèse (Della Torrè *et al.*, 2001 ; Touré *et al.*, 1998 ; Coluzzi *et al.*, 1985). Il a ainsi été mis en évidence au Mali, que lorsque la forme chromosomique Mopti colonisait une zone (delta du fleuve Niger et zones irriguées), elle excluait l'établissement de la forme chromosomique Savanna et même celui d'*An. arabiensis*. Ceci est prouvé par le fait que dans des zones de même faciès écologique où la forme Mopti était absente, ces autres formes et espèces sont rencontrées. Le mécanisme par lequel les populations de ces deux formes moléculaires arrivent à s'exclure l'une de l'autre n'a pas fait l'objet de notre étude. Mais si cette hypothèse est vérifiée, il serait intéressant de chercher à comprendre comment ce mécanisme d'exclusion fonctionne, bien que certains modes d'exclusion rencontrés entre différentes espèces de moustiques comme entre *Aedes*

aegypti et *Aedes albopictus* (Black *et al.*, 1989) suggèrent que c'est probablement au niveau larvaire que l'exclusion s'effectue.

I.3.1.2- Facteurs éco-climatiques

La distribution des deux formes M et S au Cameroun et particulièrement au grand Sud indique que les facteurs éco-climatiques y joueraient aussi un rôle. En effet, nous avons observé que la forme M était plus prédominante dans des localités où le couvert végétal était moins dégradé, contrairement à la forme S qui est trouvée en majorité dans les localités avec une végétation plus dégradée. Autour de Yaoundé, il semble qu'il existerait une ligne de rupture entre les deux formes moléculaires car au nord de cette ligne, la forme S est très prédominante notamment à Obala, Nkoteng, Dschang et Bankim. Par contre, au Sud de Yaoundé, nous constatons que la forme M est plus prédominante (Simbock, Mbalmayo, Mfou, Douala, Buéa, Olama). La région située au nord de Yaoundé est caractérisée par une pluviométrie plus faible, une végétation de type savane périforestière alors que la région Sud est caractérisée par une végétation de type forêt congolaise, semi-décidue ou encore de façade maritime. La pluviométrie y est plus forte. Cette corrélation observée entre les facteurs éco-climatiques et la distribution des deux formes moléculaires M et S est néanmoins troublée par le résultat de la localité de Nyabessan située en plein Sud où contrairement au profil décrit ci-dessus, c'est la forme S qui y est présente. Ce fait pourrait signifier que la distribution de ces deux formes n'est pas influencée par un facteur mais plutôt par un ensemble de facteurs dans lesquels nous pouvons aussi inclure le phénomène d'isolement écologique ou «niche-partitionning».

I.3.1.3- Le phénomène d'isolement écologique ou «niche-partitionning»

Ce phénomène suppose que les deux formes moléculaires M et S coloniseraient des gîtes différents pour leur développement larvaire. En effet, il a été montré dans les études précédentes notamment au Mali (Touré *et al.*, 1998), que la forme S (forme chromosomique Savanna) avait pour gîtes des collections d'eau temporaires issues des pluies alors que la forme M (Cytotype Mopti), avait pour gîtes des collections d'eau permanentes telles que les bordures des fleuves et des rivières et les zones irriguées. Nous pensons que, bien que la plupart des spécimens des localités étudiées soit de la forme chromosomique Forest, la disponibilité des gîtes larvaires pourrait aussi bien justifier la distribution des formes telle qu'observée dans l'étude. Pour clarifier cette hypothèse, il serait intéressant d'étudier l'écologie larvaire de chacune des formes dans les différentes localités. Nous avons

néanmoins dans la localité de Simbock trouvé dans un même gîte, une larve de forme S parmi d'autres larves de forme M sans que ceci puisse éliminer l'hypothèse de l'isolement écologique.

I.3.2- Relation formes chromosomiques/formes moléculaires au Cameroun

La relation existant entre les formes chromosomiques et les deux formes moléculaires a fait l'objet de quelques études antérieures (Della Torrè *et al.*, 2001 ; Touré *et al.*, 1998). La découverte des formes moléculaires a d'ailleurs eu lieu lors de la tentative de Favia *et al.* (1997) de trouver des marqueurs moléculaires diagnostics pouvant permettre d'identifier aisément par PCR toutes les formes chromosomiques. Cette étude conduite avec des spécimens provenant du Burkina Faso et du Mali avait montré que les formes chromosomiques Savanna et Bamako correspondaient à la forme moléculaire S tandis que la forme chromosomique Mopti correspondait à la forme moléculaire M. Mais depuis lors, on s'est rendu compte que cette correspondance ne tenait qu'au sein de ces deux pays et que hors de ceux-ci, la correspondance n'était plus automatique.

De l'étude que nous avons menée, il ressort que les populations de la forme moléculaire M au Cameroun (Simbock, Tibati) ont présenté une variété d'arrangements chromosomiques correspondant à ceux des formes chromosomiques Forest, Savanna et Mopti. En effet, dans les localités du Sud où les spécimens de forme M ont été trouvés, ils ont été tous de la forme chromosomique Forest avec l'arrangement standard $2R^{+/-}2L^{+/-}$ sans aucune inversion. La population M de la localité de Tibati quant à elle, a présenté un arrangement chromosomique correspondant à la forme Mopti (arrangement $2Rbc/+$ ou $2Rbc/u$) ou à la forme chromosomique Savanna ($2Rb/+$, $2Rbcd/d$, $2Rb/b$ et autres). On y a trouvé un spécimen M qui était un hybride Forest/Savanna. Ce résultat confirme le fait que hors du Mali et du Burkina, la forme moléculaire M ne correspond pas uniquement à la forme chromosomique Mopti.

Les populations de la forme moléculaire S ont présenté des arrangements chromosomiques correspondant aux formes chromosomiques Forest et Savanna. Dans le Sud, à l'exception d'un spécimen qui s'est avéré être de la forme Savanna, tous les individus de la forme S sont tous de la forme chromosomique Forest avec la présence unique dans la localité de Mbébé, de l'arrangement standard ($2R^{+/-} ; 2L^{+/-}$). Mais par contre, dans tous les autres sites, l'on note la présence d'inversions en proportion de plus en plus importante lorsqu'on approche des zones de savane. Ce sont notamment les inversions $2Rb$ et $2La$. Pour l'inversion $2La$, nous avons pour les localités de Mbébé, Ebebda, Obala, Bankim et Nkoteng

respectivement les fréquences de zéro ; 9 ; 9 ; 19 et 37,5%. Pour l'inversion 2Rb, pour les localités de Mbébé, Ebebda, Obala, Nkoteng et Bankim, nous avons respectivement les fréquences de 0 ; 0 ; 8,8 ; 12,5 et 28 %. Les populations S des zones de savane notamment à Banyo et à Tibati, sont presque toutes de la forme chromosomique Savanna. A Banyo, 3 spécimens sont de la forme Forest avec l'arrangement standard 2R+/+ mais possèdent à chaque fois l'inversion 2La à l'état hétérozygote. L'on y trouve aussi 4 spécimens hybrides dont 3 Savanna/Mopti et 1 Savanna/Forest. A Tibati, tous les spécimens S sont de la forme chromosomique Savanna.

La relation entre les formes chromosomiques d'*An. gambiae* et les formes moléculaires M et S au Cameroun est schématiquement indiquée dans la figure 17.

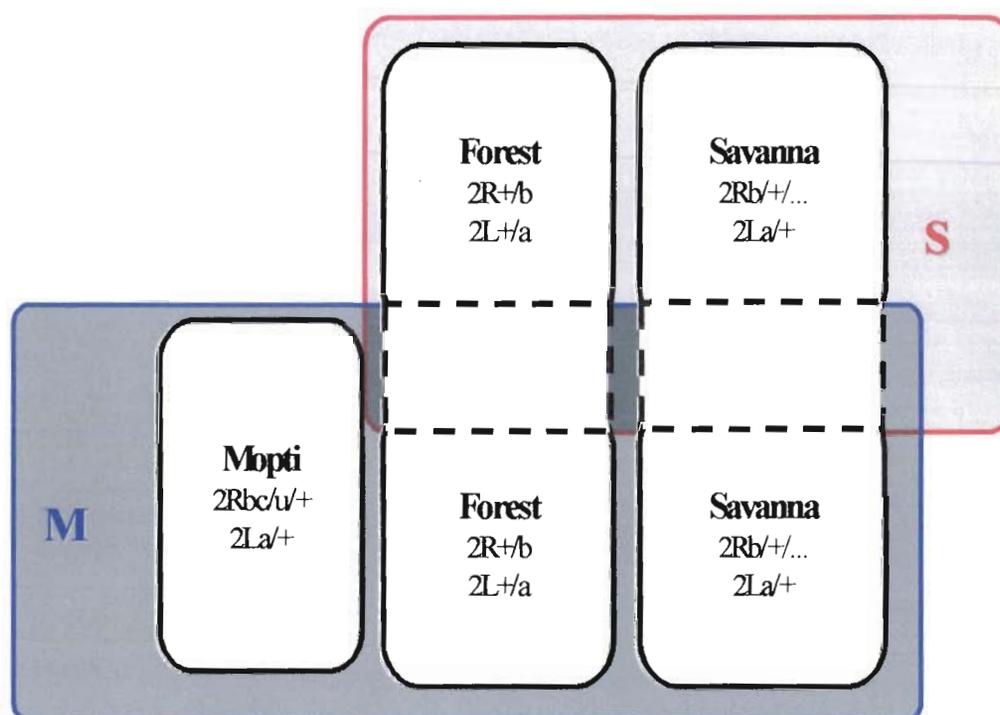


Figure 17 : Relation entre formes moléculaires et formes chromosomiques au Cameroun

M : Forme moléculaire M, **S** : Forme moléculaire S, **2Rb/+** ; **2L+/a**... représentent les différentes inversions spécifiques de chaque forme chromosomique.

De manière plus précise, nous pouvons caractériser sur le plan cytogénétique les formes moléculaires M et S au Cameroun de la façon suivante :

I.3.2.1- La forme moléculaire S

Elle est caractérisée dans les localités de forêt du Sud, par une fréquence élevée de l'arrangement standard sur le bras 2R (ce qui est typique de la forme chromosomique Forest). Par ailleurs, on note une augmentation significative de la fréquence des inversions 2Rb et 2La

à mesure qu'on approche de la région de savane. Cette forme S est caractérisée dans les localités de savane par des arrangements chromosomiques correspondant à la forme Savanna (arrangements 2Rb/b, 2Rbcd/b, 2La). Quelques spécimens S ont présenté des arrangements correspondant à des hybrides entre les formes chromosomiques. Il faut noter que la forme moléculaire S correspond au type I d'*An. gambiae* défini par Gentile *et al.* (2001) sur la base d'une substitution de 3 nucléotides dans la région des ITS de l'ADN ribosomal.

1.3.2.- La forme moléculaire M

Elle est caractérisée par une fréquence élevée de l'arrangement standard typique de la forme chromosomique Forest dans les localités du Sud. Aucune inversion dans notre étude n'a été observée dans ces localités mais ceci peut être dû en partie au faible nombre de spécimens appartenant à cette forme par rapport au nombre de spécimens de forme S dans ce travail. Dans la région de savane et particulièrement à Tibati, elle est caractérisée par des arrangements chromosomiques du bras 2R correspondant à la forme chromosomique Mopti (2Rbc/+, 2Rbc/u et 2La) ou à des arrangements correspondant à la forme Savanna (2Rb, 2Rbcd et 2La). Cette forme moléculaire M correspond au type II d'*An. gambiae* défini par Gentile *et al.* (2001).

1.3.3- Isolement reproductif entre les deux formes moléculaires

Comme l'a montré Touré *et al.*, (1998), l'isolement reproductif entre les formes moléculaires M et S est indépendant de leur forme chromosomique. La correspondance entre Mopti et M tout comme entre Savanna et S n'est valable qu'au Mali et au Burkina Faso et hors de ces deux pays elle ne tient plus. Notre étude l'a également prouvé surtout qu'au sein de la forme chromosomique Forest nous retrouvons les deux formes moléculaires M et S, tout comme au sein de la forme Savanna. De ce fait, les systèmes de reconnaissance intervenant avant le croisement entre individus ne sont pas du tout liés aux inversions situées sur le bras droit du chromosome 2. Par contre, ces inversions seraient plutôt les sites pour des déterminants génétiques gouvernant l'adaptation à l'environnement du moustique.

L'absence d'hybride dans notre étude sur 1521 spécimens est indicatif d'un fort isolement reproductif entre les formes moléculaires M et S. Ce fait a été aussi mis en évidence par d'autres auteurs (Della Torre *et al.*, 2001, 2002) à l'échelle du continent ou uniquement dans d'autres pays (Tripet *et al.*, 2001). Des études ont été menées pour comprendre les mécanismes impliqués dans cet isolement.

I.3.3.1- Hypothèse d'isolement post-zygotique

Des études de croisements expérimentaux des spécimens de différentes formes chromosomiques au laboratoire ainsi que des deux formes moléculaires ont été conduites (Touré *et al.*, 1998 ; Persiani *et al.*, 1984 ; Di Deco *et al.*, 1980 et Favia *et al.*, 1997). Elles ne signalent pas la présence de barrières post-zygotiques entre ces formes. En effet, les hybrides de la génération F1 sont bien viables et vigoureux avec des organes reproducteurs normaux et aucune déviation n'a été observée par rapport au sex-ratio 1:1. Aucune stérilité ou réduction de la viabilité n'a été observée dans les descendances issues des différents croisements. Il faut signaler que ces études ont surtout été menées sur les formes chromosomiques plus que sur les formes moléculaires. Par contre, il y a une indication d'une non-viabilité des hybrides sur le terrain en ce sens que des études ont rapporté une décroissance au niveau de la proportion d'hybrides de la copulation jusqu'au stade adulte en passant par le stade larvaire. En effet, Tripet *et al* (2001) ainsi que Taylor *et al.* (2001) ont indiqué que la proportion de croisements entre les deux formes M et S était de 1,2%, celle d'hybrides au stade larvaire de 1,1% alors que seulement 0,3% d'hybrides étaient rencontrés chez les adultes. Cette observation constitue une indication de ce qu'un isolement post-zygotique pourrait s'être établi et jouer un rôle dans le processus de spéciation entre M et S chez *An. gambiae*.

I.3.3.2- Hypothèse d'isolement prézygotique

Trois mécanismes d'isolement prézygotique non mutuellement exclusifs peuvent être considérés. Ce sont l'isolement saisonnier quand les différentes formes surviennent à des saisons différentes, l'isolement par l'habitat quand il existe une différence dans le choix du site de reproduction (couplé à une étroite association entre les sites de reproduction et de croisement) et enfin comme troisième mécanisme, l'isolement éthologique tel que les différences dans les systèmes spécifiques de reconnaissance intervenant avant le croisement entre individus. N'ayant pas prouvé dans notre étude l'existence des deux premiers mécanismes (absence de variation saisonnière entre les formes M et S et existence de zones de sympatrie entre elles), nous pensons que l'isolement éthologique entre la forme M et S serait le moyen par lequel l'isolement prézygotique opérerait. Les preuves de l'existence de l'isolement prézygotique entre les formes M et S ont été clairement fournies par Tripet *et al.* (2001). En effet, ils ont étudié le sperme contenu dans la spermathèque de 251 femelles collectées dans une même localité du Mali et appartenant aux deux formes moléculaires M et S. L'objectif de cette étude était de déterminer la source du sperme (quelle était la forme moléculaire du mâle qui a copulé la femelle). Les résultats qu'ils ont obtenus indiquent que

parmi les femelles de la forme M, 2/195 (1,03%) seulement possédaient du sperme de la forme S et que parmi 55 femelles de forme S, 1 seul sperme (1,82%) provenait d'un mâle de la forme M. Ces résultats montrent donc que les croisements entre individus de formes moléculaires différentes sont très limités et confirment que les individus de formes M et S ne se croisent pas de manière aléatoire, mais plutôt en fonction de leur forme moléculaire. Nous pensons que des différences au niveau des mécanismes de reconnaissance intervenant avant les croisements entre individus peuvent en être responsables. Parmi ces mécanismes, nous pouvons mentionner les modes de parade (type de vol avant la copulation) entre mâles et femelles qui pourraient être différents ou alors les différents stimuli émis par les partenaires qui se seraient différenciés entre les deux formes moléculaires au cours de la spéciation. Milligan *et al.* (1993) ont mis en évidence une différence de profil d'hydrocarbures cuticulaires au Mali entre les formes chromosomiques Mopti et Savanna qui dans ce pays correspondent respectivement aux formes M et S. Ce profil de différenciation est également observé entre les espèces jumelles de drosophiles (Coyne & Orr, 1989).

I.3.4- Impact sur la transmission du paludisme

Les études entomologiques menées dans quatre localités différentes (chacune des deux formes étant prédominante dans 2 des localités), indiquent une implication assez semblable de chacune des formes moléculaires M et S dans la transmission du paludisme. En effet, les indices circumsporozoïtiques des deux formes sont assez proches entre 5,2 et 7,4 à l'exception de celui de la population de forme S de Dschang qui est de 1,7. Les taux d'infection obtenus dans notre étude par ELISA sont tout de même un peu plus élevés que ce que la dissection des glandes salivaires pourrait fournir. En effet, Fontenille *et al.* (2001) ont montré qu'il existe un rapport de 1,12 entre les résultats obtenus par ELISA et ceux obtenus par la dissection des glandes salivaires. La faible valeur de l'indice d'infection à Dschang serait davantage le résultat de l'impact de l'environnement sur les populations d'anophèles qu'un trait propre à la forme S. En effet, cette localité étant située en altitude (1400m), la température ambiante y est modérée (18-20°C) et de ce fait, allonge le temps d'incubation extrinsèque du parasite chez les anophèles et diminue ainsi l'espérance de vie infectante des anophèles. Ces conditions entraînent dans cette localité une transmission du parasite plus faible chez les populations humaines et en retour une faible infection des moustiques lors de leur prise de repas de sang. La différence observée au niveau des taux d'inoculation dans les différentes localités est due à la différence des densités des populations d'*An. gambiae*. En effet, dans la localité de Nkoteng, *An. gambiae* est minoritaire par rapport à un autre vecteur *An. funestus*, et face à cette faible agressivité, le nombre de piqûres infectantes transmises par

An. gambiae ne peut être que faible sans que cela soit une caractéristique propre à la forme S. Le taux d'inoculation observé à Dschang subit quant à lui l'effet de son faible indice circumsporozoitique. Les populations des deux formes sont toutes infectées par *P. falciparum* à l'exception de deux cas de *P. malariae*. Toutes les deux formes sont fortement anthropophiles dans les quatre localités, mettant ainsi en évidence la forte association des deux formes moléculaires d'*An. gambiae* à l'homme ce qui augmente leur rôle dans la transmission du paludisme.

Ainsi, l'isolement reproductif observé entre les formes moléculaires M et S ne semble pas se traduire par une différence significative au niveau de leur capacité vectorielle et de leur implication dans la transmission du paludisme. Ceci n'est pas très surprenant si ces deux formes proviennent de la subdivision en deux pools géniques de l'espèce mère bien adaptée à l'homme. Ce processus de spéciation entre M et S apparaît aussi comme une différenciation écotypique au sein d'une espèce, ce qui contribue à augmenter la diversité du système vectoriel et sa flexibilité à exploiter les ressources environnementales. De ce fait, ce processus de spéciation va dans le sens d'une augmentation de la stabilité globale de la capacité vectorielle telle qu'observée à Simbock avec un taux d'inoculation élevé et constant au sein de la population vectorielle (Antonio-Nkondjio *et al.*, 2002).

CHAPITRE II

Structure génétique d'*An. gambiae*
au Cameroun

II.1- Rappel de la problématique

De nombreuses études ont été menées dans le passé pour une meilleure compréhension de la complexité génétique chez *An. gambiae*. En effet, comme nous l'avons signalé dans la première partie de ce mémoire, cette espèce comprend 5 formes chromosomiques et est également subdivisée en deux formes moléculaires M et S. Les travaux précédents ont révélé une hétérogénéité génétique entre les formes chromosomiques Mopti, Savanna et Bamako. Néanmoins, l'isolement reproductif observé est incomplet. Les hétérocaryotypes correspondant aux hybrides Bamako-Mopti n'ont jamais été observés sur le terrain alors que des hybrides naturels Savanna-Mopti ont été trouvés plusieurs fois (Coluzzi *et al.*, 1985 ; Robert *et al.*, 1989 ; Lanzaro *et al.*, 1998). Au laboratoire, toutes ces formes chromosomiques se croisent entre elles et donnent une descendance viable et fertile (Di Deco *et al.*, 1980 ; Persiani *et al.*, 1984 ; Favia *et al.*, 1997). Après la définition des deux formes moléculaires par Favia *et al.* (1997), des études réalisées au Mali et visant à estimer les flux de gènes à l'aide des marqueurs microsatellites entre ces deux formes ont révélé qu'elles se différenciaient essentiellement au niveau des locus situés à l'intérieur des inversions chromosomiques (Lanzaro *et al.*, 1998). Cette différenciation a été expliquée par les effets de la sélection qui s'exerceraient sur ces locus étant donné le caractère adaptatif à l'environnement que confèrent les inversions. Ces auteurs ont donc argumenté contre une véritable subdivision d'*An. gambiae* en deux entités pouvant avoir le statut d'espèce. Par contre, les études menées par Chandre *et al.* (1999) dans le but de déterminer la fréquence du gène *kdr* chez les populations d'*An. gambiae* en Côte d'Ivoire ont révélé que seule la forme moléculaire S possédait ce gène de résistance aux pyréthrinoides et que la forme M en était dépourvue. Ceci a été aussi le cas dans les localités où les deux formes étaient trouvées en sympatrie. Ce résultat mettait ainsi en évidence un isolement reproductif entre les formes M et S contrairement aux travaux effectués au Mali par Lanzaro *et al.* (1998). En outre, ces auteurs n'ont trouvé aucun hybride M/S sur le terrain en effectuant des PCR-RFLP, renforçant ainsi cette observation d'une forte hétérogénéité entre M et S.

Face à ce schéma contradictoire de la structuration génétique au sein de l'espèce *An. gambiae* entre les différentes formes chromosomiques comme entre les formes moléculaires, nous avons entrepris une étude au Cameroun afin d'estimer le niveau de flux de gènes entre les formes moléculaires M et S. Les études précédentes ont été menées sur des formes chromosomiques présentant des arrangements chromosomiques différents. Ces inversions, soumises à une pression de sélection environnementale, pourraient constituer un facteur de confusion pour des études visant à estimer les niveaux réels d'échanges géniques entre les

populations de formes moléculaires ou de formes chromosomiques différentes. La forme chromosomique Forest ne présente aucune ou très peu d'inversions, ce qui permet d'éviter une potentielle confusion due aux inversions dans le profil de différenciation génétique entre M et S. Cette forme chromosomique Forest comprend aussi les formes moléculaires M et S ce qui permet une estimation des flux de gènes entre ces deux formes sans la présence de biais causé par la différence de forme chromosomique entre les populations comparées.

L'hypothèse de cette étude est la suivante : si la forme M est en réalité reproductivement isolée de la forme S, l'on doit observer une discontinuité génétique au sein de la forme chromosomique Forest et ceci en particulier dans les zones de sympatrie. D'un autre côté, si la sélection environnementale sur les inversions est la principale force induisant l'hétérogénéité génétique au sein et entre populations d'*An. gambiae*, l'on devrait observer une homogénéité génétique au sein des populations de forme chromosomique Forest puisqu'elles vivent dans le même environnement.

Pour vérifier ces hypothèses, nous avons mené une étude de la structure génétique d'*An. gambiae* en choisissant des populations appartenant à 5 localités du Sud. Ces localités sont Simbock, Mbalmayo, Nkoteng, Dschang et Santchou. Ces populations appartiennent à la forme chromosomique Forest. On retrouve en sympatrie les deux formes moléculaires M et S à Simbock. Nous avons pour cette étude, utilisé comme marqueurs génétiques, des locus ADN microsatellites.

II.2- Résultats

II.2.1- Sélection des locus ADN microsatellites pour notre étude

La sélection des locus a été effectuée sur la base des données de la bibliographie et des paramètres spécifiques aux locus microsatellites. La caractérisation des locus microsatellites a été réalisée par Zheng *et al.*, (1996) qui ont pu caractériser 131 locus chez *An. gambiae*. Ces locus ont servi à la construction d'une carte génétique de cette espèce. Cette mise au point a été poursuivie par Wang *et al.*, (1999). Sur la base des travaux précédents sur *An. gambiae* et *An. arabiensis*, menés en Afrique de l'Ouest et de l'Est, nous avons présélectionné 26 locus au départ de l'étude et répartis sur tout le génome d'*An. gambiae* et à partir desquels nous avons choisi des locus pour notre étude. Les amorces de ces locus ont été obtenues par commande auprès d'une entreprise de biotechnologie (Eurogentec).

II.2.1.1- Résultats des tests sur l'ensemble des 26 locus

II.2.1.1.1- Le polymorphisme des locus

Il est important pour une étude de génétique de populations que les locus soient polymorphes afin d'accroître leur capacité à mettre en évidence la moindre différenciation ou variabilité génétique intra ou inter-populationnelle. La plupart des locus testés étaient polymorphes (3 à 14 allèles observés). Certains de ces locus polymorphes ont présenté des allèles à forte fréquence. C'est le cas du locus H803 qui présente deux allèles majeurs de fréquence 0.48 chacun contre 0.02 pour deux autres allèles observés à ce locus. Les locus à faible polymorphisme ou possédant des allèles à fréquence très élevée n'ont pas été sélectionnés (H29C, H803, H603).

II.2.1.1.2- Qualité de l'amplification

La moitié des locus a présenté une bonne amplification (13 locus sur 26 testés). Pour certains locus, ceci s'est réalisé dès le premier essai (H80, H750, H325, H170, H125, H417, H776). Pour d'autres par contre, il a été nécessaire d'adapter les conditions opératoires de l'amplification. Ainsi la température d'hybridation a été selon les cas soit élevée quand des bandes parasites étaient présentes (H93, H555, H786) ou abaissée quand l'amplification était faible (H525) ceci pour faciliter l'appariement des amorces sur la matrice. Nous avons aussi agi au niveau de la quantité d'ADN pour optimiser la qualité de l'amplification. Les quantités de Taq polymérase et de $MgCl_2$ ont été très souvent modifiées selon les locus pour améliorer la qualité (spécificité, sensibilité) de leur amplification. Ainsi, dans certains cas, cette quantité a été augmentée (H85, H197) ou bien diminuée (H93, H22C, H45C) selon que les bandes obtenues étaient respectivement très peu visibles ou alors très épaisses (voir figure 19 pour un exemple de profil de migration sur gel de polyacrylamide).

Plusieurs problèmes d'amplification ont été rencontrés :

i) Pour certains locus, malgré les diverses modifications opératoires effectuées, la qualité de l'amplification n'a pas été améliorée. Nous avons ici le cas des locus ayant présenté des bandes très épaisses et pour lesquels la distinction des différents allèles d'un spécimen s'est avérée très difficile. C'est le cas de H22C et de H45C. Quand nous avons essayé de déterminer les différents allèles et avons calculé le taux d'hétérozygotie, nous avons remarqué que ce taux était très élevé par rapport à la valeur attendue (ce qui suggère que des erreurs ont eu lieu pendant la détermination des allèles sur le gel).

ii) Une autre catégorie de locus a présenté une amplification partielle. En effet, avec ces locus, certains spécimens n'ont pas pu être amplifiés alors qu'ils l'étaient bien sur d'autres

locus. C'est le cas des locus H24D, H117, H525, H46. Certains locus quant à eux n'ont présenté aucune amplification et ceci pour tous les échantillons. C'est le cas pour H85, H197, H135.

iii) Nous avons observé pour certains locus, des échantillons présentant trois bandes alléliques de même intensité. Après de nouvelles amplifications et modifications des conditions opératoires visant à faire disparaître la bande parasite, les trois bandes ont toujours été retrouvées. Ceci est un réel problème car *An. gambiae* étant une espèce diploïde et les marqueurs sélectionnés étant de copies uniques, il ne peut avoir qu'un maximum de deux bandes alléliques chez un hétérozygote. Ceci a été le cas pour les locus H555, H49, H786, H125, H525.

Les locus n'ayant pas présenté une amplification ont été abandonnés. C'est le cas de H85, H197, H135. De même, les locus ayant présenté une amplification mais dont la détermination des allèles s'est avérée difficile voire impossible (car les bandes alléliques étaient très épaisses), ont été laissés de côté. C'est le cas des locus H45C, H22C, H53. Les locus pour lesquels certains individus n'ont pas été amplifiés alors que c'était le cas avec d'autres locus (signalant ainsi la présence d'allèles nuls), ont été laissés de côté afin d'éviter les biais dans l'interprétation des résultats (erreur au niveau de l'estimation de la différenciation génétique entre populations). Parmi ceux-ci, nous citons le H525, H24D, H117, H603.

II.2.1.1.3- Taux d'hétérozygotie

Les taux d'hétérozygotie des locus ayant présenté une amplification ont été calculés. Les proportions d'hétérozygotes observées (le nombre d'individus hétérozygotes sur le nombre d'individus total) se sont avérées dans la plupart des cas inférieures aux proportions d'hétérozygotes attendues calculées à partir des fréquences génotypiques. Ceci a été le cas pour la plupart des locus, sauf quelques exceptions telles que H803 ou H22C et H45C (pour ces deux locus nous signalons de possibles erreurs dans la détermination des allèles due à la grosse épaisseur des bandes sur le gel). Pour H417, la différence observée n'est pas significative ($H_e=69$; $H_o=71\%$). La plupart des locus ont présenté des taux d'hétérozygotie supérieurs à 50%. Ceux ayant présenté des taux inférieurs à ce seuil sont : H603 (30%), H117 (17%), H786 (48%). Nous avons porté notre choix sur des locus ayant montré une hétérozygotie assez élevée (50-70%) pour être à même de révéler les plus petites variabilités génétiques au sein ou entre des populations.

II.2.1.1.4- Localisation cytologique des locus sur le génome

Nous avons eu pour objectif de voir la distribution des locus embrasser tout le génome du moustique afin d'avoir une idée globale de la variabilité et de la différenciation génétique au sein et entre populations. Pour cette raison, nous avons choisi des locus sur chacun des trois chromosomes (X, II et III) (figure 18). Nous avons choisi un locus au moins sur chaque bras chromosomique. Au niveau de chaque chromosome, nous avons veillé à ce que les locus choisis soient à une distance génétique considérable les uns des autres pour éviter les cas de déséquilibre de liaison dû à la proximité pouvant introduire un biais dans notre étude par redondance d'un état donné (différenciation ou bien homogénéité génétique entre deux populations).

La localisation des locus par rapport aux inversions chromosomiques a également orienté la sélection. En effet, nous avons choisi des locus en grande partie en dehors des inversions chromosomiques afin d'éviter d'éventuels biais dûs à une sélection qui pourrait s'exercer au niveau des inversions (Lanzaro *et al.*, 1998). En effet, la localisation des locus à l'intérieur des inversions peut entraîner une surestimation de la différenciation génétique entre les populations étudiées.

II.2.1.1.5- Taille des allèles

Les locus ont été choisis aussi en fonction de la taille des allèles attendus. En effet, pour éviter les longues durées de migration, notre choix s'est porté sur des locus ayant des allèles de taille comprise entre 80 et 150 pb.

Ce travail nous a permis de sélectionner 10 locus pour l'étude de la structure génétique des populations d'*An. gambiae* au Cameroun (voir tableau 11). Par cette sélection, nous pensons que de nombreux biais ont pu être évités dans notre étude conformément aux critères présentés par Walton *et al.* (1998)

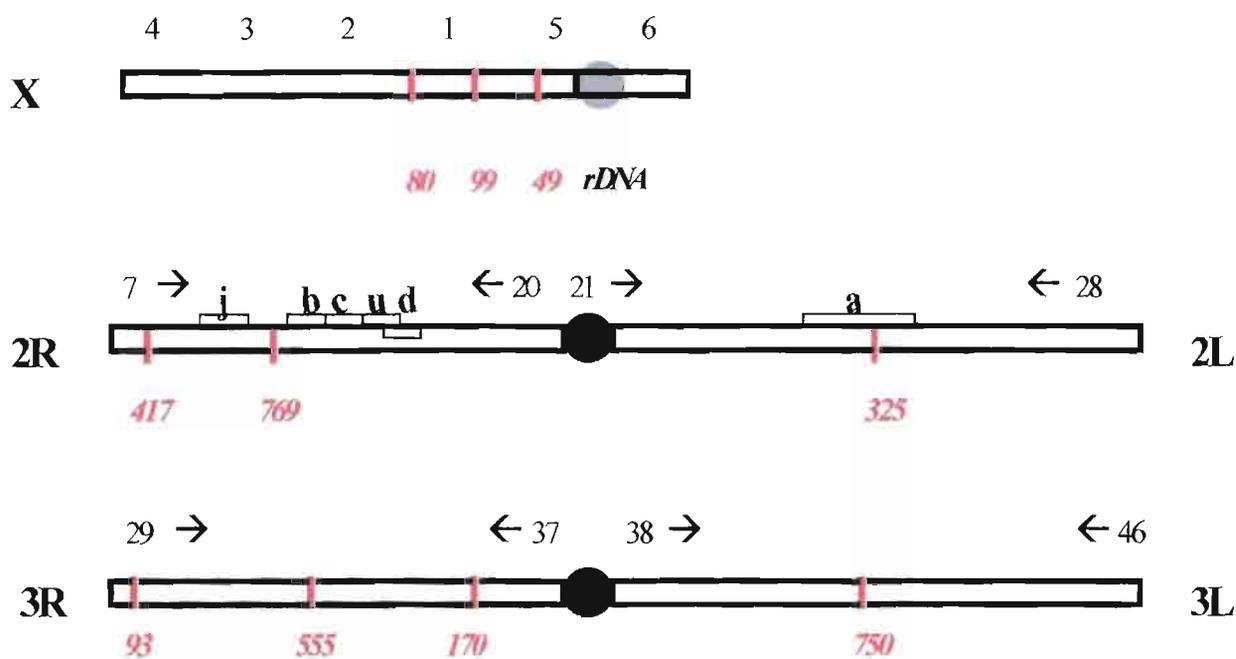


Figure 18 : Localisation cytologique des 10 locus microsatellites d'*An. gambiae* choisis.

Les chiffres représentent les divisions cytologiques des chromosomes alors que les lettres représentent les inversions chromosomiques les plus étudiées et diagnostiques de forme.

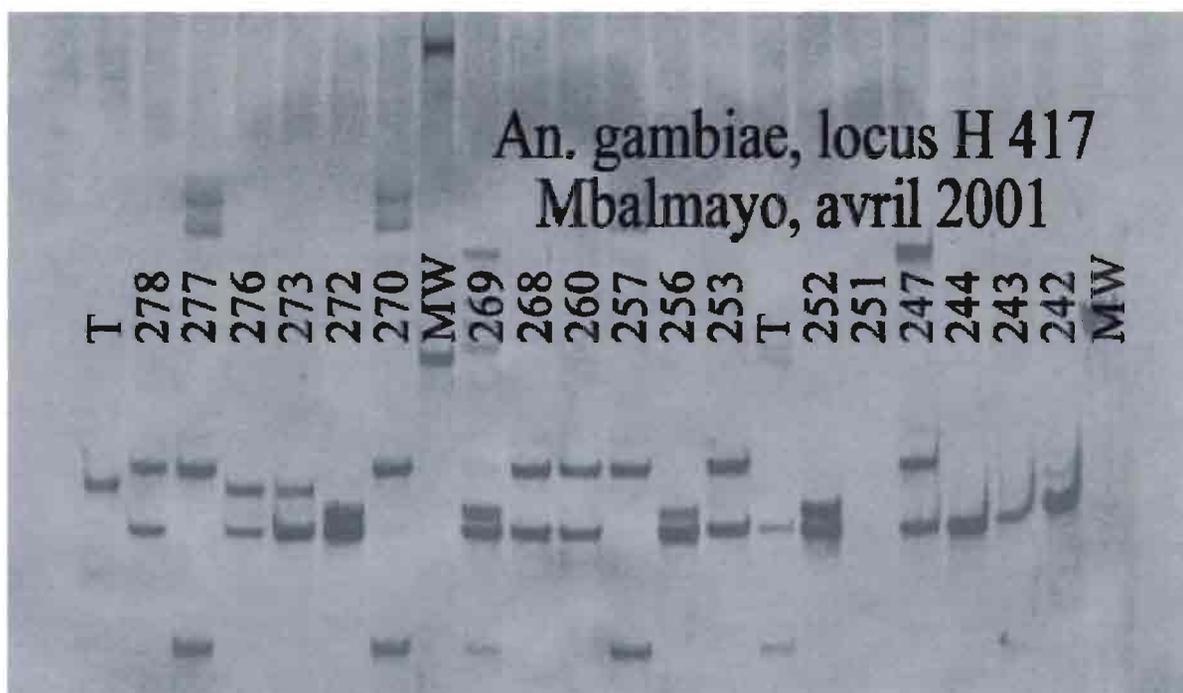


Figure 19 : Profil de migration du locus H417 sur gel de polyacrylamide

Chaque colonne représente un spécimen qui est homozygote en cas de présence d'une bande ou hétérozygote en cas de deux bandes

II.2.2- Structure génétique au sein des populations d'*An. gambiae* au Cameroun

Nous avons entrepris l'étude de la structure génétique d'*An. gambiae* au Cameroun en analysant l'ensemble des populations collectées sur le terrain. Nous avons collecté pour cette étude les populations d'*An. gambiae* provenant de 5 localités à savoir celles de Simbock, de Dschang, Nkoteng, Santchou et Mbalmayo. Notre choix a été orienté par le faciès écologique diversifié de ces localités ce qui permet de comparer les populations d'*An. gambiae* vivant dans des environnements différents. Nous avons au départ, procédé à une analyse des données sans tenir compte des différences de forme moléculaire entre les populations.

II.2.2.1- Caractérisation génétique des populations d'*An gambiae* étudiées

Toutes les populations de ces localités ont été génétiquement caractérisées afin de déterminer à quelle espèce du complexe *gambiae* et ensuite à quelle forme moléculaire d'*An. gambiae* elles appartenaient. Les résultats obtenus par PCR indiquent que les populations de toutes ces localités appartiennent à l'espèce *An. gambiae s.s.* Nous avons toutefois trouvé un individu appartenant à l'espèce *An. arabiensis* à Nkoteng. Il faut signaler que la localité de Nkoteng est située dans une zone de transition entre la forêt et la savane, donc la présence d'*An. arabiensis* n'est pas très surprenante puisque cette espèce est reconnue comme étant plus présente dans les zones sèches ou arides. Nous avons retrouvé la forme moléculaire S présente seule ou prédominante dans 3 localités (Santchou, Dschang, Nkoteng), la forme M était prédominante à Mbalmayo alors qu'à Simbock, nous avons trouvé les formes moléculaires M et S en sympatrie avec une prédominance de la forme M (90%) à Simbock. Des résultats et analyses plus détaillés ont été présentés dans le premier chapitre de ce travail. A partir de ces 5 localités, nous avons constitué 6 populations pour notre étude. Nous avons donc 4 populations de la forme moléculaire S (Dschang, Nkoteng, Santchou et Simbock). Nous avons également deux populations de la forme moléculaire M appartenant à la forme chromosomique Forest (Simbock et Mbalmayo).

Les génotypes des individus de ces populations ont été déterminés pour les 10 locus sélectionnés pour cette étude.

Tableau 11 : Caractéristiques des locus sélectionnés pour l'étude de la variabilité génétique de *An. gambiae*

Locus	Localisation cytologique	Distance Génétique en CM	Motif répété	Amorces	Taille moyenne des allèles
H80	X : 2B	37,7	(GT)9	5'TGCTCTCTCCTACATCGAGG 3' 3'GCCAGTGCTCTAGATTAACG 5'	81
H99	X : 2C	39,5	(GT)8	5'CGGGAATTTGTTGCTTCCTG 3' 3'TCGCCCTCTTTCTCCATCTC 5'	127
H49	X : 1D	40,8	(GT)5+ 4	5'CAGCGCCTCCATATAGAATG 3' 3'GATCATTGAGCTGAACCTGC 5'	97
H417	2R : 7	0.0	(GT)9	5'GTGCAGGAGATGTTTCCTACC 3' 3'GATATGGTGTTTAGGCCAG 5'	89
H769	2R : 9B	20,4	(TG)19	5'TTATGGGCAACTGCGGATGG 3' 3'AGTGTGCTATCCTTCGTCCG 5'	108
H325	2La : 23C-D	72,4	(GT)13	5'CCGGTGTCCGTGTTG 3' 3'GCGCGAAAGCAAATGACACG 5'	100
H93	3R : 29A	0	(GT)4+ 7	5'TCCCCAGCTCACCCCTCAAG 3' 3'GGTTGCATGTTTGGATAGCG 5'	150
H555	3R : 35B	37,1	(GT)8	5'GCAGAGACACTTCCGAAAC 3' 3'TGTCAACCCACATTTTGCGC 5'	81
H170	3L : 35B	45,9	(GT)10	5'TATACCCGATATCGTGCGCC 3' 3'GGAAAAAACATCCCCAAC 5'	108
H750	3L : 39A	68	(GT)8	5'GCAAAAAAGCTTCTCCCC 3' 3'TTAGCTACCGTCGACGCTTC 5'	87

II.2.2.2- Polymorphisme des locus microsatellites

396 spécimens d'*An. gambiae* ont été génotypés dans les 5 localités : 53 à Nkoteng, 54 à Dschang, 54 à Santchou, 49 de la forme M à Simbock, 33 de la forme S à Simbock, 53 à Mbalmayo.

Les dix locus microsatellites se sont avérés être très polymorphes au niveau de toutes les populations avec un nombre d'allèles distincts allant de 8 pour le locus H99 à 16 pour les locus H80, H170 et H769. Le nombre d'allèles le plus faible a été obtenu avec le locus H49 dans la population de Mbalmayo où seulement 4 allèles ont été observés. L'allèle majeur ou prédominant par locus n'a pas toujours été le même dans les différentes populations. Nous avons observé une certaine homogénéité entre les deux populations de la forme M en ce sens

que l'allèle majeur a été le même pour 8 des 10 locus dans les deux populations (Simbock et Mbalmayo) (Tableau 12). Par contre, les 4 populations S n'ont eu le même allèle majeur qu'au niveau d'un des 10 locus (H417). Beaucoup de locus ont aussi présenté plusieurs allèles à faible fréquence à l'exemple du locus H49 (cas de 7 allèles sur 12) et de H80 (8 sur 16). La fréquence des allèles prédominants (c'est-à-dire les allèles de plus grande fréquence au sein de la population) au niveau des différents locus est variable allant de 17 à 50% avec une moyenne autour de 30%. On observe aussi des allèles privés c'est-à-dire des allèles qui ne sont rencontrés que dans une seule population, au niveau de presque tous les locus sauf au niveau du H93. Chaque population possède des allèles privés, le nombre allant de 1 (Santchou et Simbock S) à 4 (Dschang).

Tableau 12 : Polymorphisme des locus microsatellites

Locus	Nombre d'allèles	Allèle majeur Population totale	Proportion allèle majeur	Nombre d'allèles de faible fréquence	Allèle majeur dans les 2 populations M	Allèle majeur dans les populations S
H80	16	Allèle 6	23%	8/16	Pas identique	Identique pour 2/4
H99	8	Allèle 1	31%	3/8	Identique	2/4
H49	12	Allèle 8	33%	7/12	Identique	3/4
H417	10	Allèle 6	50%	5/10	Identique	Identique
H769	16	Allèle 7	17%	5/16	Identique	2/4
H325	9	Allèle 2	34%	3/9	Identique (allèle 2)	Identique (allèle 6)
H93	11	Allèle 4	31%	2/11	Identique	3/4
H555	11	Allèle 2	30%	4/11	Identique	3/4
H170	16	Allèle 1	22%	6/16	Pas identique	3/4
H750	13	Allèle 2	19%	4/13	Identique	3/4

Pour les nombres d'allèles de faible fréquence, les valeurs entre parenthèses représentent les proportions de ces allèles au niveau de l'ensemble des spécimens échantillonnés dans l'étude.

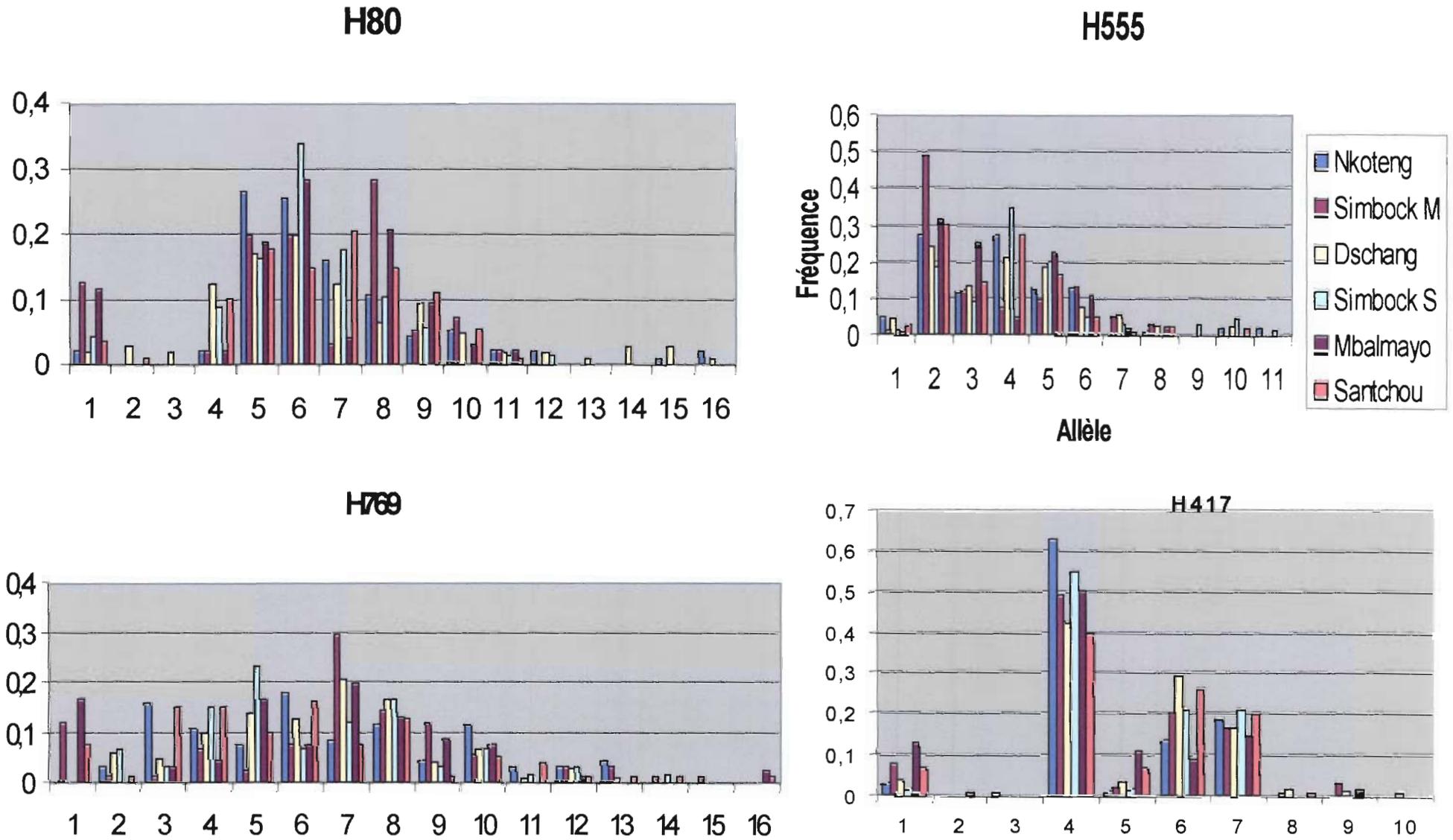
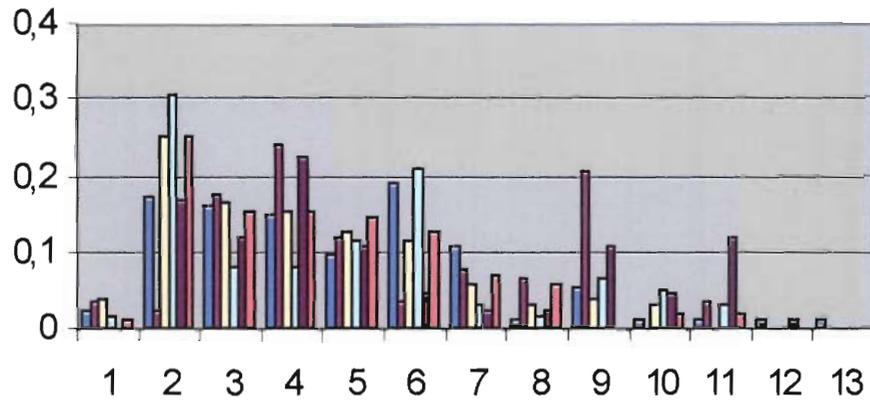
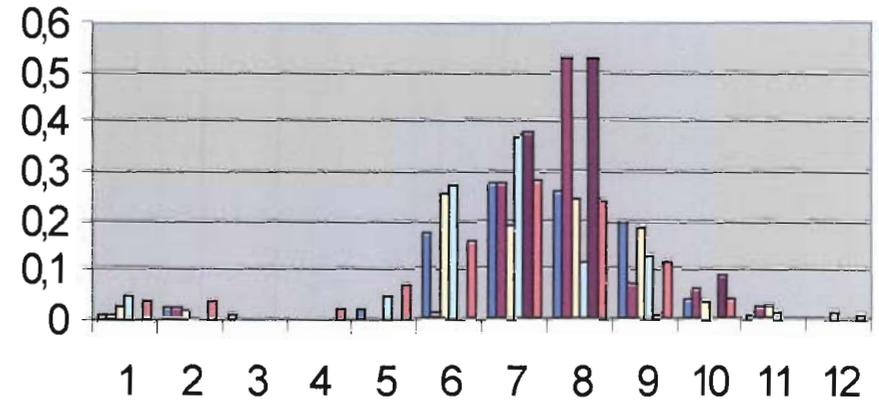


Figure 20 : Fréquences alléliques des 10 loci au sein des populations d'*An. gambiae* forme Forest 1 à 16 représentant le numéro des allèles

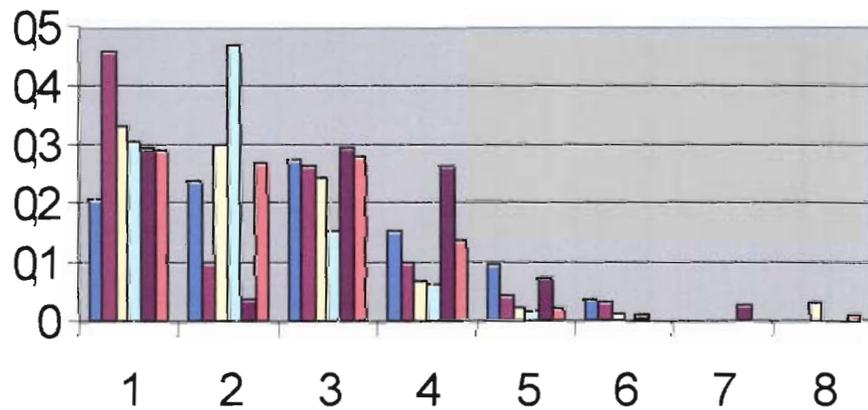
H750



H49



H99



H170

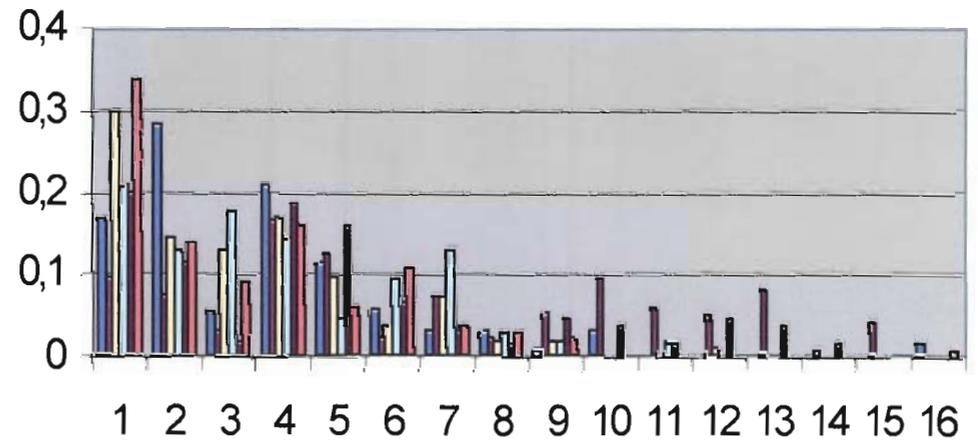


Figure 20 : Fréquences alléliques des 10 loci au sein des populations d'*An. gambiae* forme Forest
(suite)

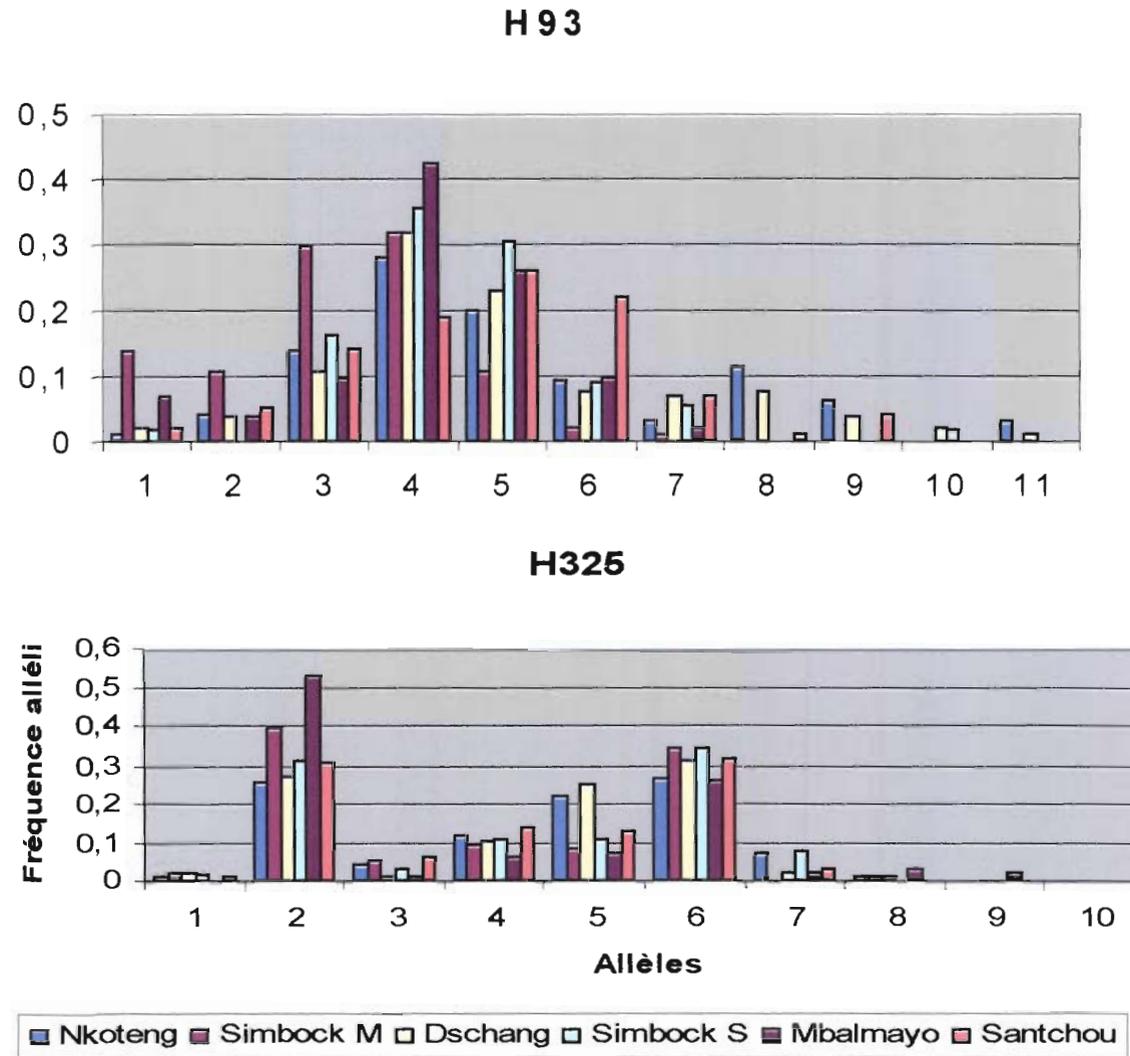


Figure 20 : Fréquences alléliques des 10 locus au sein des populations d'*An. gambiae* forme Forest (suite)

II.2.2.3- Variabilité génétique intra populationnelle

Le test d'homogénéité des fréquences alléliques entre toutes les populations indique que la distribution des fréquences alléliques n'est pas uniforme entre les différentes localités au niveau de tous les locus avec des valeurs de $P < 0,001$. Par contre, lorsque ce test est effectué entre les différentes paires de populations, nous observons une homogénéité de la distribution des fréquences alléliques entre les populations de la même forme moléculaire alors qu'une hétérogénéité est observée quand les populations appartiennent à des formes moléculaires différentes avec des valeurs de $P < 0,01$. Le test G de Goudet *et al.* (1996) nous a permis de tester la distribution des génotypes entre les différentes populations et nous avons eu des résultats identiques à ceux obtenus pour la distribution des fréquences alléliques. Le nombre moyen d'allèles par locus varie très peu d'une population à l'autre (8,3 à 10) et peut être classé selon la hiérarchie suivante : Dschang > Nkoteng > Santchou > Simbock M > Mbalmayo > Simbock S. Néanmoins il faut signaler que la population S de Simbock est de taille plus petite ($n=33$) que les autres populations ($49 < n < 54$), n étant le nombre de spécimens par localité. Lorsque toutes les populations sont groupées, cette moyenne s'élève à 12,2 allèles/locus. Le Tableau 13 (variabilité génétique) et la figure 20 (fréquences alléliques) présentent les résultats des 10 locus microsatellites que nous avons utilisés. Les différentes populations ont présenté une hétérozygotie observée (H_o) moyenne allant de 0,65 à 0,78. La population S de Simbock a une fois de plus présenté la plus petite valeur due à son effectif plus réduit. La population de Mbalmayo a seule présenté une hétérozygotie observée moyenne supérieure à l'hétérozygotie calculée (H_e) à partir des fréquences alléliques sous l'hypothèse d'un équilibre de Hardy-Weinberg. Les populations de la forme S ont présenté les plus fortes valeurs de H_e alors que les populations de la forme M présentaient les plus faibles.

II.2.2.4- Test de la panmixie

Nous avons utilisé un processus hiérarchique pour détecter la subdivision des populations en analysant la conformité de ces populations par rapport aux proportions attendues sous l'équilibre de Hardy-Weinberg. Ainsi, nous avons d'abord considéré que toutes nos populations ne formaient qu'une seule population panmictique donc appartenant au même pool de gènes. Sous cette considération, nous avons obtenu des valeurs de F_{is} positives au niveau de 9 locus sur 10 (seul H417 a eu un F_{is} négatif). Le déficit en hétérozygotes indiqué par les valeurs positives de F_{is} a été statistiquement significatif au niveau de ces 9 locus (Tableau 14) avec un seuil de $P < 0,01$ pour 6 d'entre eux. Cet excès d'homozygotes peut-être indicatif d'un important effet Wahlund suggérant une hétérogénéité du pool de

Tableau 13 : Variabilité génétique des différentes populations au niveau des 10 locus utilisés

<u>Locus</u>		Populations						
		Nkoteng (2n=106)	Dschang (2n=108)	Simbock S (2n=66)	Santchou (2n=108)	Simbock M (2n=98)	Mbalmayo (2n=106)	Total (2n=592)
AGXH80	N	12	16	9	10	9	9	16
	Ho	0,87	0,85	0,71	0,79	0,96	0,88	0,84
	He	0,83	0,89	0,82	0,87	0,82	0,82	0,84
AGXH99	N	6	7	5	6	6	7	8
	Ho	0,74	0,72	0,64	0,67	0,59	0,79	0,69
	He	0,80	0,74	0,67	0,75	0,70	0,76	0,74
AGXH49	N	10	9	7	10	8	4	12
	Ho	0,89	0,80	0,74	0,84	0,57	0,69	0,75
	He	0,80	0,81	0,77	0,82	0,64	0,58	0,74
AG2H417	N	6	8	5	6	7	7	10
	Ho	0,65	0,65	0,68	0,61	0,76	0,86	0,70
	He	0,56	0,71	0,62	0,73	0,69	0,70	0,67
AG2H769	N	12	12	12	15	14	11	16
	Ho	0,77	0,88	0,80	0,75	0,76	0,87	0,81
	He	0,90	0,88	0,88	0,87	0,86	0,88	0,88
AG2H325	N	8	8	7	7	7	8	9
	Ho	0,66	0,75	0,78	0,64	0,77	0,64	0,71
	He	0,80	0,76	0,77	0,77	0,71	0,65	0,74
AG3H93	N	10	11	7	9	7	7	11
	Ho	0,79	0,85	0,43	0,6	0,55	0,75	0,66
	He	0,85	0,82	0,76	0,82	0,78	0,73	0,79
AG3H555	N	9	11	10	10	8	8	11
	Ho	0,83	0,81	0,73	0,86	0,73	0,87	0,81
	He	0,81	0,84	0,81	0,79	0,72	0,77	0,79
AG3H170	N	11	10	10	10	15	14	16
	Ho	0,76	0,64	0,58	0,64	0,90	0,75	0,71
	He	0,84	0,84	0,88	0,82	0,92	0,88	0,86
AG3H750	N	13	10	11	10	10	11	13
	Ho	0,59	0,69	0,42	0,67	0,70	0,69	0,63
	He	0,88	0,86	0,85	0,84	0,85	0,88	0,86
Moyenne	N	9,7	10	8,3	9,2	9,1	8,6	12,2
Pour tous	Ho	0,76	0,76	0,65	0,71	0,73	0,78	0,73
Les loci	He	0,81	0,81	0,78	0,81	0,77	0,77	0,79

«Total» correspond à l'ensemble des populations mélangées

N : Nombre d'allèles par locus

2n : Nombre de chromosome

Ho : Hétérozygotie observée ; He : Hétérozygotie attendue sous l'équilibre d'Hardy-Weinberg (Nei, 1978)

gènes constitué par l'ensemble des populations. Toutefois, d'autres facteurs peuvent également en être responsables tels que la consanguinité entre les individus collectés ou la présence d'allèles nuls. Notre échantillonnage ayant été aléatoire, nous pouvons rejeter l'hypothèse de la consanguinité comme facteur responsable de cet excès d'hétérozygotes. Si c'est la présence des allèles nuls qui justifie cette déviation de l'équilibre de Hardy-Weinberg, les tests de déséquilibre de liaison devraient indiquer peu de paires de locus en liaison car les locus auraient les mêmes probabilités de présenter ces allèles nuls.

Ainsi, nous avons effectué les analyses de déséquilibre de liaison entre les paires de locus. Les allèles de différents locus sont en déséquilibre de liaison s'ils surviennent ensemble chez les individus avec des probabilités plus élevées ou plus faibles que celles attendues par association aléatoire. L'analyse par le test exact de Fisher indique que sur 45 tests effectués, 8 sont significatifs (soit 17,8%) avec des valeurs de probabilité allant de $P < 10^{-4}$ à 0,04. Une liaison physique des locus ne peut pas expliquer ces déséquilibres puisque ces locus liés ne sont pas toujours situés sur le même chromosome. Cette proportion élevée de déséquilibres de liaison entre les paires de locus sur la population globale permet d'éliminer la présence des allèles nuls comme facteur responsable du déficit observé en hétérozygotes. De ce fait, l'existence d'un effet Wahlund est certainement à l'origine du fort déficit en hétérozygotes que nous avons observé sur la population globale.

Parce que la distance géographique entre populations est une cause majeure pouvant conduire à une différenciation génétique, nous avons ensuite divisé la population globale en 5 sous-populations, suivant les localités considérées ici comme des unités reproductives de base.

II.2.2.4.1- Test de panmixie dans chacune des populations

L'équilibre de Hardy-Weinberg est respecté dans les populations de Nkoteng, Dschang, Santchou et Mbalmayo qui chacune appartiennent à une seule forme moléculaire. La plupart des locus dans ces populations se conforment aux attentes sous l'équilibre de Hardy-Weinberg (Tableau 14). Après l'application de la procédure de Bonferroni (Holm, 1979 ; Rice, 1989) pour prendre en compte le nombre de tests effectués, seuls quelques déficits en hétérozygotes statistiquement significatifs ont encore été observés au sein de ces populations au niveau de 1, 2 ou 3 locus suivant les populations. Pour la population de Mbalmayo, aucun locus n'a présenté de déviation par rapport à l'équilibre de Hardy-Weinberg. Le peu de déficits décelés au niveau des locus particuliers au sein de ces populations locales, découle plus des paramètres différents de ceux régissant la structuration génétique au sein des populations que d'un effet Wahlund indiquant la séparation du pool

génique constitué par chacune de ces quatre populations. Parmi ces paramètres, nous citons la présence d'allèles nuls au niveau des locus (Callen *et al.*, 1993 ; Lehmann *et al.*, 1996a ; Simard *et al.*, 2000), les phénomènes de sélection (Garcia de Leon *et al.*, 1997), l'amplification préférentielle de l'un des allèles par rapport à l'autre chez certains individus de génotype hétérozygote (Wattier *et al.*, 1998).

D'ailleurs, l'analyse du déséquilibre de liaison entre paires de locus au sein de chacune de ces populations ne soutient pas l'hypothèse de l'hétérogénéité de leur pool génique. Sur 45 tests effectués dans chaque population, nous avons observé après l'application de la procédure de Bonferroni, un maximum de 2 paires de locus qui présentait une liaison statistiquement significative, ce qui est dans l'ordre de ce qui pourrait être obtenu par chance uniquement ($P > 0,18$). Les populations de Nkoteng, de Mbalmayo et de Dschang n'ont présenté aucune paire de locus en déséquilibre de liaison alors que la population de Santchou n'a eu qu'une paire de locus en déséquilibre de liaison. Tous ces résultats confirment que ces différentes populations constituent des unités reproductives panmictiques.

La population de Simbock qui est composée d'individus appartenant aux deux formes moléculaires M et S, présente un déficit en hétérozygotes hautement significatif au niveau de six des dix locus. Cette observation indique que la population de cette localité est subdivisée en au moins deux pools géniques. Avec cette hypothèse, nous avons subdivisé la population de Simbock en deux parties constituées l'une des spécimens de la forme moléculaire M et l'autre des spécimens de la forme moléculaire S. Cette subdivision a été faite en supposant que l'existence de deux formes moléculaires au sein de cette population pourrait être la source de cette déviation par rapport à l'équilibre de Hardy-Weinberg tel qu'observé. Chacune de ces deux populations a aussi été testé pour la conformité à l'équilibre de Hardy-Weinberg. Les deux populations ont chacune présenté des déficits en hétérozygotes au niveau de trois locus, H49, H93 et H750 pour la forme M et H93, H750 et H170 pour la forme S. Ce niveau de déviation est comparable à celui observé au sein des autres populations. Aucun déséquilibre de liaison n'a été observé au sein des deux populations confirmant ainsi qu'elles sont chacune en équilibre de Hardy-Weinberg.

Chacune des 6 populations a présenté des déficits en hétérozygotes à certains locus, à notre avis dûs aux contraintes spécifiques à ces locus. En effet, les écarts de l'équilibre de Hardy-Weinberg sont fréquents dans les études de génétique de populations d'*An. gambiae* et *An. arabiensis* et d'autres organismes, lorsque les ADN microsatellites sont utilisés.

Tableau 14 : Test de déviation par rapport à l'équilibre de Hardy-Weinberg : Fis et valeurs de la probabilité P des tests d'équilibre de Hardy Weinberg, basés sur l'hypothèse nulle d'un déficit d'hétérozygotes pour toutes les populations d'*Anopheles gambiae*

Populations		Locus										Moyenne des locus	
		H80	H99	H49	H417	H769	H325	H93	H555	H170	H750		
Population globale	2n=592	+0,01	+0,093	+0,018	-0,027	+0,1	+0,066	+0,169	-0,004	+0,176	0,267	+0,087	
M	Mbalmayo	2n=106	-0.061	-0.045	-0.201	-0.237	+0.010	+0.012	-0.029	-0.122	+0.145	+0.212	-0.031
	Simbock	(2n=98)	-0.163	+0.159	+0.107	-0.091	+0.115	-0.081	+0,289	-0.024	+0.024	+0,184	+0.051
	Ensemble	2n=204	-0.111	+0.064	-0.042	-0.164	+0,066	-0.030	+0.148	-0.058	+0.088	+0,204	+0.016
Zone de sympatrie de M et S	Simbock	2n=164	-0.018	+0,163	+0,147	-0.094	+0.126	-0.058	+0,356	+0,080	+0,155	+0,340	+0,119
	Simbock	2n=66	+0.137	+0.052	+0.031	-0.093	+0.091	-0.022	+0,438	+0.097	+0,337	+0,505	+0,157
S	Nkoteng	2n=106	-0.053	+0.082	-0.113	-0.154	+0.143	+0.179	+0.063	-0.023	+0.096	+0,319	+0.054
	Dschang	2n=108	+0.047	+0,033	+0.017	+0.083	-0.002	+0.017	-0.029	+0.032	+0,232	+0,195	+0,062
	Santchou	2n=108	+0,079	+0,11	-0,025	+0,163	+0,167	+0,168	+0,277	-0,097	+0,222	0,213	+0,128
Ensemble	2n=388	+0,053	+0,081	-0,22	+0,41	+0,102	+0,095	+0,162	-0,006	+0,215	0,283	0,117	

Le *Fis* a été calculé selon Weir & Cockerham, 1984. Lorsque toutes les populations sont mélangées, $Fis=Fit$. Les valeurs de la probabilité P ont été corrigées selon la procédure de Bonferroni pour prendre en compte les tests multiples (10 tests par population). Valeurs en gras $P<0.05$; Valeurs en gras et soulignées $P<0.01$. 2n, Nombre de chromosomes étudiés.

II.2.2.5- Différenciation génétique entre populations

La différenciation génétique entre les populations d'*An. gambiae* a été estimée à l'aide des statistiques F (Wright, 1951), calculées selon Weir & Cockerham (1984). Les valeurs de la probabilité P associées aux différentes valeurs de Fst ont été estimées grâce au test exact de différenciation génotypique de Goudet *et al.* (1996).

II.2.2.5.1- Différenciation génétique au sein de chacune des formes moléculaires

Nous avons dans un premier temps évalué la différenciation génétique entre les populations appartenant à la même forme moléculaire.

II.2.2.5.1.1- Forme moléculaire S

Au sein de la forme chromosomique Forest, les résultats concernant les quatre populations de la forme S (Tableau 15) indiquent des valeurs de Fst très faibles, presque toutes non significatives à l'exception de quelques-unes au niveau de certains locus. En effet, sur 60 tests effectués entre les paires de populations au niveau de chacun des 10 locus, seuls 2 ont indiqué une hétérogénéité au sein de la forme S. Cette hétérogénéité est observée entre les populations de Nkoteng et de Simbock, au niveau des locus AGXH99 et AG3H93 à un seuil de $P < 0,05$. Sur les six comparaisons effectuées entre les populations de la forme S, les locus situés sur le chromosome X ont présenté la valeur moyenne de Fst la plus élevée à trois reprises, suivis des locus du chromosome 2 à deux reprises et un cas où les locus du chromosome 3 et 2 ont eu des valeurs similaires les plus élevées. Nous avons aussi estimé les différentes valeurs de Fst en excluant les locus qui ont présenté une déviation par rapport à l'équilibre de Hardy-Weinberg (AG3H750, AG3H170 et AG3H93). Les résultats obtenus ne sont pas différents des précédents. Cette même situation a été observée quand nous avons exclu les locus situés à l'intérieur des inversions chromosomiques et qui sont susceptibles d'être soumis à la sélection (AG2H325 et AG2H769).

II.2.2.5.1.2- Forme moléculaire M

Entre les 2 populations de la forme M (Simbock et Mbalmayo), le test effectué a mis en évidence une différenciation génétique significative au niveau de 3 locus sur 10, dont 2 à $P < 0,05$ et 1 à $P < 0,01$. La valeur de Fst se situe autour de 0,017. Ce sont les locus situés sur le chromosome 3 qui ont présenté la valeur moyenne de Fst la plus élevée (Fst = 0,024).

Tableau 15 : Différenciation génétique (estimée par Fst) entre les populations de la même forme moléculaire d'*An. gambiae*

Locus	A l'intérieur de la forme moléculaire S						A l'intérieur de la forme moléculaire M
	Simbock S & Santchou	Nkoteng & Santchou	Dschang & Santchou	Nkoteng & Dschang	Nkoteng & Simbock S	Dschang & Simbock S	Simbock M & Mbalmayo
AgXH80	0,013	0,011	0,001	<u>0,008</u>	0,0045	0.006	0.001
AgXH99	0,028	0,000	0,000	0.012	<u>0,0237</u>	0.014	0.031
AgXH49	0,014	0,000	0,009	0.002	0,0000	0.024	0.003
Chromosome X	0,019	0,003	0,003	0.007	<u>0.023</u>	0.014	0.012
Ag2H417	0,009	0,044	0,000	0.041	0,0162	0.006	0.008
Ag2H769	0,014	0,000	0,013	0.012	0,0021	0.000	0.009
Ag2H325	0,000	0,000	0,002	0.000	0,0119	0.004	0.012
Chromosome II	0,007	0,015	0,005	0.014	0.008	0.003	0.010
Ag3H93	0,013	0,013	0,015	0.000	0.008	0.000	<u>0.047</u>
Ag3H555	0,005	0,000	0,000	0.000	0.006	0.004	0.034
Ag3H170	0,003	0,020	0,000	0.018	0.020	0.000	0.004
Ag3H750	0,001	0,000	0,000	0.000	0.003	0.000	0.015
Chromosome III	0,005	0,011	0,005	0.002	0.009	0.000	0.024
Tous les locus	0,009	0,007	0,001	0.007	0.013	0.005	0.017
Moyenne pour 7 locus ^a	0,011	0,008	0,003	0.009	<u>0.014</u>	0.008	<u>0.014</u>
Locus situés hors des inversions ^b	0,010	0,010	0,003	0.008	<u>0.014</u>	0.006	<u>0.018</u>

Après application de la procédure de Bonferroni, valeurs en gras : $P < 0.05$; valeurs en gras soulignés : $P < 0.01$; ^a En excluant les loci présentant des déficits d'hétérozygotes significatifs au niveau de la population (i.e., loci Ag3H93, Ag3H170, Ag3H750) ; ^b En excluant les loci situés à l'intérieur des inversions polymorphiques (i.e., loci Ag2H769 et Ag2H325)

Les différentes valeurs de F_{st} obtenues au niveau de chaque locus entre les populations de la même forme moléculaire ont donc été très faibles et pas significatives à l'exception de quelques-unes. Néanmoins, lorsque les valeurs moyennes de F_{st} de tous les locus ont été estimées, elles ont été statistiquement significatives sauf entre les populations de forme S de Dschang et Simbock. Ceci indique que malgré la faible valeur des F_{st} ($0,002 < F_{st} < 0,017$), ces populations présentent néanmoins une différenciation génétique entre elles et constituent ainsi des pools géniques différentes

II.2.2.5.2- Différenciation génétique entre les populations de formes moléculaires différentes

8 paires de populations ont été comparées (tableau 16). Au niveau de tous les locus à l'exception de AG2H417 et dans une moindre mesure de AG2H325, les valeurs de F_{st} sont élevées et significatives à un seuil de $P < 0,01$ contrairement au cas de paires de populations de même forme moléculaire. L'exclusion des locus ayant présenté une déviation par rapport à l'équilibre de Hardy-Weinberg ou ceux situés à l'intérieur des inversions chromosomiques ne modifie pas le profil de différenciation observé. Cette observation indique que cette différenciation génétique entre les populations de formes moléculaires différentes, est un phénomène englobant tout le génome. C'est ainsi que les valeurs moyennes de F_{st} sur tous les locus sont situées entre 0,032 et 0,058 dans le cas des paires de populations allopatriques alors qu'elle est plus élevée [0,06 ($P < 0,01$)] entre les populations M et S de Simbock (en sympatrie).

II.2.3- Tests d'assignation pour l'évaluation de la différenciation génétique entre populations d'*An. gambiae*

Les études de génétique des populations ont presque toujours été menées en s'appuyant sur les données issues des fréquences alléliques obtenues au niveau des locus d'un marqueur génétique. De nos jours, la technologie moléculaire offre un grand nombre de marqueurs génétiques hypervariables (comme les locus microsatellites) qui permettent d'obtenir les génotypes des individus au niveau de plusieurs locus (multilocus). De ce fait, on a commencé à considérer comme unité génétique plutôt l'individu et non la population. Les fréquences alléliques au sein des populations sont ainsi remplacées par les génotypes multilocus des individus. En se basant sur le principe que les individus appartenant à la même population auront des génotypes proches, la différenciation génétique entre populations peut être évaluée en observant les proportions d'assignation des individus entre ces populations.

Tableau 16 : Différenciation génétique entre les populations de formes moléculaires différentes

Locus	Nkoteng & Simbock M	Nkoteng & Mbalmayo	Dschang & Simbock M	Dschang & Mbalmayo	Santchou & Simbock M	Santchou & Mbalmayo	Simbock S & Simbock M	Simbock S & Mbalmayo	Total S & Total M
AgXH80	0.030	0.016	0.037	0.023	0.030	0.025	0.041	0.014	0.025
AgXH99	0.044	0.029	0.029	0.063	0.028	0.038	0.102	0.140	0.047
AgXH49	0.068	0.094	0.095	0.131	0.063	0.082	0.151	0.167	0.104
Chromosome X	0.047	0.046	0.054	0.072	0.040	0.048	0.096	0.106	0.058
Ag2H417	0.012	0.022	0.001	0.035	0.001	0.024	0.000	0.021	0.008
Ag2H769	0.046	0.035	0.018	0.013	0.044	0.024	0.046	0.019	0.023
Ag2H325	0.023	0.059	0.021	0.061	0.000	0.034	0.000	0.032	0.032
Chromosome II	0.029	0.039	0.014	0.035	0.015	0.027	0.016	0.024	0.022
Ag3H93	0.037	0.021	0.040	0.008	0.061	0.037	0.039	0.000	0.013
Ag3H555	0.049	0.044	0.049	0.025	0.048	0.036	0.105	0.074	0.044
Ag3H170	0.033	0.013	0.035	0.011	0.048	0.015	0.027	0.017	0.017
Ag3H750	0.034	0.015	0.043	0.011	0.051	0.013	0.078	0.026	0.030
Chromosome III	0.038	0.023	0.042	0.014	0.052	0.025	0.063	0.030	0.026
Ensemble des locus	0.038	0.035	0.038	0.037	0.038	0.037	0.060	0.052	0.035
Moyenne pour 7 locus ^a	0.040	0.043	0.037	0.050	0.030	0.037	0.065	0.068	0.041
Locus situés hors des inversions ^b	0.040	0.032	0.042	0.038	0.041	0.039	0.069	0.058	0.037

Valeurs en gras : $P < 0.05$; Valeurs en gras soulignés : $P < 0.01$; ^a En excluant les locus présentant des déficits d'hétérozygotes significatifs au niveau de la population (i.e., locus Ag3H93, Ag3H170, Ag3H750) ; ^b En excluant les locus situés à l'intérieur des inversions polymorphiques (i.e., locus Ag2H769 et Ag2H325).

II.2.3.1- Principe

Les tests d'assignation se font suivant deux types de méthodes : la méthode des probabilités et celles des distances génétiques. Dans la première, les individus sont assignés à la population dans laquelle leurs génotypes ont la plus grande probabilité d'appartenir alors que dans la deuxième, les individus sont assignés à la population la plus proche génétiquement.

II.2.3.1.1- Méthodes basées sur la probabilité

II.2.3.1.1.1- La méthode des fréquences

Cette méthode qui a été mise au point par Peatkau *et al.* (1995), assigne un individu à la population dans laquelle il est plus probable que son génotype appartienne. En supposant que J locus indépendants ont été génotypés dans I populations, la fréquence de l'allèle k au locus j dans la population i est p_{ijk} . Sous l'hypothèse de l'équilibre de Hardy-Weinberg, la probabilité d'un génotype $A_k A_{k'}$ d'appartenir à la population i au locus j , est proportionnelle à $(p_{ijk})^2$ si $k=k'$ et à $2p_{ijk}.p_{ijk'}$ dans le cas contraire.

II.2.3.1.1.2- La méthode Bayésienne

Cette méthode est semblable à la précédente à la différence que la probabilité pour un individu d'appartenir à une population donnée, est estimée en s'appuyant sur les formules (particulièrement la formule 9) développées par Rannala & Mountain (1997). Cette méthode est considérée par plusieurs auteurs comme étant la plus performante dans bon nombre de situations (Cornuet *et al.*, 1999) pour correctement assigner les individus à leur population d'origine. Cette méthode est similaire à la méthode des fréquences et a été décrite par Rannala & Mountain (1997) qui ont utilisé une approche bayésienne pour détecter des immigrants grâce aux génotypes multilocus des individus.

II.2.3.1.2- Méthodes des distances

Cette méthode assigne un individu à la population la plus proche génétiquement et c'est pourquoi l'on parle de distance génétique entre un individu et les populations. Plusieurs distances génétiques ont été définies. La plupart d'entre elles sont définies entre populations (ex. distances de Nei, distance de Cavalli-Sforza & Edwards chord...) à l'exception d'une seule, la DAS pour «Shared allele distance», Chakraborty & Jin (1993) qui est définie entre les individus.

II.2.3.2- Assignation des individus des six populations d'*An. gambiae*

Nous avons cherché à voir si le profil de la structuration génétique observé entre les formes M et S d'*An. gambiae* en utilisant les statistiques F (Fst) basées sur les fréquences alléliques des populations était confirmé ou non lorsqu'une autre approche était considérée. Les tests d'assignation constituent donc cette nouvelle approche de l'étude de la structuration génétique chez *An. gambiae* en s'intéressant aux génotypes multilocus des individus des différentes populations.

Nous avons effectué des tests d'assignation en utilisant les 6 populations de la forme chromosomique Forest déjà testées avec les microsatellites. Cette assignation a été également effectuée en ne considérant que deux grandes populations, l'une comprenant les populations de la forme moléculaire M et l'autre celles de la forme moléculaire S. Nous signalons que ces six populations sont en équilibre de Hardy-Weinberg et qu'à l'exception de quelques paires de locus, la présence du déséquilibre de liaison entre les locus est négligeable. Nous avons utilisé les trois méthodes mentionnées ci-dessus avec néanmoins une préférence pour la méthode bayésienne compte tenu de la meilleure qualité des résultats qu'elle fournit. Les tests ont été effectués grâce au logiciel GeneClass mis au point par Cornuet *et al.* (1999).

Les résultats obtenus avec la méthode bayésienne sont présentés dans le tableau 16 et la figure 21.

298 individus ont été analysés. En utilisant la méthode Bayésienne, nous avons obtenu un pourcentage de bonne assignation des individus de l'ordre de 44,29% (c'est la proportion d'individus qui ont été assignés à la population d'où ils ont été échantillonnés). La méthode des fréquences fournit un pourcentage de 45,64% (136 bonnes assignations sur 298) et la méthode de distance génétique de Nei 43,29% pour 129 bonnes assignations sur 298. Tous les individus ont été génotypés au niveau des 10 locus que nous avons utilisés pour estimer la différenciation génétique entre les populations grâce aux statistiques F.

La répartition des individus incorrectement assignés ne se fait pas en relation avec la distance géographique notamment entre les individus des populations de forme moléculaire S. Par contre, l'on observe qu'en cas de mauvaise assignation d'un individu, il est assigné dans une population de la même forme moléculaire que sa population d'origine (dans 88% de cas pour la forme M et 94% pour la forme S ; Tableau 17C). En effet, 13 individus M sur 103 ont été assignés à une population de la forme moléculaire S alors que 12 individus de la forme S sur 195 ont été assignés à une population M. On note ainsi que lorsqu'on tient compte de la forme moléculaire, la proportion d'individus correctement assignés est très élevée, indiquant ainsi une homogénéité génétique au sein de chacune de ces formes moléculaires. Le fait que

sur 195 ont été assignés à une population M. On note ainsi que lorsqu'on tient compte de la forme moléculaire, la proportion d'individus correctement assignés est très élevée, indiquant ainsi une homogénéité génétique au sein de chacune de ces formes moléculaires. Le fait que très peu d'individus soient assignés dans une population appartenant à la forme moléculaire opposée indique et confirme l'hétérogénéité génétique observée entre les populations de la forme moléculaire M et celles de la forme S. Ce profil de structuration génétique est similaire à celui qui a été obtenu grâce aux statistiques F. Aucun individu de la population de la forme M de Simbock n'a été assigné à celle de la forme S de la même localité alors que 6 individus de cette population M étaient assignés à des populations de forme S distantes de plus de 250 km. Une fois de plus ceci va dans le même sens que ce que nous avons observé à l'aide des statistiques F où la plus grande valeur de F_{st} a été obtenue entre les populations de forme S et M de Simbock.

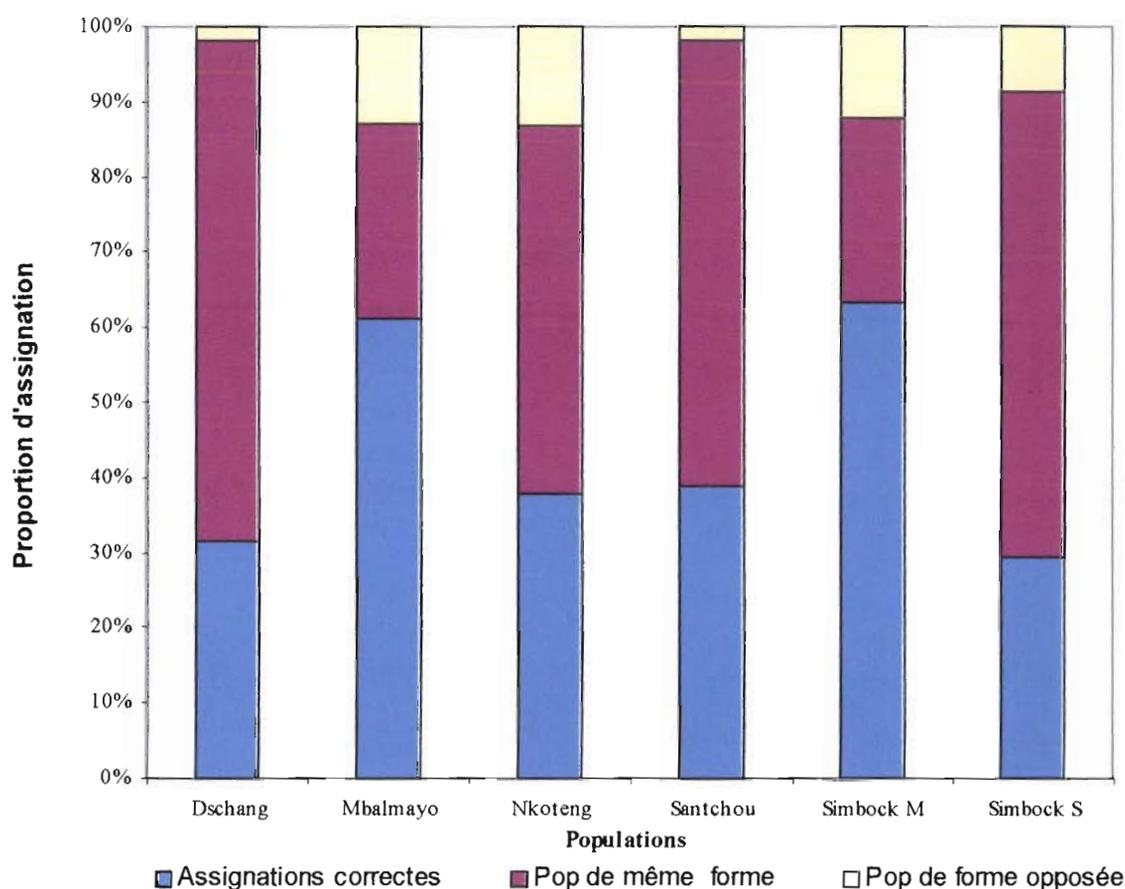


Figure 21 : Histogramme des proportions d'assignation des spécimens des différentes localités

Tableau 17 : Résultats des tests d'assignation de tous les individus dans les localités étudiées**A) Assignation globale en fonction des formes moléculaires**

	Assignations S	Assignations M	Total
Forme S	183 0,94	12 0,06	195
Forme M	13 0,12	90 0,87	103
Total	196	102	298

B) Assignation des individus par population

Localités	Dschang S	Simbock S	Nkoteng S	Santchou S	Simbock M	Mbalmayo M	Total
Dschang S	17 0,31	15 0,27	11 0,20	10 0,19	1 0,019	0 0,00	54
Simbock S	9 0,26	10 0,29	4 0,117	8 0,23	1 0,029	2 0,059	34
Nkoteng S	14 0,26	4 0,075	20 0,37	8 0,15	3 0,056	4 0,07	53
Santchou S	13 0,24	10 0,19	9 0,17	21 0,38	1 0,019	0 0,00	54
Simbock M	2 0,04	0 0,00	3 0,06	1 0,02	31 0,63	12 0,24	49
Mbalmayo M	1 0,0185	1 0,019	2 0,037	3 0,055	14 0,25	33 0,61	54
Total	54	34	53	54	49	54	298

nombre d'individus assignés indiqué en haut et proportion indiquée en bas

C) Assignation des individus des différentes localités selon les formes moléculaires M et S

Localités	Population de forme S	Population de forme M	Total
Dschang S	53 0,98	1 0,018	54
Simbock S	31 0,91	3 0,09	34
Nkoteng S	46 0,86	7 0,14	53
Santchou S	53 0,98	1 0,02	54
Simbock M	6 0,12	43 0,88	49
Mbalmayo M	7 0,13	47 0,87	54
Total	196	102	298

II.2.4- Etude de la corrélation entre les tests d'assignation et les statistiques F (Fst)

Nous avons cherché à déterminer quelle était la corrélation entre les résultats fournis par les tests d'assignation et les valeurs de Fst qui indiquent le degré de différenciation génétique. Le but de ce test était de déterminer le degré d'association entre ces deux paramètres aidant à l'étude de la différenciation génétique. Nous avons aussi recherché à savoir s'il y avait une corrélation entre les taux d'assignation des individus entre deux populations (à savoir le taux d'individus de la population 1 assignés à la population 2 et le taux d'individus de la population 2 assignés à la population 1) et la valeur de leur Fst.

Les résultats obtenus sont présentés dans le tableau 18 et la figure 22.

Les valeurs de MA1.2 (proportion d'individus de la population 1 assignés à la population 2) et de MA2.1 (proportion d'individus de la population 2 assignés à la population 1) ont été calculées pour chacune des 15 paires de populations. Les valeurs de MA.MA (proportion moyenne d'individus assignés de part et d'autre entre 2 populations) ont été obtenues en faisant juste la moyenne de MA1.2 et MA2.1 pour chaque paire de populations. Ces valeurs associées à celles de Fst ont permis d'obtenir la droite de corrélation telle que présentée par la figure 22. De même, les différents coefficients de corrélation obtenus sont indiqués dans le tableau 18. La droite de corrélation obtenue en comparant les valeurs de MA.MA et de Fst, présente une pente d'une valeur négative de $-0,1554$ et significative à un seuil de probabilité de $P < 0,0001$. Le coefficient de corrélation r est égal à $-0,86$ indiquant

ainsi une très forte corrélation entre les résultats obtenus par les tests d'assignation et les valeurs de F_{st} reflétant la différenciation génétique entre les populations. Cette corrélation est négative puisque, plus la valeur de F_{st} est élevée, moins important est le nombre d'individus assignés entre les deux populations concernées. Cette forte corrélation entre les tests d'assignation et les valeurs de F_{st} est obtenue aussi bien avec les valeurs de MA1.2 que celles de MA2.1. Il ressort aussi de ces résultats qu'il existe une forte corrélation entre les valeurs de MA1.2 et celles de MA2.1 c'est-à-dire que les proportions d'individus d'une population 1 assignés à une population 2 sont comparables à la proportion d'individus de cette population 2 assignés à la même population 1. Ainsi la proportion d'individus assignés de part et d'autre entre deux populations est comparable avec une valeur de probabilité $P < 0,0001$.

Tableau 18 : Corrélation entre proportion d'assignation et valeurs de F_{st}

	MA1.2	MA2.1	Fst	MA.MA
MA1.2	1			
MA2.1	0,92 <0,0001	1		
Fst	-0,82 0,0001	-0,86 0,0001	1	
MA.MA	0,98 <0,0001	0,98 <0,0001	-0,86 <0,0001	1

MA1.2 étant la proportion d'individus de la population 1 assignés à la population 2

MA2.1 étant la proportion d'individus de la population 2 assignés à la population 1

MA.MA étant la proportion moyenne d'individus assignés de part et d'autres entre 2 populations.

La valeur supérieure du tableau représente le coefficient de corrélation alors que celle inférieure représente la probabilité qui y est associée.

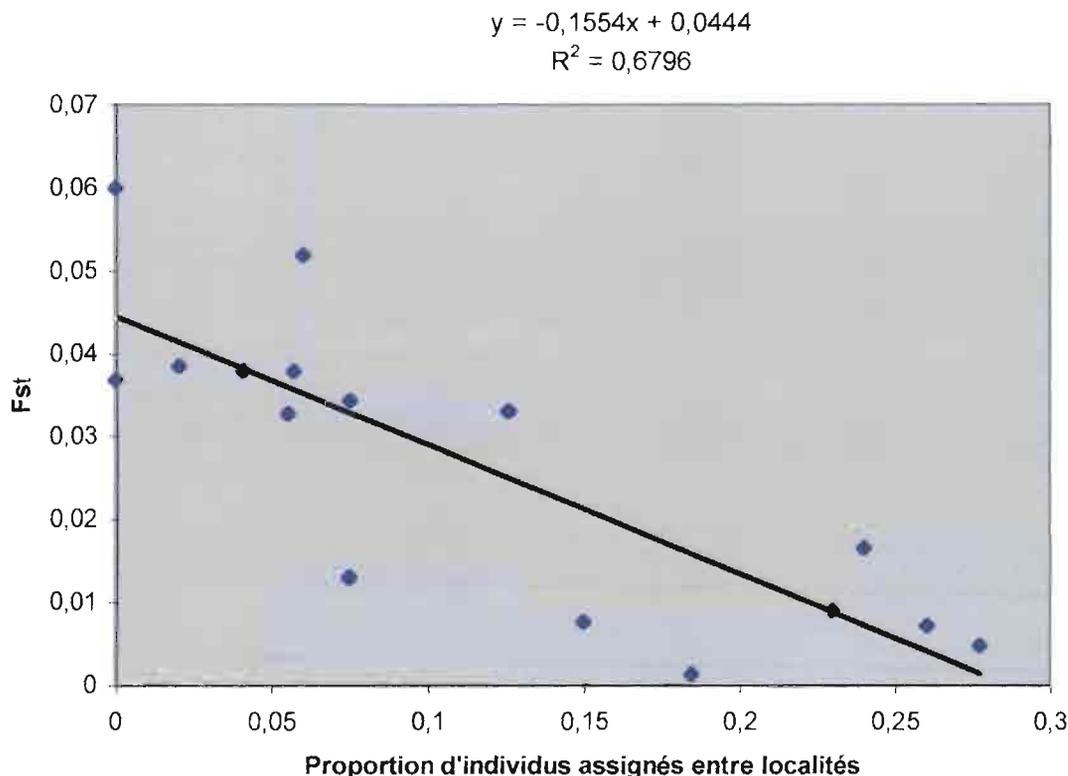


Figure 22 : Droite de corrélation entre les valeurs de F_{st} et les résultats des tests d'assignation

II.2.5- Etude de l'isolement par la distance entre les populations

Nous avons entrepris d'évaluer la corrélation qui existe entre la distance géographique et les différentes valeurs de F_{st} représentant la différenciation génétique entre les populations. Le modèle d'isolement par la distance (Wright, 1951 ; Slatkin, 1994) suppose qu'entre les populations, la migration s'effectue dans toutes les directions et que la distance est le principal facteur qui détermine la différenciation génétique. Ce modèle quantifie donc l'effet de la distance sur la différenciation génétique et permet de prédire le niveau de différenciation entre les populations en fonction des distances qui les séparent. Etant donnée l'hétérogénéité génétique mise en évidence entre les populations de forme M et S, nous avons entrepris d'évaluer le rôle que la distance géographique y jouait et si ce rôle était avéré, d'en estimer l'ampleur. Nous avons utilisé des paires de populations dans un premier temps sans tenir compte des formes moléculaires ou des formes chromosomiques et dans un second temps, les paires de populations ont été traitées selon leur forme moléculaire avec d'un côté une évaluation du modèle entre les paires de populations de la forme S et de l'autre une évaluation entre les paires de populations de la forme M. L'évaluation de ce modèle s'est fait suivant la relation qui existe entre les paramètres $F_{st}/1-F_{st}$ et le Logarithme de la distance géographique

Nous avons utilisé une droite de régression dont l'équation est de la forme $y=ax+b$. Les différentes droites obtenues sont présentées dans les figures 23 et 24.

En considérant toutes les paires de populations indépendamment des formes moléculaires, nous avons obtenu l'équation suivante :

$$F_{st}/1-F_{st} = -0,0059\text{LogD} + 0,0542$$

La pente de la droite de régression est égale à $-0,0059$. Cette droite de régression a été construite à l'aide de 15 paires de populations constituées entre les 6 populations de notre étude. Les différents points sont distribués de façon dispersée autour de la droite et non de manière linéaire. Le modèle d'isolement par la distance est rejeté ici avec une probabilité associée à la pente non significative de $P = 0,133$. Ceci indique donc que la différenciation génétique entre les populations indépendamment des formes moléculaires, n'est pas due à la distance géographique qui les sépare.

Au sein de la forme S, nous avons eu 6 paires de populations qui ont permis de construire la droite de régression dont l'équation obtenue est la suivante :

$$F_{st}/1-F_{st} = 0,0017\text{LogD} - 0,0011$$

La pente obtenue est positive et est égale à $0,0017$ mais elle n'est pas statistiquement significative ($P=0,43$.) L'on observe que les points sont assez dispersés par rapport à la droite tel que présenté sur la figure 24. Ceci signifie donc que même au sein des populations de la même forme (ici la forme S), la différenciation génétique dans notre étude n'est pas fonction de la distance. Toutefois, on remarque tout de même que la plus petite valeur de F_{st} a été obtenue entre les localités les plus proches de Dschang et de Santchou qui sont séparées seulement de 20 km. Nous notons aussi que les populations considérées ici ne sont pas situées de manière linéaire (voir carte du Cameroun, figure 5). Nous signalons aussi que le faible nombre de paires de populations dans cette analyse rend ce test moins robuste. Un nombre plus important de populations seraient nécessaire pour affiner l'étude.

Pour ce qui est de la forme moléculaire M, aucun résultat n'a pu être obtenu du moment où nous n'avons qu'une paire de populations à savoir Simbock M et Mbalmayo. Il faudrait davantage de populations pour effectuer cette analyse.

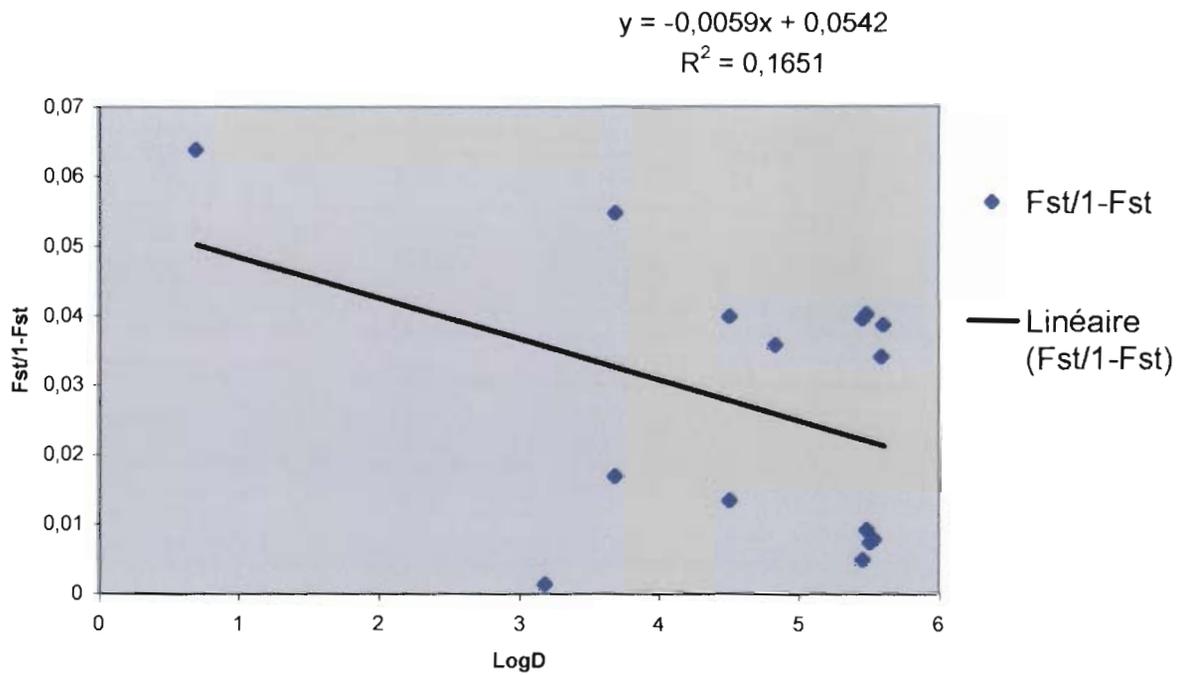


Figure 23 : Droite de régression illustrant l'isolement par la distance entre toutes les populations

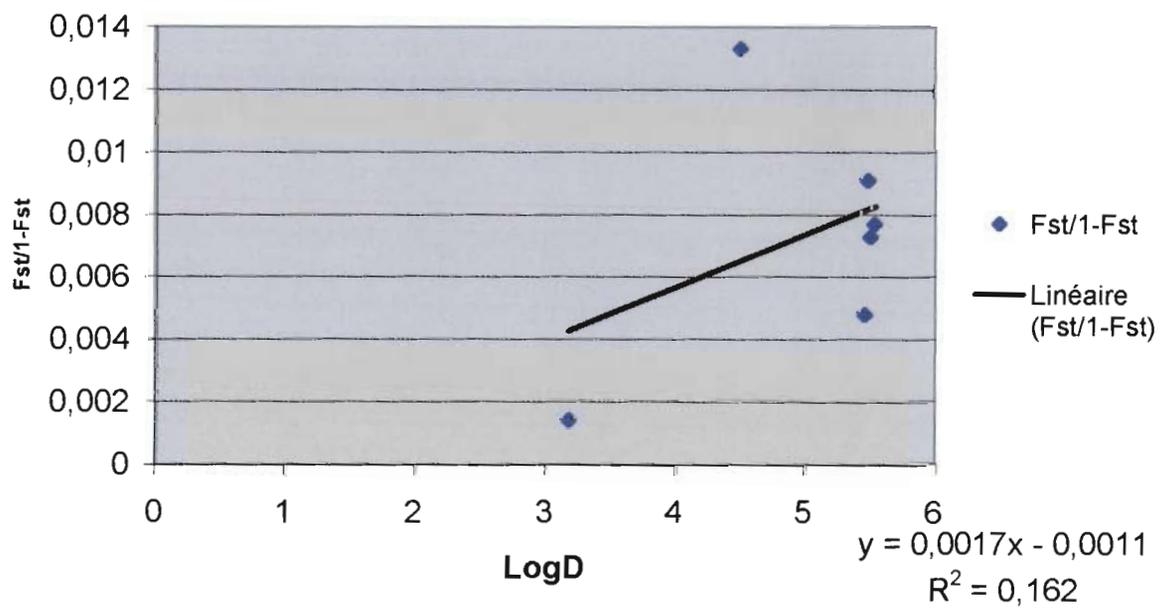


Figure 24 : Droite de régression de l'isolement par la distance entre populations de forme S

II.3- Discussion

Nous avons entrepris une étude de la structuration génétique des populations d'*An. gambiae* au Cameroun afin d'améliorer nos connaissances sur la grande complexité génétique que ce vecteur du paludisme présente sur toute son aire de répartition en Afrique. Nous avons choisi comme outil les loci microsatellites qui sont des marqueurs génétiques puissants de part leur capacité à révéler les faibles niveaux de structuration génétique au sein et entre populations. L'utilisation de ce marqueur nous a permis de tester un certain nombre d'hypothèses, de répondre aussi à des questions que nous nous posions au sujet d'*An. gambiae* au regard de sa complexité génétique.

II.3.1- Les locus d'ADN microsatellite comme outil de l'estimation de la variabilité génétique

Dans cette étude, nous avons testé 26 locus microsatellites dans le but de sélectionner les plus appropriés pour une étude de génétique des populations d'*An. gambiae* au Cameroun. Trois catégories de locus se sont dégagées de ce travail : la première, constituée de locus ayant présenté une bonne amplification, la deuxième constituée de locus présentant un défaut d'amplification pour certains spécimens et la dernière constituée de locus n'ayant présenté aucune amplification. Les différents locus testés ont pour la plupart présenté un taux d'hétérozygotie supérieur à 50%. Nous avons présenté les caractéristiques des différents locus testés notamment en précisant quel était leur degré de polymorphisme, leur qualité d'amplification et les différents problèmes rencontrés. Tout au long de ce travail, plusieurs observations ont été faites.

Au niveau des amplifications, certains locus n'ont présenté aucune amplification pour tous les spécimens. C'est le cas pour H135, H85, H197. Le problème pourrait se trouver au niveau de la séquence des amorces utilisées. Ces séquences ont été obtenues par Zheng *et al.* (1996) qui ont sélectionné ces amorces à partir des souches d'élevage d'*An. gambiae*. Ces souches d'élevage, suite à une pression de sélection et une dérive génétique, peuvent être différentes des populations sauvages que nous avons utilisées dans notre étude. L'hybridation de ces amorces sur la matrice d'ADN peut aussi être impossible empêchant ainsi toute amplification au niveau de ces locus. Par ailleurs, quelques échantillons n'ont pas pu être amplifiés à certains locus alors que ceux-ci l'ont été avec d'autres locus. Dans ce cas, nous pensons à l'existence d'allèles nuls au niveau de ces locus au sein de la population étudiée. Ces allèles sont dus à des mutations ponctuelles au niveau de la séquence des amorces et ainsi l'appariement de celles-ci sur la matrice d'ADN devient impossible lors de la PCR (Callen *et*

al., 1993). L'existence de ces allèles nuls induit une baisse du taux d'hétérozygotie et entraîne une déviation par rapport à l'équilibre de Hardy-Weinberg, même dans le cas d'une population panmictique. Lorsque ces allèles nuls sont détectés au niveau d'un locus, l'on peut synthétiser une séquence alternative pour l'amorce en question. C'est le cas pour le locus H46 (Lehmann *et al.*, 1996).

Nous avons également enregistré de très mauvaises amplifications qui ont rendu la détermination des différentes bandes alléliques très compliquée voire impossible (H22C et H45C). En effet, ces locus ont présenté des bandes très épaisses et il a été impossible de distinguer les individus homozygotes de ceux qui sont hétérozygotes. Malgré les ajustements visant à réduire l'épaisseur de ces bandes, le résultat a été inchangé (réduction de la quantité de DNA par une dilution au 1/10^{ème} voire au 1/100^{ème}, baisse de la quantité de la Taq polymérase, augmentation de la température d'appariement pour plus de spécificité de l'amplification). Les locus H22C et H45C sont pourtant des locus utilisés sans grande difficulté dans des études similaires ayant eu lieu sur d'autres populations du complexe *gambiae* soit en Afrique de l'Ouest (Lanzaro *et al.*, 1998) soit en Afrique de l'Est (Kamau *et al.*, 1998b ; Simard *et al.*, 1999).

La présence des triples bandes au niveau de certains locus tels que H555, H49 et H750 lors de l'amplification de certains spécimens, est probablement due à un problème d'amplification de structure secondaire de l'ADN (Queller *et al.*, 1993). Face à cette situation, nous avons fait le constat selon lequel la bande supplémentaire serait celle qui ne se retrouvait jamais en dehors du cas des triplets. Au niveau de tous ces locus, lorsque ces triplets apparaissent, c'est toujours avec un écart d'une paire de bases et 2 paires de bases respectivement entre la première et la deuxième bande et la 2^{ème} et 3^{ème} bandes.

II.3.2- Variabilité génétique des populations

Les locus microsatellites utilisés ont tous montré une forte variabilité génétique au niveau des 6 populations d'*An. gambiae* confirmant ainsi le fort degré de polymorphisme généralement reconnu pour ce marqueur génétique chez les populations d'anophèles (Lanzaro *et al.*, 1998). En effet, à cause de leur taux de mutation μ plus élevé, les microsatellites sont des outils puissants pour l'étude de la structuration génétique des populations contrairement à d'autres marqueurs comme les isoenzymes.

Toutes les 6 populations constituent des unités panmictiques, malgré l'existence de quelques déviations observées par rapport à l'équilibre de Hardy-Weinberg au niveau de certains locus dans la plupart des populations. Comme nous l'avons signalé plus haut, la

grande majorité de ces déviations se traduit par un déficit en hétérozygotes à ces locus. Parce que plusieurs spécimens n'ont présenté aucune amplification en PCR, et que plusieurs auteurs l'ont mis en évidence dans des études précédentes avec les microsatellites (Chakraborty *et al.*, 1992 ; Lehmann *et al.* 1996 ; Simard *et al.*, 2000), nous pensons que ces déviations peuvent se justifier par la présence d'allèles nuls.

Dans toutes les populations, les paires de locus analysées n'ont pas montré un nombre significatif de déséquilibre de liaison susceptible d'influencer le profil de différenciation génétique observé entre les populations (notamment par la redondance au niveau des locus liés d'un phénomène évolutif entre les populations comme la sélection ou la dérive génétique). Seuls quelques paires de locus ont été en déséquilibre mais chaque fois, la proportion de celles-ci était égale à celle susceptible d'être obtenue par chance uniquement en fonction du nombre de tests effectués. Ceci constitue d'ailleurs un élément renforçant l'observation selon laquelle toutes nos populations constituent des unités reproductives panmictiques.

Les résultats que nous avons obtenus grâce aux marqueurs microsatellites montrent clairement qu'*An. gambiae* est constitué de deux entités génétiques différentes que nous avons appelé «les formes moléculaires M et S». En effet, les valeurs de F_{st} obtenues indiquent une homogénéité génétique au sein des populations de la même forme moléculaire alors qu'une hétérogénéité génétique est mise en évidence chaque fois que nous avons comparé deux populations de formes moléculaires différentes. Il faut signaler que cette étude a été menée afin d'estimer le niveau réel de flux de gènes entre les populations de la forme moléculaire M et S sans une influence provenant de potentiels facteurs de confusion au premier chef desquels la présence d'inversions chromosomiques. C'est cette raison qui nous a poussé à choisir des populations de la forme chromosomique Forest qui ont un arrangement chromosomique standard sans inversion pour la très grande majorité des individus. Le profil de différenciation génétique entre M et S tel que présenté par nos résultats ne met pas en évidence des facteurs comme la sélection pour justifier le niveau de différenciation observé entre M et S. En effet, à chaque comparaison de paire de populations en vue d'estimer le niveau de flux de gènes, la valeur de F_{st} obtenue a été une contribution de tous les locus indépendamment de leur localisation cytologique que ce soit à l'intérieur ou à l'extérieur des inversions chromosomiques. Des travaux précédents ont eu à mettre au premier plan le rôle des inversions (Lanzaro *et al.*, 1998 ; Taylor *et al.*, 2001) ou une plus grande contribution d'un des locus microsatellites à la valeur de F_{st} obtenue (Wang *et al.*, 2001) pour justifier la restriction de flux géniques entre M et S.

Les études de génétique des populations chez *An. gambiae* ont en effet été menées par plusieurs autres auteurs dans le but de comprendre la structuration génétique des populations de cet important vecteur du paludisme en Afrique. Au Mali, Lanzaro *et al.* (1998) ont évalué les échanges géniques entre les formes chromosomiques Savanna et Mopti qui y correspondent aux formes moléculaires S et M respectivement. Contrairement à nos résultats, ils ont montré que la différenciation génétique observée entre ces deux formes moléculaires était causée seulement par certains locus et n'était pas uniformément répartie sur tout le génome à l'opposé de ce que nous avons constaté dans notre étude. En effet, ils ont montré que les deux formes M et S se différençaient plus au niveau des locus situés sur le bras droit du chromosome 2 et qu'au niveau des autres régions du génome, les populations de ces deux formes étaient presque homogènes. La différence avec nos résultats est suffisamment importante pour que nous nous interrogeons sur ce qui pourrait en être la cause. Il faut signaler que le bras droit du chromosome 2 est le siège d'un grand polymorphisme des inversions chromosomiques chez *An. gambiae* et que c'est sur la base de certains arrangements chromosomiques sur ce bras que les formes chromosomiques ont été définies. Ces inversions sont connues pour leur rôle dans l'adaptation des anophèles aux conditions environnementales auxquelles ils sont soumis. Par conséquent, ce sont des régions du génome soumises à forte pression de sélection. Les gènes situés dans ces inversions sont préservés de toute recombinaison génétique de telle sorte que les populations portant des arrangements chromosomiques spécifiques comme c'est le cas des formes chromosomiques, peuvent génétiquement évoluer différemment en ce qui concerne ces gènes. Nous pensons que c'est ce fait qui expliquerait le profil de différenciation génétique que ces auteurs ont mis en évidence et que la forte pression de sélection à laquelle les populations de différentes formes M et S sont soumises au Mali introduit un facteur qui brouille le niveau réel de flux de gènes entre M et S dans ces régions. Par contre dans notre étude, nous avons choisi des populations M et S de la forme Forest ; donc non soumises à l'influence de la sélection agissant au niveau des inversions. Les résultats obtenus soutiennent cette affirmation en ce sens que des deux locus situés dans les inversions, aucun n'a indiqué une valeur de F_{st} significativement différente de celle des autres locus situés hors des inversions. De ce fait, nous pensons que les résultats que nous avons obtenus reflètent le profil réel de différenciation génétique entre les formes moléculaires M et S d'*An. gambiae*.

II.3.3- Structure génétique des populations d'*An. gambiae* au Cameroun

La différenciation génétique entre les populations M et S a été plus forte au niveau des locus situés sur le chromosome sexuel X (H80, H49 et H99) qu'au niveau de ceux situés sur les chromosomes 2 et 3 qui sont autosomaux. Ce résultat est observé pour chaque paire de populations S et M comparées. Ce phénomène a été plus marqué dans d'autres études comme celles de Wang *et al.* (2001) et Lehmann *et al.* (2003), qui indiquent que certains locus situés sur le chromosome X sont le siège d'une plus forte différenciation génétique entre les populations de forme moléculaire M et S. En effet, l'étude menée par Wang *et al.* (2001) sur des populations du Mali en utilisant environ 25 locus microsatellites, a révélé une plus nette différenciation génétique entre M et S au niveau de 2 locus situés à proximité de l'inversion fixée Xag à savoir les locus H678 et H614 (divisions chromosomiques 5D et 6 respectivement). Il convient de signaler que les marqueurs permettant de distinguer les formes M et S se situent dans la même région autour de la division cytologique 6 de l'ADNr (F.H. Collins, communication personnelle). Black & Lanzaro (2001) ont suggéré que les séquences répétées en tandem de l'ADNr sont soumises à une évolution concertée qui contribue à l'apparition et à la fixation des génotypes différents au sein des différentes sous-populations malgré le fait que les flux de gènes entre elles demeurent non négligeables. Ce phénomène d'évolution concertée a été mis en exergue pour justifier la plus grande différenciation génétique entre M et S telle qu'observée par ces auteurs. Mais les recherches dans les banques de données génétiques incluant notamment les séquences cibles des clones BAC («Bacterial Artificial Chromosome»), ont indiqué que le locus H678 ne faisait pas partie de l'unité répétée (18S-ITS1-5,8S-ITS2-28S-IGS) de l'ADNr (Lehmann *et al.* 2003) et par conséquent l'évolution concertée ne peut pas expliquer la forte différenciation observée à ce locus. L'effet de la sélection sur la division 6 du chromosome X a aussi été mis en exergue pour expliquer cette plus forte différenciation génétique entre M et S observée au niveau des locus qui y sont situés (Lehmann *et al.* 2003). Cette hypothèse n'a pas encore été confirmée.

II.3.3.1- Phénomène de renforcement entre M et S à Simbock

Nous avons observé la plus forte différenciation génétique entre M et S lorsque les populations de ces deux formes moléculaires étaient sympatriques dans la localité de Simbock. Ce fait est assez surprenant pour que nous y portions un regard. En effet à Simbock, la valeur de F_{st} obtenue entre la population M et la population S a été de 0,06 avec la valeur de $P < 0,01$, ce qui constituait la plus grande valeur obtenue entre toutes les comparaisons de populations effectuées, notamment entre les populations allopatriques séparées de plusieurs centaines de km. Bien que la différence ne soit pas statistiquement significative (les

intervalles de confiance de F_{st} s'étalant sur une gamme assez large (Wondji *et al.*, 2002)), cette situation a déjà été trouvée par d'autres auteurs utilisant les profils d'hydrocarbones cuticulaires des spécimens d'*An. gambiae*. Milligan *et al.* (1993) ont en effet trouvé des différences diagnostiques dans les profils d'hydrocarbones cuticulaires entre les formes chromosomiques Savanna et Mopti dans une localité de sympatrie au Mali. Par contre, dans le cas où elles étaient en allopatrie la différence était moindre. Les hydrocarbones cuticulaires sont reconnus comme jouant un rôle dans les phénomènes de reconnaissance intervenant avant le croisement entre individus dans l'embranchement des arthropodes (Fletcher & Michener, 1987). Cette différenciation plus marquée dans les situations de sympatrie, implique un phénomène de renforcement qui se caractérise par une plus grande différenciation entre deux espèces jumelles lorsqu'elles cohabitent dans le même lieu que quand elles sont en allopatrie. Ce phénomène de renforcement pourrait être causé par la sélection naturelle qui pourrait rapidement augmenter la divergence dans les systèmes de reconnaissance avant les croisements entre individus au sein des populations d'espèces sympatriques en instance de spéciation. Le renforcement aboutira dans ce cas à une différence de plus en plus importante au niveau des caractères reproductifs entre les deux espèces en spéciation dans les zones de sympatrie que dans les zones d'allopatrie ou en d'autres termes, l'isolement sexuel est plus prononcé en zone de sympatrie qu'en zone d'allopatrie. Ce fait est connu sous l'expression de «reproductive character displacement» (Higgie *et al.*, 2000). Les études menées chez la drosophile ont montré des résultats similaires aux nôtres. En effet, Coyne & Orr (1989, 1997) en analysant 171 paires d'espèces de drosophiles, ont trouvé que les paires d'espèces ayant récemment divergé présentaient un plus grand isolement sexuel quand elles étaient en sympatrie qu'en allopatrie. Comme exemple, nous pouvons citer la paire constituée par *Drosophila serrata* et *D. birchii* (Higgie *et al.*, 2000). Néanmoins, il faudrait encore expliquer par quel mécanisme la sélection qui agit spécifiquement sur certains gènes comme les gènes d'hydrocarbones cuticulaires, parvient à se manifester à travers tout le génome comme nous l'avons vu dans notre étude avec les microsatellites.

II.3.3.2- Rôle de la distance dans les profils de différenciation observés

Aucune évidence d'un isolement par la distance n'a été trouvée dans cette étude entre les différentes populations quelque soit la forme moléculaire ou chromosomique. En effet, il n'y a eu aucune corrélation entre les distances géographiques et les différentes valeurs de F_{st} obtenues entre les paires de populations. D'autres études menées sur la structure génétique d'*An. gambiae* ont aussi relevé cette absence d'isolement par la distance (Kamau *et al.*,

1998a). En fait, cette absence d'isolement par la distance entre populations d'*An. gambiae* dans notre étude n'est pas surprenante en soi étant donné que cette espèce présente déjà une forte hétérogénéité génétique selon les deux formes moléculaires M et S. Le fait que la plus forte différenciation génétique ait été observée dans un site de sympatrie atteste bien de ce que d'autres facteurs autres que la distance influencent la structuration génétique chez *An. gambiae*. Par contre, au sein de chacune des formes moléculaires, nous ne saurions être aussi catégoriques dans la mesure où peu de populations ont été étudiées pour permettre d'avoir un test robuste. Le fait que la plus faible valeur de F_{st} ait été observée entre les localités les plus proches de Dschang et Santchou pourrait indiquer qu'un effet de la distance entre les différentes populations influencerait tout de même la structuration génétique au sein de chacune des formes moléculaires. Nous avons aussi observé des valeurs de F_{st} significatives, bien que faibles, entre les populations de la même forme et ceci pourrait aussi être le produit de la distance entre ces populations. En effet, la distance est parfois une barrière à un flux génique total entre populations, les individus d'une localité se croisant de préférence à d'autres de la même localité plutôt qu'à ceux d'une localité voisine ou plus éloignée. De ce fait, la dérive génétique est la principale force génétique qui concourt à une divergence génétique entre les populations de la même espèce. L'augmentation du nombre de populations dans chacune des formes moléculaires permettrait de mieux évaluer le phénomène d'isolement par la distance.

II.3.3.3- Les tests d'assignation comme indicateurs de différenciation génétique entre M et S

La question d'assigner des individus à des populations a été spécifiquement posée pour plusieurs types d'études notamment dans la biologie de conservation des poissons (Taylor *et al.*, 1994) ou des abeilles (Cornuet *et al.*, 1996), dans l'évaluation de la différenciation génétique chez les ours polaires (Peatkau *et al.*, 1995) ou même pour détecter les immigrations récentes au sein des populations humaines (Rannala & Mountain, 1997). Nous avons quant à nous utilisé ces tests dans le but d'évaluer le niveau de différenciation génétique entre les formes moléculaires M et S d'*An. gambiae*.

L'ensemble des assignations de tous les individus indique une proportion de bonnes assignations assez moyenne autour de 46%. Les populations de forme S ont eu une proportion de bonnes assignations plus faible que celle de la forme moléculaire M. En effet, pour les deux populations M à savoir Simbock et Mbalmayo, cette proportion de bonnes assignations se situe autour de 60% alors que pour les 4 populations de la forme S, elle se situe autour de

30% Cette différence peut s'expliquer par le nombre de populations de chaque forme. Plus le nombre de populations d'une unité génétique est élevé, moins la proportion de bonnes assignations par population est importante. Néanmoins, il faut remarquer que les deux populations M sont plus différenciées entre elles que ne le sont les populations S entre elles, ce qui pourrait augmenter la proportion de bonnes assignations au sein des populations de la forme M. La proportion modérée de bonnes assignations s'explique par le niveau modéré de la divergence génétique entre les populations comme l'indique les valeurs de F_{st} obtenues entre les 6 populations considérées. En effet, il a été prouvé que plus la valeur de F_{st} était élevée, plus la proportion d'individus correctement assignés était grande (Cornuet *et al.*, 1999 ; Fréville *et al.*, 2001). L'étude de la corrélation entre les assignations et les valeurs de F_{st} l'indique très clairement avec des coefficients de corrélation très élevés autour de 0,96.

Grâce à leur grande variabilité génétique, les microsatellites sont des excellents outils pour les tests d'assignation comme l'ont signalé certains auteurs (Waser & Strobeck, 1998). Il a été prouvé que les scores d'assignation sont influencés par la variabilité génétique du marqueur moléculaire utilisé et les meilleurs scores sont obtenus avec les marqueurs les plus variables (Cornuet *et al.*, 1999). Ceci a été vérifié avec les microsatellites qui sont hypervariables contre les isoenzymes qui ont une variabilité modérée, puisque la proportion de bonnes assignations est plus élevée avec les locus microsatellites (Estoup *et al.*, 1998). La proportion de bonnes assignations dans notre étude a pu aussi être influencée par le modèle de mutation de nos locus microsatellites selon que ce soit l'IAM (Infinite Allele Model) ou le SMM (Step-wise Mutation Model) qui est considéré. En effet, il est établi que lorsque des populations avec une taille effective N_e ont divergé depuis un nombre de générations donné, le niveau de différenciation est plus élevé pour des locus évoluant sous le modèle de l'IAM que pour ceux évoluant sous le modèle SMM. En effet, dans le modèle SMM, le phénomène d'homoplasie (absent sous l'IAM mais présent sous le SMM) tend à réduire les différences entre taxons et par conséquent induit une proportion de bonnes assignations plus faible. Il faut noter qu'il est admis que les locus microsatellites suivent plus le modèle de mutation SMM avec donc de possibles contraintes de taille qui augmente l'homoplasie (Estoup & Cornuet, 1999). Néanmoins, l'ensemble des résultats des tests d'assignation va dans le sens d'une subdivision nette d'*An. gambiae* en ses deux formes moléculaires formant deux pools géniques et séparés l'un de l'autre bien que proches sur le plan phylogénétique.

II.3.4- Statut taxonomique des formes moléculaires M et S : Représentent-elles deux espèces à part entière?

Depuis l'introduction du concept génétique pour définir les espèces (Dobzhansky, 1937), l'analyse génétique des populations naturelles a été importante pour différencier les espèces qui n'auraient pas pu être reconnues sur la simple base des traits classiques comme la morphologie. La preuve d'une distinction génétique ou d'une différenciation génétique entre les populations naturelles, particulièrement dans les zones de sympatrie, est souvent utilisée comme un critère pour ériger de nouvelles espèces. Le concept biologique des espèces définit les espèces comme étant des entités reproductivement isolées c'est-à-dire incapables d'échanger des gènes à travers les croisements sexuels. C'est sur la base de ce concept que nous nous appuyons pour répondre à la question du statut taxonomique des formes moléculaires M et S au sein de l'espèce *An. gambiae* s.s.

En nous basant sur nos résultats, nous pouvons clairement affirmer qu'il existe une barrière de flux de gènes entre les populations des formes moléculaires M et S. Cette observation est plus nette en zone de sympatrie comme à Simbock. Les profils de différenciation génétique entre ces deux formes indiquent tous que les formes M et S constituent deux unités génétiques distinctes que ce soit par les statistiques F ou par les tests d'assignation. La distribution géographique des deux formes moléculaires vient renforcer ce point de vue en mettant en évidence un fort isolement reproductif entre M et S car aucun hybride n'a été trouvé entre elles sur 1521 spécimens identifiés dans plus de 19 localités du Cameroun. De ce fait, il est évident que les formes moléculaires M et S constituent deux unités reproductives bien distinctes. D'après le concept biologique des espèces, cet isolement reproductif constitue l'élément clé à la reconnaissance de telles unités comme étant des espèces distinctes.

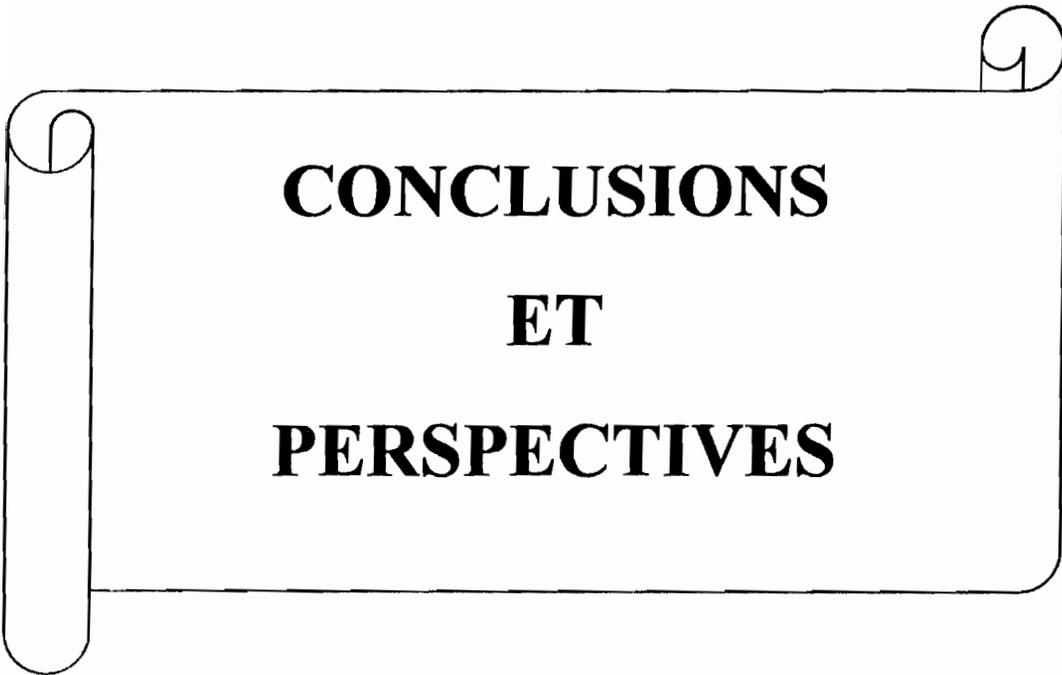
Les résultats que nous avons obtenus sont conformes à de nombreux autres provenant des études sur le même sujet de la part de différents auteurs et dans différents pays. Nous rapporterons les résultats obtenus par Favia *et al.* (1997) en identifiant par PCR-RFLP, la forme moléculaire de 203 spécimens d'*An. gambiae* collectés au Mali dans une zone de sympatrie de trois formes chromosomiques à savoir Savanna, Bamako et Mopti. Aucun hybride M/S n'y avait été trouvé malgré le fait que des individus portant des caryotypes hybrides Mopti/Savanna avaient été introduits dans l'échantillon. Par contre, les femelles hybrides obtenues au laboratoire ont présenté le profil M/S par PCR-RFLP sans ambiguïté tout comme avec la nouvelle technique de PCR directe (Favia *et al.*, 2001).

Un autre élément renforçant l'isolement reproductif entre les formes M et S est la distribution du gène de résistance aux pyréthrinoïdes (*kdr*). En effet, Chandre *et al.* (1999) ont révélé à travers une large étude de la distribution du gène *kdr* en Côte d'Ivoire, que ce gène était trouvé uniquement chez les spécimens de la forme S et jamais chez ceux de la forme M même dans les zones de sympatrie. Parfois la fréquence de ce gène était élevée chez la forme S (96%) mais absente chez M. Des études ont mis en évidence deux différences de nucléotides entre M et S au niveau de l'intron I du gène canal sodium voltage dépendant, en amont de la mutation conférant la résistance aux pyréthrinoïdes (Weill *et al.*, 2000). Dans une autre étude, (Della Torre, Communication personnelle) un fragment de 535pb de cet intron I a été séquencé sur plus d'une centaine de spécimens provenant de 10 pays d'Afrique de l'Ouest et du Centre allant du Sénégal à l'Angola. Les résultats montrent qu'à la position 702, le nucléotide T était toujours retrouvé chez tous les spécimens S alors que chez tous les spécimens de la forme M, un C était présent. A la position 896, un C était fixé chez tous les spécimens S analysés alors que les spécimens M montraient soit un C comme pour les échantillons S ou un polymorphisme A/C, selon l'origine géographique des spécimens. La mutation à la position 702 constitue une autre indication de la différenciation entre M et S et cette fois hors de la région IGS/ITS de l'ADN ribosomal sur le chromosome X où ces formes ont été définies. Cette observation est d'autant plus significative qu'elle a été mise en évidence sur une large échelle géographique. La récente découverte du gène *kdr* chez des spécimens de la forme M au Bénin, Burkina Faso et en Côte d'Ivoire (Chandre, communication personnelle) ne remet pas en cause l'isolement reproductif entre M et S. Le fait que l'apparition du gène *kdr* chez la forme M soit le fruit d'une introgression à partir de la forme S (Weill *et al.*, 2000) ne signifie pas absolument qu'il y a un flux de gènes significatif entre les deux formes. En effet, les phénomènes d'introgression ont été observés entre des espèces bien établies à l'exemple d'*An. gambiae* et *An. arabiensis* (Besansky *et al.*, 1997), entre *An. gambiae* et *An. bwambae* (Thewell *et al.*, 2000). Nos résultats et ceux obtenus dans d'autres études supportent l'idée selon laquelle après que l'introgression ait eu lieu entre M et S dans une zone de sympatrie de leur aire de distribution, la diffusion du gène *kdr* s'est ensuite effectuée au sein de la forme moléculaire M. Ceci justifierait pourquoi le gène *kdr* est rencontré actuellement au Bénin et en Côte d'Ivoire au sein de la forme M dans les zones du littoral où elle est seule présente et par contre dans les zones de savanne où les deux formes sont sympatriques, seule la forme S possède le gène *kdr*.

Toutefois, l'existence d'hybrides sur le terrain a été mentionnée dans certaines études. Della Torre *et al.* (2001) ont trouvé 3 hybrides sur un total de 1161 spécimens adultes d'*An.*

gambiae testés à travers l'Afrique soit une fréquence d'hybrides de 0,26%. Il faut signaler que deux de ces hybrides ont été trouvés dans une localité où seule la forme S est présente et que par conséquent les auteurs ont pensé qu'il pourrait s'agir de contaminations durant la PCR. Taylor *et al.* (2001) ont identifié des hybrides M/S au stade larvaire à une proportion comprise entre 0 et 1,29% dans la localité de Banambani au Mali. Près de cette même localité, Tripet *et al.* (2001) ont trouvé une femelle hybride M/S inséminée démontrant ainsi que les femelles hybrides M/S peuvent être produites sur le terrain, survivre jusqu'au stade adulte. La détermination du génotype de cette femelle grâce à 4 locus microsatellites a permis aux auteurs de rejeter toute possibilité de contamination lors de l'identification de son statut moléculaire et la proportion totale d'hybrides dans leur étude fut de 1/329 soit 0,3%. Néanmoins, il faut signaler qu'entre espèces jumelles bien établies comme *An. gambiae* et *An. arabiensis*, il existe parfois des hybrides dans la nature dans des zones de sympatrie. La proportion de ces hybrides est toujours faible (0,1-0,2‰) entre *An. gambiae* et *An. arabiensis* selon Touré *et al.* (1998).

Les valeurs de F_{st} que nous avons observées entre les populations des formes M et S dans notre étude comme dans d'autres menées sur la structure génétique des populations d'*An. gambiae* (Lanzaro *et al.*, 1998) sont modérées. Ceci peut être dû plus au fait que les deux formes M et S conserveraient encore un polymorphisme ancestral qu'à un échange de gènes contemporains entre elles. Ce polymorphisme ancestral serait la proportion de variabilité que les deux formes moléculaires auraient conservée de leur ancêtre commun et qui malgré leur séparation en deux pools géniques distincts, leur conférerait encore un degré de similitude indépendamment du profil actuel de flux de gènes. Il semble évident que ces deux formes moléculaires M et S ont subi de récents changements dans leur démographie et leur histoire et ceci principalement à cause des changements anthropiques au niveau de l'environnement (Coluzzi, 1999 ; Powell *et al.*, 1999). Ces changements sont probablement à l'origine de cette spéciation chez *An. gambiae* et qui à notre avis, a conduit à la formation de deux espèces jumelles que sont la forme moléculaire M et la forme S. Cette spéciation est récente et probablement n'est pas encore achevée à tous les points du génome, ce qui peut expliquer les valeurs de F_{st} encore modérées et d'autres similitudes observées entre M et S notamment au niveau de plusieurs régions de l'ADN nucléaire (Mukabayire *et al.*, 2001).



**CONCLUSIONS
ET
PERSPECTIVES**

Ce travail dont nous venons de présenter les résultats avait pour but i) de caractériser les populations d'*An. gambiae* au Cameroun selon les formes moléculaires M et S et les différentes formes chromosomiques et ensuite ii) d'étudier la structure génétique de cet important vecteur du paludisme au Cameroun.

L'utilisation des locus microsatellites a été très utile pour mettre en évidence l'hétérogénéité génétique entre les formes M et S d'*An. gambiae*.

Au terme de ce travail, plusieurs conclusions et perspectives peuvent être proposées.

1- Apport des locus d'ADN microsatellites dans la détermination de la structure génétique des populations du complexe *An. gambiae*

Au cours de notre travail nous avons utilisé les locus microsatellites pour étudier la variabilité génétique des populations d'*An. gambiae* responsable de la majeure partie de la transmission du paludisme en Afrique. 10 locus ont été utilisés pour cette étude.

La procédure de sélection des locus avant le début de chacune de ces études a permis de choisir des locus possédant des caractéristiques appréciables pour déterminer la structure génétique des populations. En général, presque tous les locus ont présenté un niveau élevé de polymorphisme ce qui leur confère l'habileté à révéler les faibles structurations au sein des populations étudiées.

En fait, les microsatellites sont des marqueurs génétiques présentant plusieurs avantages. Parmi ces avantages, nous pouvons citer leur taux de mutation élevé qui leur confère un polymorphisme important contrairement à d'autres marqueurs comme les isoenzymes (Lanzaro *et al.*, 1995). Grâce à leur fort polymorphisme, les microsatellites conservent moins les traces du polymorphisme ancestral commun à deux espèces. Par conséquent, ils sont reconnus être de bons indicateurs de la différenciation génétique entre les populations (Walton *et al.*, 1998). Les locus microsatellites sont des marqueurs neutres, ce qui leur permet de fournir une information génétique plus fiable, non biaisée par la sélection. Contrairement à d'autres marqueurs génétiques, les locus microsatellites sont caractérisés par leur ubiquité. En effet, ils sont distribués à travers tout le génome. Chez les anophèles qui possèdent $2n=6$ chromosomes, ils sont retrouvés sur tous les chromosomes permettant d'avoir un profil de différenciation génétique entre les sous-populations ou entre les espèces qui reflète tout le génome. Cette ubiquité est importante car elle permet d'éviter les potentiels biais de sélection qui peuvent agir sur une région particulière du génome et non sur les autres et fournir ainsi un profil de différenciation qui serait loin de la réalité. Chez *An. gambiae*, les

études ont justement mis en évidence une action de la sélection sur les locus microsatellites situés près ou à l'intérieur des inversions chromosomiques du bras droit du chromosome 2 (Lanzaro *et al.*, 1998). Cette influence de la sélection amplifiait le niveau de différenciation génétique observé dans cette étude et induisait ainsi une conclusion erronée sur la structure génétique de la dite population. Dans notre travail, nous avons en général sélectionné des locus situés hors des inversions et pour ceux qui y étaient proches, nous n'avons pas observé de différences avec les autres que ce soit pour l'équilibre de Hardy-Weinberg (valeurs de Fis) que pour le niveau de différenciation génétique (valeurs de Fst). Par conséquent dans toute notre étude les profils de différenciation génétique observés ont toujours été un reflet de tout le génome.

Malgré ces nombreux avantages, il a été démontré que de nombreuses contraintes agissaient sur les locus microsatellites et étaient de nature à affecter leur potentiel résolutif. Au nombre de ces contraintes, nous mentionnons la présence d'allèles nuls au niveau de certains locus. La présence d'allèles nuls au niveau des locus microsatellites dans des études chez *An. gambiae* et *An. arabiensis* est bien documentée dans la littérature (Lehmann *et al.*, 1996 ; Kamau *et al.*, 1998a et b ; Walton *et al.*, 1998 ; Wang *et al.*, 1999, 2001). Dans notre étude, nous avons observé que trois locus présentaient très souvent un déficit en hétérozygotes (H170, H93 et H750). Ces trois locus sont situés sur le chromosome 3 à l'extérieur de toute inversion chez *An. gambiae*. L'hypothèse d'un effet Wahlund est aussi rejetée tout comme celle d'une consanguinité (apparemment) des individus car l'équilibre de Hardy-Weinberg est généralement respecté au niveau des autres locus. Les tests de déséquilibre de liaison n'indiquent pas que ces locus soient particulièrement liés à d'autres. Par conséquent la présence d'allèles nuls au niveau de ces locus est la cause probable de ces déficits observés en hétérozygotes. Toutefois, ces déficits n'ont pas influencé de façon significative les niveaux de différenciation génétique observés entre les populations de ces deux espèces vectorielles.

Malgré l'hétérogénéité génétique et l'isolement reproductif observés entre les formes moléculaires M et S chez *An. gambiae*, les valeurs de Fst ont été modérées. Cette situation peut être la conséquence des contraintes de taille agissant sur les locus microsatellites ou un taux de mutation plus faible que ce qui est supposé pour ce marqueur génétique. En effet, les phénomènes d'homoplasie au niveau des locus microsatellites entraînent une surestimation de l'homologie entre les populations testées si les allèles identiques par état (allèles de même taille dans notre étude) ne le sont pas par descendance (c'est-à-dire issus de processus évolutifs différents). Ces phénomènes d'homoplasie au niveau de certains locus microsatellites ont été mis en évidence chez *An. gambiae* par Walton *et al.* (1998) notamment

pour le locus H147. Ces auteurs ont mis en évidence plusieurs substitutions de bases et des insertions/délétions dans les régions flanquantes de ce locus entre les populations de différentes formes chromosomiques (Forest et Mopti). Toutefois il a été montré que ces phénomènes d'homoplasie étaient plus courants au niveau interspécifique qu'à l'intérieur des complexes d'espèces (Estoup *et al.*, 1995a). Ceci nous amène à relativiser cette influence sur les locus microsatellites dans notre étude.

Tout en tenant compte des nombreuses incertitudes qui demeurent sur les taux et processus de mutation des séquences microsatellites, du fait que plusieurs hypothèses sur lesquelles les méthodes d'analyse des données sont basées ne sont pas respectées chez les populations naturelles d'anophèles, nous pouvons affirmer que les microsatellites ont été des marqueurs génétiques très utiles pour répondre aux questions que nous nous sommes posées dans cette étude.

2- Spéciation chez *An. gambiae*

L'ensemble des résultats que nous avons obtenus confirme la grande complexité génétique d'*An. gambiae*. Il ressort clairement de notre étude que cette espèce est constituée de deux unités génétiques bien différentes (les formes moléculaires M et S).

A l'aide des locus microsatellites, nous avons montré que les populations d'*An. gambiae* étaient structurées en fonction des deux formes moléculaires M et S. Au sein de chaque forme moléculaire, les populations sont génétiquement homogènes malgré de grandes distances qui les séparent. Cette homogénéité génétique implique un niveau élevé de flux de gènes entre ces populations ceci indépendamment des formes chromosomiques auxquelles elles appartiennent. La distance géographique entre ces populations joue un rôle négligeable dans le profil de différenciation génétique puisque nous n'avons trouvé aucune corrélation entre ces deux paramètres. Néanmoins, les populations des différentes localités étudiées ne sont pas panmictiques lorsqu'elles sont groupées ce qui suppose tout de même une action de la dérive génétique sur ces populations au niveau local. Cette homogénéité génétique entre les populations de la même forme moléculaire, a été confirmée par d'autres méthodes notamment par les tests d'assignation qui ont l'avantage de s'intéresser aux individus et non aux populations. Il en est ressorti que les individus d'une forme moléculaire sont assignés majoritairement dans une population de la même forme (individus S assignés dans une population S individus M assignés dans les populations de forme M

Par contre, nous avons observé une forte hétérogénéité génétique entre les populations de forme M et S. Avec les locus microsatellites, cette hétérogénéité s'est traduite par un

niveau de flux génique très faible entre les populations de ces deux formes moléculaires avec des valeurs élevées et significatives de *F_{st}*. L'hétérogénéité génétique entre les formes M et S a été observée et même de façon plus prononcée lorsqu'elles étaient sympatriques notamment à Simbock. Cette situation indique l'existence d'un phénomène de renforcement entre les populations de ces deux formes moléculaires d'*An. gambiae*. L'hétérogénéité génétique entre les formes M et S est également confirmée par les tests d'assignation.

La différenciation génétique observée entre M et S a été confortée par les résultats de l'étude de la distribution géographique de ces formes au Cameroun. Nous avons montré que ces deux formes M et S étaient reproductivement isolées puisque sur 1521 spécimens, aucun hybride n'a été trouvé et l'une des formes est toujours prédominante sur l'autre sinon même seule présente dans les localités étudiées. La distribution de ces deux formes semble influencée par des facteurs éco-climatiques et par l'exclusion compétitive. Néanmoins, il serait intéressant de mener des études bio-écologiques pour mieux comprendre les facteurs qui régissent la distribution des formes M et S. Nos résultats sont corroborés par ceux de Della Torrè *et al.* (2001) qui dans une étude dans plusieurs pays africains sont arrivés aux mêmes conclusions que nous. Nous avons confirmé l'extraordinaire complexité génétique d'*An. gambiae* en montrant l'existence de trois formes chromosomiques à savoir Forest, Savanna et Mopti au Cameroun. Cette complexité est renforcée par le fait qu'en dehors de la forme chromosomique Mopti, les deux autres sont constituées d'individus appartenant aux deux formes moléculaires M et S.

La somme des résultats et observations obtenues sur la caractérisation moléculaire, chromosomique et sur la structure génétique d'*An. gambiae* au Cameroun, nous a amené à nous poser la question du statut taxonomique des formes moléculaires M et S. Nous avons montré par plusieurs méthodes (ADN microsatellites, tests d'assignation) qu'elles étaient génétiquement hétérogènes et qu'elles étaient reproductivement isolées l'une de l'autre dans la nature. Fort de toutes ces observations et celles faites par d'autres auteurs (Della Torrè *et al.*, 2001, 2002), nous suggérons que ces deux formes moléculaires d'*An. gambiae* seraient deux espèces vectrices ayant divergé récemment à la suite d'un processus de spéciation qui pourrait ne pas encore être terminé. Les similitudes encore observées entre ces deux formes tant au niveau de leurs paramètres entomologiques, qu'au niveau génomique suggèrent que leur séparation est récente.

Il faut signaler que la reconnaissance de nouvelle espèce au sein du complexe *An. gambiae*, n'a pas uniquement un intérêt fondamental. Les membres de ce complexe sont les insectes les plus importants d'un point de vue médical. Environ 2,7 millions de personnes

meurent du paludisme chaque année et parmi ces personnes une majorité a été infectée par les anophèles membres du complexe *An. gambiae* tels *An. gambiae* et *An. arabiensis*. Les programmes de contrôle des vecteurs doivent prendre en compte le nombre d'unités taxonomiques ou génétiques qui sont ciblées par le contrôle, surtout en considérant le fait que ces unités pourraient être écologiquement distinctes.

3- Apport des études de la structure génétique des populations du complexe *An. gambiae* sur les mesures de contrôles anti-vectoriels

Les études de la structure génétique des populations d'anophèles fournissent des informations sur le niveau d'échange des gènes entre populations. Ces informations sont utiles pour déduire les profils de dispersion des anophèles, paramètre important dans le contrôle des vecteurs. Si le niveau d'échange génique observé entre populations de différentes localités géographiques est significativement élevé, les gènes d'intérêt, s'ils sont présents dans une des populations migreront dans les autres avec le temps. Si ces gènes d'intérêt sont par exemple les gènes responsables de la résistance aux insecticides ou des gènes de la compétence vectorielle, il est possible de prédire non seulement jusqu'où migreront ces gènes, mais aussi avec quelle vitesse ils le feront.

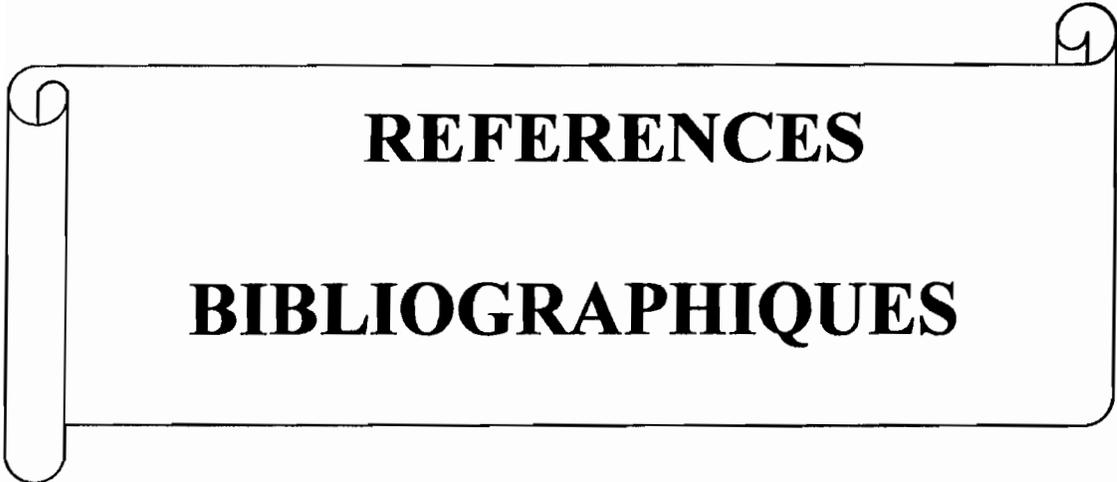
Notre étude indique que les populations d'*An. gambiae* sont génétiquement structurées notamment en fonction des deux formes moléculaires M et S. Chez ce vecteur, si les gènes de résistance sont présents au sein d'une population d'une forme moléculaire précise, ils ne pourront migrer qu'au sein des populations de la même forme moléculaire. En effet, nous avons montré qu'il existe une barrière de flux de gènes entre les formes M et S. Ce fait a été confirmé par les études de Chandre *et al.* (1999b) et Della Torre *et al.* (2001) qui ont trouvé des fréquences différentes du gène *kdr* chez les populations de forme M et S. Dans la plupart des cas, quand le gène de résistance *kdr* était trouvé au sein de la population S, il était absent au sein de la population M de la même localité. Ainsi, les populations de forme M sont en général plus sensibles aux pyréthrinoïdes que celles de forme S. Cette information est à prendre en compte dans la mise sur pied de tout programme de contrôle de ce vecteur. C'est dans cette perspective que notre étude sur la distribution géographique des formes moléculaires M et S s'avère importante et utile. Celle-ci gagnerait à être étendue à toute l'aire de répartition de ce vecteur au Cameroun. La connaissance de la structure génétique des populations vectorielles est par conséquent importante dans le contrôle effectif de ces populations à travers une utilisation rationnelle des insecticides. Les profils de dispersion des

espèces que nous avons obtenus peuvent aussi fournir des indications sur l'habilité de ces vecteurs à propager les souches de *Plasmodium* chimiorésistantes.

Il est de plus en plus fait mention de l'utilisation des moustiques génétiquement modifiés pour lutter contre le paludisme. Cette méthode de lutte ne vise pas l'élimination ou la réduction des populations naturelles de vecteurs, mais plutôt une modification de l'aptitude de ces populations à permettre le développement des *Plasmodiums*. L'idée de base est d'introduire dans les populations naturelles de vecteurs, des gènes qui les rendront réfractaires aux infections de *Plasmodiums*. Les recherches sont menées pour trouver les meilleurs moyens pour transformer les anophèles et véhiculer ces gènes dans les populations naturelles. Actuellement les études sur ce sujet ont déjà permis de transformer *An. stephensi* et *An. gambiae* en utilisant des éléments transposables minos et *piggyBac* respectivement (Caterrucia *et al.*, 2000 ; Grossman *et al.*, 2001). L'on a aussi observé chez la drosophile, que les symbiotes du genre *Wolbachia* pourraient envahir très rapidement et uniformément des populations naturelles réparties sur une grande zone géographique (Turelli & Hoffman, 1991 ; Turelli *et al.*, 1992). Le succès d'un programme de contrôle des vecteurs par utilisation des moustiques génétiquement modifiés nécessite une grande connaissance de la structure génétique des populations de vecteurs ainsi que des barrières de flux de gènes. Cette connaissance est indispensable pour déterminer l'échelle temporelle et géographique nécessaire pour l'introduction de tels gènes. Nous pensons que dans cette perspective, le travail que nous avons mené est d'avant garde et pourrait permettre plus aisément la mise sur pied au Cameroun d'un programme de contrôle des vecteurs par ces moustiques génétiquement modifiés si au préalable, les considérations d'ordre éthique étaient minutieusement étudiées.

Le séquençage du génome entier d'*An. gambiae* très récemment (Holt *et al.*, 2002) ouvre d'énormes perspectives dans la recherche sur ce vecteur majeur du paludisme en Afrique. Pour ce qui nous concerne, il s'agira d'étudier les bases moléculaires de l'isolement reproductif que nous avons observé entre les formes moléculaires M et S de cette espèce. Nous comptons également étudier de manière intensive les niveaux de flux géniques entre ces formes avec de nouveaux marqueurs qui découleront du séquençage (ADN microsatellites, gènes discriminants particuliers) mais aussi à l'aide du séquençage des gènes de l'ADN mitochondrial qui est reconnu évoluer plus rapidement que l'ADN nucléaire. Nous envisageons fortement d'étudier également les réponses immunitaires d'*An. gambiae* face à l'infection de *Plasmodium*. Nous pensons utiliser pour cela les techniques de puces d'ADN («microarray») qui permettent d'avoir accès à l'expression des gènes du moustique dans des

situations spécifiques et ainsi de déterminer quels sont les gènes impliqués dans le phénomène étudié. Il serait aussi important d'évaluer l'impact des moustiquaires imprégnées sur la structure génétique des populations de vecteurs étant donné leur utilisation croissante comme moyens de prévention contre le paludisme par les populations.

A decorative border resembling a scroll, with a vertical strip on the left side and rounded ends on the top and right. The text is centered within this scroll.

REFERENCES

BIBLIOGRAPHIQUES

Références Bibliographiques

- Antonio Nkondjio C., Awono-Ambene P., Toto J.C., Nyamban R., Kemleu-Zébazé S. Meunier J.Y., Wondji C., Tchuinkam T & Fontenille D (2002). High malaria transmission in a village close to Yaoundé, the capital city of Cameroon. *J. Med. Ent.* **39** (2) : 350-355).
- Ayala F.J. (ed.). (1976). Molecular evolution. Sinauer Associates, Sunderland, Mass.
- Beier J.C., Perkins P.V., Wirtz R.A., Koros J., Diggs D., Gargan T.P.II and Koech DK. (1988). Bloodmeal identification by direct enzyme-linked immunosorbent assay (ELISA) tested on *Anopheles* (Diptera: Culicidae) in western Kenya. *J. Med. Ent.* **33**: 613-618.
- Bell C.J. & Ecker J.R. (1994). Assignment of 30 microsatellite locus to the linkage map of *Arabidopsis*. *Genomics* **19**: 137-144.
- Besansky N.J., Lehmann T., Fahey G.T., Fontenille D., Braack L.E.O., Hawley W.A. & Collins F.H. (1997). Patterns of mitochondrial variation within and between African malaria vectors, *Anopheles gambiae* and *Anopheles arabiensis*, suggest extensive gene flow. *Genetics* **147**: 1817-1828.
- Black W.C., Rai I.V. Turco B.J. & Arroyo D.C. (1989). Laboratory study of competitive between United States strains of *Aedes albopictus* and *Aedes aegypti* (Diptera : Culidae). *J. of Med. Entomol.* **26** : 260-271.
- Black W.C. & Lanzaro G.C. (2001). Distribution of genetic variation among chromosomal forms of *Anopheles gambiae* s.s: introgressive hybridization, adaptive inversions, or recent reproductive isolation? *Insect Mol Biol* **10** : 3-7.
- Boëte C. & Koella J.C. (2003). Evolutionary ideas about genetically manipulated mosquitoes and malaria control. *Trends in Parasitology* **19**(1) : 32-38.
- Bossart J.L. & Pashley Prowell D. (1998). Genetic estimates of population structure and gene flow: limitations, lessons and new directions. *T.R.E.E.* **13**(5): 202-206.
- Bowcock A.M., Ruiz Linares A., Tomfohrde J., Minch E., Kidd J.R. & Cavalli-Sforza L.L. (1994). High resolution of human evolutionary trees with polymorphic microsatellites. *Nature* **368**: 455-457.
- Brahmachari S.K., Meera G., Sarkar P.S., Balagurumoorthy P., Tripathi J., Raghavan S., Shaligram U. & Pataskar S. (1995). Simple repetitive sequences in the genome: Structure and functional significance. *Electrophoresis* **16**(9): 1705-1714.
- Brinkmann B., Klintschar M., Neuhuber F., Hühne J. & Rolf B. (1998). Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am. J. Hum. Genet.* **62**: 1408-1415.
- Brookfield J.F.Y. (1996). A simple new method for estimating null allele frequency from heterozygote deficiency. *Mol. Ecol.* **5**: 453-455.
- Bullini L. & Coluzzi M. (1978). Applied and theoretical significance of electrophoretic studies in mosquitoes (Diptera: Culicidae). *Parassitologia* **20** : 7-21.

- Bullini L. (1984). Enzyme variants in the identification of parasites and vectors: methodological aspects of the electrophoretic approach. *In: New Approach to the Identification of Parasites and their Vectors*. Eds Newton BN et Michal F. *World Health Organ. Tropical Disease Research, Geneva. Series 5*: 53-69.
- Callen D.F., Thompson A.D., Shen Y., Phillips H.A., Richards R.I., Mulley J.C. & Sutherland G.R. (1993). Incidence and origin of "null" alleles in the (AC)_n microsatellite markers. *Am. J. Hum. Genet.* **52**: 922-927.
- Carnevale P., Robert V., Boudin C., Halna J.M., Pazart L., Richard A. & Mouchet J. (1988). La lutte contre le paludisme par les moustiquaires imprégnées de pyréthrinoïdes au Burkina Faso. *Bull. Soc. Path. Ex.* **81** : 832-846.
- Carson H.L., Clayton F.E. & Stalker H.D. (1967). Karyotypic stability and speciation in Hawaiian *Drosophila*. *Proc. Nat. Acad. Sci. USA* **57** : 1280-1285.
- Catterucia F., Nolan T., Loukeris T.G., Blass C., Savakis C., Kafatos F.C. & Crisanti A. (2000). Stable germline transformation of the malaria mosquito *Anopheles stephensi*. *Nature* **405** : 959-962.
- Chakraborty R., De Andrade M., Daiger S.P. & Budowle B. (1992). Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Ann Hum Genet* **56**: 45-57.
- Chakraborty R. & Jin L. (1993). An unified approach to study hypervariable polymorphisms : statistical considerations of determining relatedness and populations distances, pp. 153-175 in *DNA fingerprinting : State of the Science*, edited by S.D.J. Pena, R. Chakraborty, J.T. Eppelen and A.J. Jeffreys. Birkhauser verlag Basel, Switzerland.
- Chakraborty R. & Zhong Y. (1994). Statistical power of an exact test of Hardy-Weinberg proportions of genotypic data at a multiallelic locus. *Hum. Hered.* **44**: 1-9.
- Chandre F., Manguin S., Brengues C., Dossou Yovo J., Darriet F., Diabaté A., Carnevale P. & Guillet P. (1999a). Current distribution of pyrethroid resistance gene (kdr) in *Anopheles gambiae* complex from West Africa and further evidence for reproductive isolation of the Mopti form. *Parassitologia*, **41**, 319 – 322.
- Chandre F., Darriet F., Manguin S., Brengues C., Carnevale P. & Gillet P. (1999b). Pyrethroid cross resistance spectrum among populations of *Anopheles gambiae* from Côte d'Ivoire. *J. Mosq. Control Assoc.* **15**: 15: 53-59.
- Chung M., Ranum L.P.W., Duvick L.A., Servadio A., Zoghbi H.Y. & Orr H.T. (1993). Evidence for a mechanism predisposing to intergenerational CAG repeat instability in spinocerebellar ataxia type I. *Nat. Genet.* **5**: 254-258.
- Cianchi R., Villani F., Touré Y.T., Petrarca V. & Bullini L. (1983). Electrophoresis study of different chromosomal forms within *Anopheles gambiae* s.s. *Parassitologia* **25**: 239-241.
- Collins F.H., Mendez M.A., Rasmussen M.O., Mehaffey P.C., Besansky N.J. & Finnerty V. (1987). A ribosomal RNA gene probe differentiates member species of the *Anopheles gambiae* complex. *Am. J. Trop. Med. Hyg.* **37**(1): 37-41.
- Collins F.H., Petrarca V., Mpofo S., Brandling-Bennett A.D., Were J.B.O., Rasmussen M.O., Finnerty V. (1988). Comparison of DNA probe and cytogenetic methods for identifying

- field collected *Anopheles gambiae* complex mosquitoes. *Am. J. Trop. Med. Hyg.* **39** : 545-550.
- Collins F.H. et James A.A. (1996). Genetic modification of mosquitoes. *Science & Medicine* (Nov-Dec): 52-61.
- Colson I. & Golstein D.B. (1999). Evidence for complex mutations at microsatellite locus in *Drosophila*. *Genetic* **152** : 617-27
- Coluzzi M. (1966). Osservazione comparative sul cromosoma X delle specie A e B del complesso *Anopheles gambiae*. *Accad. Naz. Dei Lincei, fasc. 4, serie 8* : 40
- Coluzzi M. & Sabatini A. (1967). Cytogenetic observations on sp. A and B of the *Anopheles gambiae* complex. *Parassitologia* **9**: 73-88.
- Coluzzi M. & Sabatini A. (1968a). Cytogenetic observations on species C of the *Anopheles gambiae* complex. *Parassitologia* **10**: 155-165.
- Coluzzi M. & Sabatini A. (1969). Cytogenetic observations on the salt-water species *Anopheles merus* and *An. melas*, of the *gambiae* complex. *Parassitologia* **11**: 177-187.
- Coluzzi M. (1970). Sibling species in *Anopheles* and their importance in malariology. *Misc. Publ. Ent. Soc. Am.* **7** : 63-77.
- Coluzzi M., Sabatini A., Petrarca V. & Di Deco M.A. (1979). Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* **73**: 483-497.
- Coluzzi M. (1982). Spatial distribution of chromosomal inversions and speciation in Anopheline mosquitoes. In: *Mechanisms of Speciation*. Ed Barigozzi C. Liss, New York: 143-153.
- Coluzzi M., Petrarca V. & Di Deco M.A. (1985). Chromosomal inversion intergradation and incipient speciation in *Anopheles gambiae*. *Boll. Zool.* **52**: 45-63.
- Cornel, A. & Collins, F.H. (1996) PCR of the Ribosomal DNA Intergenic Spacer Regions as a method for identifying mosquitoes in the *Anopheles gambiae* complex. *Meth in Mol Biol* **50**: 321-332.
- Cornuet J.M., Piry S., Luikart G., Estoup A. & Solignac M. (1999). New methods employing multilocus to select or exclude populations as origins of individuals. *Genetics* **153** : 1989-2000.
- Costantini C., Li S.G., Della-Torre A., Sagnon N'F., Coluzzi M. & Taylor C.E. (1996). Density, survival and dispersal of *Anopheles gambiae* complex mosquitoes in a west African Sudan savanna village. *Med. Vet. Ent.* **10**: 203-219.
- Coyne J.A. & Orr H.A. (1989). Patterns of speciation In *Drosophila*. *Evolution* **43**, 362-381.
- Coyne JA & Orr H.A. (1997). Patterns of speciation In *Drosophila* revisited. *Evolution* **51** : 295-303.
- Crampton J.M., Morris A.C., Lycett G.J., Warren A. & Eggleston P. (1990). Transgenic mosquitoes: a future vector control strategy? *Parasitology Today* **6**: 31-36.
- Crow J.F. & Kimura M. (1970) An introduction to population genetics theory. Harper & Row, New York.

- Danis M. & Mouchet J. (1991). Paludisme. Paris, Universités francophones UREF. 240pp.
- Darriet F., Robert V., Tho Vien N. and Carnevale P. (1984). Evaluation of the efficacy of permethrin-impregnated intact and perforated mosquito nets against vectors of malaria. WHO mimeographed document, WHO/VBC/84. 899, 19pp.
- Darriet F. (1991). Evaluation sur le terrain de l'efficacité de trois pyréthrinoïdes dans le cadre de la lutte contre les vecteurs du paludisme. *Parassitologia* **33** : 111-119.
- Davidson G. (1962). *Anopheles gambiae* complex. *Nature* **196**: 907.
- Davidson G. & Jakson C.E. (1962). Incipient speciation in *Anopheles gambiae* Giles. *Bull. Wld. Hlth. Org.* **27**: 303-305.
- Davidson G. (1964a). The five mating types in the *Anopheles gambiae* complex. *Riv. Malar.* **13**: 167-183.
- Davidson G. & White G.B. (1972). The crossing characteristics of a new, sixth species in the *Anopheles gambiae* complex. *Trans. R. Soc. Trop. Med. Hyg.* **66**: 531-532.
- Davidson G. & Hunt R. (1973). The crossing and chromosome characteristics of a new sixth species in the *Anopheles gambiae* complex, *Parassitologia*, **15** : 121-128.
- Deka R., Shriver M.D., Yu L.M., Aston C.E., Chakraborty R. & Ferrell R.E. (1994). Conservation of human chromosome 13 polymorphic human (CA)_n repeats in chimpanzees. *Genomics* **22**: 226-230.
- Della Torre, A., Fanello C., Akogbeto M., Dossou-yovo J., Favia G., Petrarca V. & Coluzzi M. (2001). Molecular evidence of incipient speciation within *Anopheles gambiae* s.s. in West Africa. *Insect Mol Biol* **10** : 9-18.
- Della Torre A., Costantini C., Besansky N.J., Caccone A., Petrarca V., Powell J.R. & Coluzzi M. (2002). Speciation within *Anopheles gambiae* : the glass is half full. *Science* **298** : 115-117.
- Di Deco M.A., Petrarca V., Villani F. & Coluzzi M. (1980). Polimorfismo cromosomico da inversioni paracentriche ed eccesso degli eterocariotipi in ceppi di *Anopheles* allevati in laboratorio. *Parassitologia* **22**: 304-306.
- Di Deco M.A., Sabatinelli G., Camiz S. & Touré Y.T. (1983). Studio biometrico di due nuove entità del complesso *Anopheles gambiae*. *Parassitologia* **25**: 260-267.
- Di Rienzo A., Peterson A.C., Garza J.C., Valdes A.M., Slatkin M. & Freimer N.B. (1994). Mutational processes of simple-sequences repeat locus in human populations. *Proc. Natl. Acad. Sci. USA* **91**: 3166-3170.
- Dobzhansky T. (1937). Genetics and the origin of species, *Columbia University Press*.
- Dobzhansky T. (1970). Genetics of the evolutionary Process. Columbia Univ. Press, New York.
- Doyle J.J., Morgane M., Tingey S.V. & Powell W. (1998). Size homoplasy in chloroplast microsatellites of wild perennial relatives of Soybean (*Glycine* Subgenus *Glycine*). *Mol. Biol. Evol.* **15**(2): 215-218.
- Edwards A., Hammond H.A., Jin L., Caskey C.T. & Chakraborty R. (1992). Genetic variation at five trimeric and tetrameric tandem repeat locus in four human population groups. *Genomics* **12**: 241-253.

- Ellegren, H. (2000). Microsatellite mutations in the germline: implications for evolutionary inference. *TIG* **16**: 551-557.
- Epperson B.K. (1993). Recent advances in correlation studies of spatial patterns of genetic variation. *In: Evolutionary Biology*. Ed Hecht MK. New-York, Plenum Press.
- Epplen C., Melmer G., Siedlaczek I., Schwaiger F.W., Mäueler W. & Epplen J.T. (1993). On the essence of "meaningless" simple repetitive DNA in eukaryote genomes. *In: DNA Fingerprint: State of the Science*. Eds Pena SDJ, Chakraborty R, Epplen JT and Jeffreys AJ. Birkhauser Verlag, Basel. Pp 29-45.
- Estoup A., Garnery L., Solignac M. & Cornuet J.M. (1995a). Microsatellite variation in honey bee (*Apis mellifera* L.) populations : hierarchical genetic structure and test of the infinite allele and stepwise mutation models. *Genetics* **140** : 679-695.
- Estoup A., Tailliez C., Cornuet J.M. & Solignac M. (1995b). Size homoplasmy and mutational process of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol. Biol. Evol.* **12(6)**: 1074-1084.
- Favia G., Dimopoulos G. & Louis C. (1994). Analysis of the *Anopheles gambiae* genome using RAPD markers. *Insect. Biol. Mol.* **3**: 149-157.
- Favia G., Della Torre A., Bagayoko M., Lanfrancotti A., Sagnon N'F., Touré Y.T. & Coluzzi M. (1997). Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their isolation. *Insect Molecular Biology* **6(4)**: 377-383.
- Favia, G., Lanfrancotti A., Spanos L., Siden-Kiamos I. & Louis C. (2001). Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. *Insect Mol. Biol.* **10** : 19-23.
- Field D. & Wills C. (1998). Long, polymorphic microsatellites in simple organisms. *Proc. of the Roy. Soc. of Lond. B* **263** : 209-215.
- Fisher R.A. & Ford E.B. (1947). The spread of a gene in natural conditions in a colony of the moth *Panaxia dominula*. *Heredity* **1**: 143-174.
- Flechter, D.J.C. & Michener, C.D. (1987) *Kin recognition in animals*. John Wiley & Sons, New York, NY.
- Foltz D.W. (1986). Null alleles as a possible cause of heterozygote deficiencies in the oyster *Crassostera virginica* and other bivalves. *Evolution* **40**: 869-870.
- Fondjo E. (1996). Etude du comportement du complexe *An. gambiae* et de la transmission du paludisme dans deux faciès éco-climatiques au Mali et au Cameroun. Thèse de doctorat, ISFRA, Bamako, Mali. 92 pages.
- Fontenille D., Faye O., Konate L., Sy N. & Collins F.H. (1993). Comparaison des techniques PCR et cytogénétique pour la détermination des membres du complexe *Anopheles gambiae* au Sénégal. *Ann. Parasitol. Hum. Comp.* **68(5/6)**: 239-240.
- Fontenille D, Lochouart L, Diagne N, Sokhna C, Lemasson JJ, Diatta M, Konate L, Faye F, Rogier C & Trape JF. 1997. High annual and seasonal variations in malaria transmission by anophelines and vector species composition in Dielmo, a holoendemic area in Senegal. *Am. J. Trop. Med. Hyg.* **56(3)**: 247-253.

- Fontenille D., Meunier J.Y., Antonio-Nkondjio C. & Tchuinkam T. (2001). Use of Circumsporozoite Protein Enzyme-linked Immunosorbent Assay compared with microscopic examination of salivary glands for calculation of malaria infectivity rates in mosquitoes (Diptera : Culicidae) from Cameroon. *J. Med. Ent.* **38** (3) : 451-454.
- Fréville H., Justy F. & Olivieri I. (2001). Comparative allozyme and microsatellite population structure in a narrow endemic plant species, *Centaurea corymbosa* Pourret (Asteraceae). *Mol. Ecol.* **10** : 879-889.
- FritzSimmons N.N., Moritz C. & Moore S.S. (1995). Conservation and dynamics of microsatellite locus over 300 million years of marine turtle evolution. *Mol. Biol. Evol.* **12**: 432-440.
- Garcia de Leon F.J., Chikhi L. & Bonhomme F. (1997). Microsatellite polymorphism and population subdivision in natural population of European sea bass *Dicentrarchus labrax* (Linnaeus, 1758). *Mol Ecol* **6** : 51-62.
- Garza J.C., Slatkin M. & Freimer N.B. (1995). Microsatellite allele frequencies in Humans and Chimpanzees, with implications for constraints on allele size. *Mol. Biol. Evol.* **12**(4): 594-603.
- Garza J.C. & Freimer N.B. (1996). Homoplasmy for size at microsatellite locus in Humans and Chimpanzees. *Gen. Res.* **6**: 211-217.
- Gentile, G., Slotman M., Ketmaier V., Powell J.R. & Caccone A. (2001). Attempts to molecularly distinguish cryptic taxa in *Anopheles gambiae* s.s. *Insect Mol Biol* **10** : 25-32.
- Gillies M.T. & De Meillon B. (1968). The Anophelinae of Africa south of the Sahara. *Pub. South Afr. Inst. Med. Res.* **54**: 343 pp.
- Gillies M.T. & Coetzee M. (1987). A supplement to the Anophelinae of Africa south of the Sahara. *Pub. South Afr. Inst. Med. Res.* **55**: 143 pp.
- Goldstein D.B., Linares A.R., Cavalli-Sforza L.L. & Feldman M.W. (1995b). Genetic absolute dating based on microsatellites and the origin of modern humans. *Proc. Natl. Acad. Sci. USA* **92**: 6723-6727.
- Gorman M.J., Cornel A.J., Collins F.H. & Paskewitz S.M. (1996). A shared mechanism for melanotic encapsulation of CM-Spehadex beads and a malaria parasite, *Plasmodium cynomolgi* B, in the mosquito, *Anopheles gambiae*. *Exp. Parasitol.* **84**: 380-386.
- Goudet J., Raymond M., De Meeüs T. & Rousset F. (1996). Testing differentiation in diploid populations. *Genetics* **144**: 1933-1940.
- Green C.A. (1972). Cytological maps for the practical identification of females of the three freshwater species of the *Anopheles gambiae* complex. *Ann. Trop. Med. Parasitol.* **66**: 143-147.
- Grimaldi M.C. & Crouau-Roy B. (1997). Microsatellite allelic homoplasmy due to variable flanking sequences. *J. Mol. Evol.* **44**: 336-340.
- Grossman G.L., Rafferty C.S., Clayton J.R., Stevens T.K., Mukabayire O. & Benedict M.Q. (2001). Germline transformation of the malaria vector, *Anopheles gambiae*, with the piggyBac transposable element. *Insect Mol Biol* **10** (6) : 597-604.

- Guo S.W. & Thompson E.A. (1992). Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics* **48**: 361-372.
- Haldane J.B.S. (1954). An exact test for randomness of mating. *J. Genet.* **52**: 631-635.
- Hardy G.H. (1908). Mendelian proportions in a mixed population. *Science* **28**: 49-50.
- Hartl D.L. & Clark A.G. (1989). Principles of population genetics, 2nd Edition. Eds Sinauer Associates, Sunderland, Mass.
- Hervy J.P., Le Goff G., Geoffroy B., Hervé J.P., Manga L. & Brunhes J. (1998). Les anophèles de la région afrotropicale. CD-ROM, ORSTOM-éditions, Paris.
- Higgie, M., Chenoweth, S. & Blows, M.W. (2000). Natural selection and the reinforcement of mate recognition. *Science* **290**: 519-521.
- Holm S. (1979). A simple sequentially rejective multiple test procedure. *Scand. J. of Stat.* **6**: 65-70.
- Holt R.D., Subramanian G.M., Halpern A. Sutton G.G. *et al.* 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298** : 129-149.
- Hunt R.H. (1973). A cytological technique for the study of *Anopheles gambiae* complex. *Parassitologia* **15** : 137-139.
- Hunt R.H., Coetzee M. & Fittene M. (1998). The *Anopheles gambiae* complex: a new species from Ethiopia. *Trans. R. Soc. Trop. Med. Hyg.* **92**: 231-235.
- Jarne P. & Lagoda J.L. 1996. Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.* **11(10)**: 424-429.
- Jeffreys A.J., Wilson V. & Thein S.L. (1985). Hypervariable minisatellite regions in human DNA. *Nature* **314** : 67-73.
- Jeffreys A.J., Tamaki K., MacLeod A., Monckton D.G., Neil D.L. & Armour J.A.L. (1994). Complex gene conversion events in germline mutation at human minisatellites. *Nature Genetics* **6** : 136-145.
- Kamau L, Lehmann T, Hawley WA, Orago ASS et Collins FH. 1998a. Microgeographic genetic differentiation of *Anopheles gambiae* mosquitoes from Asembo bay, western Kenya : a comparison with Kilifi in Coastal Kenya. *Am. J. Trop. Med. Hyg.* **58 (1)** : 64 – 69.
- Kamau L., Hawley W.A., Lehmann T., Orago A.S.S., Cornel A., Ke Z. & Collins F.H. (1998b). Use of short tandem repeats for the analysis of genetic variability in sympatric populations of *Anopheles gambiae* and *Anopheles arabiensis*. *Heredity* **80**: 675-682.
- Kashi Y., King D. & Soller M. (1997). Simple sequence repeats of a source of quantitative genetic variation. *Trends in Genetics* **13** : 74-78.
- Kidwell M.G. & Ribeiro J.M.C. (1992). Can transposable elements be used to drive disease refractoriness genes into vector populations? *Parassitology today* **8** : 325-329.
- Kimura M. (1953). «Stepping stone» model of populations. *Ann. Rep. Natl. Inst. Genet. Japan* **3**: 62-63.

- Kimura M. & Crow J.F. (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**: 725-738.
- Kimura M. & Ohta T. (1978). Stepwise mutation model and distribution of allelic frequencies in a finite population. *Proc. Natl. Acad. Sci. USA* **75(6)**: 2868-2872.
- King D.G. & Soller M. (1998). Variation and fidelity : The evolution of simple sequence repeats as functional elements in adjustable genes. In *Evolutionary Theory and Processes : modern perspectives*. (ed. S.P. Wasser) Oxford University Press, Oxford.
- Kruglyak S., Durett R.T., Schug M.D. & Aquadro C.F. (1998). Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proc. Natl. Acad. Sci. USA* **95(18)**: 10774-10778.
- Kunsler P., Matsuo K. & Schaffner W. (1995). Pathological, physiological, and evolutionary aspects of short unstable DANN repeats in the human genome. *Biol. Chem. Hop. Sey.* **376** : 201-211.
- Lanzaro G.C., Zheng L., Touré Y.T., Traoré S.F., Kafatos F.C. & Vernick K.D. (1995). Microsatellite DNA and isozyme variability in a west African population of *Anopheles gambiae*. *Insect Mol. Biol.* **4(2)**: 105-112.
- Lanzaro G.C., Touré Y.T., Carnahan J., Zheng L., Dolo G., Traoré S., Petrarca V., Vernick K.D. & Taylor C.E. (1998). Complexities in the genetic structure of *Anopheles gambiae* populations in west Africa as revealed by microsatellite DNA analysis. *Proc. Natl. Acad. Sci. USA* **95**: 14260-14265.
- Latter B.D.H. (1973). The island model of population differentiation: a general solution. *Genetics* **73**: 147-157.
- Lehmann T., W.A. Hawley, L. Kamau, D. Fontenille, F. Simard & F.H. Collins. (1996a). Genetic differentiation of *Anopheles gambiae* populations from East and West Africa: comparison of microsatellite and allozyme locus. *Heredity*. **77**: 192-208.
- Lehmann T., Hawley W.A. & Collins F.H. (1996b). An evaluation of evolutionary constraints on microsatellite locus using null alleles. *Genetics* **144**: 1155-1163.
- Lehmann T., Besansky N.J., Hawley W.A., Fahey T.G., Kamau L. & Collins F.H. (1997). Microgeographic structure of *Anopheles gambiae* in western Kenya based on mtDNA and microsatellite. *Mol. Ecol.* **6**: 243-253.
- Lehmann T., Hawley W.A., Grebert H., Danga M., Atieli F. & Collins F.H. (1999). The Rift Valley complex as a barrier to gene flow for *Anopheles gambiae* in Kenya. *Journal of Heredity*. **90**: 613-621.
- Lehmann T., Blackston C.R., Besansky N.J., Escalante A.A, Collins F.H. & Hawley W.A. (2000). The Rift Valley complex as a barrier to gen flow for *Anopheles gambiae* in Kenya : the mtDNA perspective. *Journal of Heredity* **91** : 165-168.
- Lehmann T., Licht M., Maega B.T.A., Chimumbwa JM., Watsenga F., Wondji C., Simard F. & Hawley W.A. (2003). Population genetic structure of *Anopheles gambiae* in Africa. *Journal of Heredity* **94 (2)** : 133-147.
- Levinson G. & Gutman G.A. (1987). Slipped-strand mispairing: a major mechanism for DNA sequence evolution. *Mol. Biol. Evol.* **4**: 203-221.

- Litt M. & Luty J.A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. of Hum. Genet.* **44** : 397-401.
- Mahon R.J., Green C.A. & Hunt R.H. (1976). Diagnostic allozymes for routine identification of adults of the *Anopheles gambiae* complex (Diptera: Culicidae). *Bull. Ent. Res.* **66**: 25-31.
- Maruyama T. & Fuerst P.A. (1984). Population bottleneck and nonequilibrium models in population genetics. I Allele numbers when population evolve from zero variability. *Genetics* **108** : 745-763.
- Maruyama T. & Fuerst P.A. (1985a). Population bottlenecks and nonequilibrium models in population genetics. II. Number of alleles in a small population that was formed by a recent bottleneck. *Genetics* **111**: 675-689.
- Mayr E. (1942) Systematics and the origin of species. Columbia University Press, Cambridge.
- McLain D.K., Collins F.H., Brandling-Bennett A.D. & Were J.B.O. (1989). Microgeographic variation in rDNA intergenic spacers of *Anopheles gambiae* in western Kenya. *Heredity* **62**: 257-264.
- Mekuria Y., Petrarca V. & Tesfamariam T. (1982). Cytogenetic studies on the malaria vector mosquito *Anopheles arabiensis* in the Awash valley, Ethiopia. *Parassitologia* **24**: 238-243.
- Miles S.J. (1978). Enzyme variations in the *Anopheles gambiae* Giles group of species (Diptera, Culicidae). *Bull. Entomol. Res.* **68**: 85-96.
- Milligan P.J.M., Phillips A., Broomfield G. & Molyneux D.H. (1993). A study of the use of gas chromatography of cuticular hydrocarbons for identifying members of the *Anopheles gambiae* (Diptera : Culicidae) complex. *Bull. of Ent. Res.* **83** : 613-624.
- Mnzava A.E.P., Rwegoshora R.T., Wilkes T.J., Tanner M. & Curtis C.F. (1995). *Anopheles arabiensis* and *An. gambiae* chromosomal inversion polymorphism, feeding and resting behaviour in relation to insecticide house-spraying in Tanzania. *Med. Vet. Ent.* **9**: 316-324.
- Mouchet J., Robert V., Carnevale P., Ravaonjanahary Ch., Coosemans M., Fontenille D. & Lochouarn L. (1991). Le défi de la lutte contre le paludisme en Afrique : place et limite de la lutte antivectorielle. *Cahiers Santé* **1** : 277-288.
- Mpofu, S.M., Masendu H.T., Kanyimo K.H. & Mtetwa C.. (1993). Laboratory colonization of *Anopheles quadrimaculatus* from sympatry with other sibling species of the *Anopheles gambiae* complex in Zimbabwe. *Med. Vet. Ent.* **7**: 122-126.
- Nabarro D. (1999). Roll back malaria. *Parassitologia* **41** : 501-504.
- Nauta MJ & Weissing FJ. (1996). Constraints on allele size at microsatellite locus: implications for genetic differentiation. *Genetics* **143**: 1021-1032.
- Nei M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. USA* **70** 3321-3323.
- Nei M., Maruyama T. & Chakraborty R. (1975). The bottleneck effect and genetic variability in populations. *Evolution* **29(1)**: 1-10.

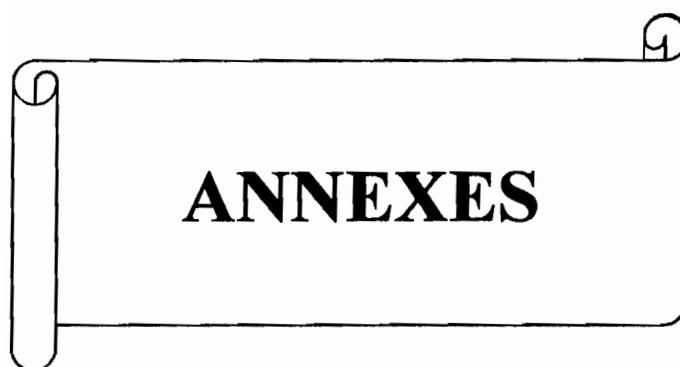
- Nei M. (1987). *Molecular evolutionary genetics*. New-York, Columbia University Press, USA.
- Nordheim A. & Rich A. (1983). The sequence (dC-dA)_n(dG-dT)_n forms left-handed Z-DNA in negatively supercoiled plasmids. *Proc. Natl. Acad. Sci. USA* **80**: 1821-1825.
- OMS. (1998). Evaluation de la santé. *In: Rapport sur la Santé dans le Monde. La vie au 21^e siècle, une perspective pour tous*. World Health Organization (ed), Genève, Suisse. Pp: 43-65.
- Paetkau D. & Strobeck C. (1995). The molecular basis and evolutionary history of a microsatellite null allele in bears. *Mol. Ecol.* **4**: 519-520.
- Paetkau D., Calvert W., Stirling I. & Strobeck C. (1995). Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* **4**: 347-354.
- Pardue M.L., Lowenhaupt K., Rich A. & Nordheim A. (1987). (dC-dA)_n(dG-dT)_n sequences have evolutionary conserved chromosomal locations in drosophila with implications for roles in chromosome structure and function. *EMBO J.* **6**: 1781-1789.
- Paskewitz S.M. & Collins F.H. (1990). Use of the polymerase chain reaction to identify mosquito species of *Anopheles gambiae* complex. *Med. Vet. Entomol.* **4**: 367-373.
- Pemberton J.M., Slate J., Bancroft D.R. & Barrett J.A. (1995). Nonamplifying alleles at microsatellite locus: a caution for parentage and population studies. *Mol. Ecol.* **4**: 249-252.
- Pépin L., Amiques Y., Lepingle A., Berthier J., Bensaid A. & Vaiman D. (1995). Sequence conservation of microsatellites between *Bos taurus* (cattle), *Capra hircus* (goat) and related species. Examples of use in parentage testing and phylogeny analysis. *Heredity* **74**: 53-61.
- Persiani A., Di Deco M.A. & Petrangeli G. (1984). Osservazioni di laboratorio su polimorfismi da inversioni originati da incroci tra popolazioni diverse di *Anopheles gambiae* s.s. *Atti XIII Congr. Soc. It. Parassitologia*. Roma. 3-5 dicembre 1984.
- Petrarca V., Petrangeli G., Rossi P. & Sabatinelli G. (1986). Etude chromosomique d'*Anopheles gambiae* et d'*Anopheles arabiensis* à Ouagadougou (Burkina Faso) et dans quelques villages voisins. *Parassitologia* **28**: 41-61.
- Petrarca V., Vercruyse J. & Coluzzi M. (1987). Observations on the *Anopheles gambiae* complex in the Senegal River Basin, West Africa. *Med. Vet. Ent.* **1**: 303-312.
- Petrarca V. & Beier J.C. (1992). Intraspecific chromosomal polymorphisms in the *Anopheles gambiae* complex as a factor affecting malaria transmission in the Kisumu area of Kenya. *Am. J. Trop. Med. Hyg.* **46**: 229-237.
- Primmer C.R., Moller A.P. & Ellegren H. (1996a). A wide-range survey of cross-species microsatellite amplification in birds. *Mol. Ecol.* **5**: 365-378.
- Queller D.C., Strassmann J.E. & Hughes C.R. (1993). Microsatellites and kinship. *Trends Ecol. Evol.* **8**: 285-288.
- Rannala B. & Hartigan J.A. (1996). Estimating gene flow in island populations. *Genet. Res. Camb.* **67**: 147-158.

- Rannala B. & Mountain J.L. (1997). Detecting immigration by using multilocus genotypes. *Proc. Natl. Acad. Sci. USA* **94** : 9197-9221.
- Raymond M. & Rousset F. (1995a). GENEPOP Version 1.2. A population genetics software for exact tests and ecumenicism. *J. of Heredity* **86**: 248-249.
- Raymond M. & Rousset F. (1995b). An exact test for population differentiation. *Evolution* **49**: 1280-1283.
- Raymond M. (1996). Les marqueurs génétiques en biologie des populations: vers une écologie moléculaire. *Médecine/Sciences* **12(n° spécial)**: 153-157.
- Rhodain F & Perez C. (1985). Précis d'entomologie médicale et vétérinaire. Paris.
- Rice WR. (1989). Analysing tables of statistical tests. *Evolution* **43(1)**: 223-225.
- Robert V. & Carnevale P. (1984). Les vecteurs du paludisme en Afrique sub-saharienne. *Etudes Médicales* **2**: 79-90.
- Robert V., Petrarca V., Carnevale P., Ovazza L. & Coluzzi M. (1989a). Analyse cytogénétique du complexe *Anopheles gambiae* dans la région de Bobo-Dioulasso (Burkina Fasi). *Ann. Parasit. Hum. Comp.* **64**: 290-311.
- Robertson A. & Hill W.G. (1984). Deviations from Hardy-Weinberg proportions : sampling variances and use in estimation of inbreeding coefficients. *Genetics* **107** : 713-718.
- Rogier C., Parzy D. & Spiegel A. (2001). Histoire naturelle et épidémiologie du paludisme. In J.-M. Sauty, ed. *Paludisme grave*. Paris : Arnette : 1-22.
- Rubinsztein D.C., Amos W., Leggo J., Goodburn S., Jain S., Li S.H., Margolis R.L., Ross C.A. & Ferguson-Smith M.A. (1995). Microsatellite evolution - evidence for directionality and variation in rate between species. *Nature Genetics* **10**: 337-343
- Sanguinetti, C.J., Neto, E.D. & Simpson, A.J.G. (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* **17(5)**: 915-919.
- Schlötterer C., Amos B. & Tautz D. (1991). Conservation of polymorphic simple sequence locus in cetacean species. *Nature* **354**: 63-65.
- Schlötterer C., Ritter R., Harr B. & Brem G. (1998). High mutation rate of a long microsatellite allele in *Drosophila melanogaster* provides evidence for allele-specific mutation rates. *Mol. Biol. Evol.* **15(10)**: 1269-1274.
- Schneider S., Kueffer J.M., Roessli D. and Excoffier L. (1997). Arlequin ver. 1.1: A software for population genetic data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Scott J.A., Brogdon W.G. and Collins F.H. (1993). Identification of single specimens of *Anopheles gambiae* complex by the polymerase chain reaction. *Am. J. Trop. Med. Hyg.* **49**: 520-529.
- Service M.W. (1993). Mosquito Ecology: Field Sampling Methods, 2nd edition. Elsevier Applied Science Publishers. London.
- Simard, F., Lehmann, T., Lemasson, J.J., Diatta, M. and Fontenille, D. (2000) Persistence of *Anopheles arabiensis* during the severe dry season conditions in Senegal: an indirect approach using microsatellite locus. *Insect Mol Biol* **9(5)** : 467-479.

- Slatkin M. (1977). Gene flow and genetic drift in a species subject to frequent local extinctions. *Theoretical Population Biology* **12** : 253-262.
- Slatkin M. (1985a). Gene flow in natural populations. *Annual Review of Ecology and Systematics* **16**: 393-430.
- Slatkin M. (1985b). Rare alleles as indicators of gene flow. *Evolution* **39(1)**: 53-65.
- Slatkin M. & Arter H.E. (1991) Spatial autocorrelation methods in population genetics. *American Naturalist* **138** : 499-517.
- Slatkin M. (1994). Linkage disequilibrium in growing & stable populations. *Genetics* **137** : 331-336.
- Slatkin M. (1995). A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**: 457-462.
- Smith G.P. (1976). Evolution of repeated DNA sequences by unequal crossover. *Science* **191**: 528-535.
- Solignac M., Periquet G., Anxolabéhère D. & Petit C. (1995). Génétique et évolution Tome II : L'espèce, l'évolution moléculaire. Collection méthodes Hermann, Editeurs des Sciences et des Arts. 367 pages.
- Stallings R.L. (1995). Conservation and evolution of (CT)_n-(GA)_n microsatellite sequences at orthologous positions in diverse mammalian genomes. *Genomics* **25**: 107-113.
- Sutherlands G.R. & Richards R.I. (1995). Simple tandem DNA repeats and human genetic disease. *Proc. Natl. Acad. Sci. USA* **92**: 3636-3641.
- Tajima F. (1989). The effect of change in population size on DNA polymorphism. *Genetics* **123** : 597-601.
- Tautz D. (1993). Notes on the definition and nomenclature of tandemly repetitive DNA sequences. DNA fingerprinting : *State of the Sciences*. S.D.J. Pena, R. Chakraborty, J.T. Epplen & A.J. Jeffreys. Switzerland, Birkhäuser Verlag Basel : 21-28.
- Tautz D. & Schlötterer C. (1994). Simple sequences. *Curr. Opin. in Genet. and Devel.* **4**: 832-837.
- Taylor C., Touré Y.T., Carnahan J., Norris D.E., Dolo G., Traoré S.F., Edillo F.E. & Lanzaro G.C. (2001). Gene flow among populations of the malaria vector, *Anopheles gambiae*, in Mali, West Africa. *Genetics* **157** : 743-750.
- Thelwell N.J., Hulsman R.A., Harbach R.E. & Butlin R.K.. (2000). Evidence for mitochondrial introgression between *Anopheles bwambae* and *Anopheles gambiae*. *Insect Mol. Biol.* **9**: 203-210.
- Touré Y.T., Petrarca V. & Coluzzi M. (1983). Nuove entità del complesso *Anopheles gambiae* in Mali. *Parassitologia* **25**: 367-370.
- Touré Y.T., Petrarca V., Traoré S.F., Coulibaly A., Maïga H.M., Sow M., Di Deco M.A. & Coluzzi M. (1994). Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae* s.str. in Mali, West Africa. *Genetica* **94**: 213-223.

- Touré Y.T., Dolo G., Petrarca V., Traore S.F., Bouare M., Dao A., Carnahan J. & Taylor C.E. (1998). Mark-release-recapture experiments with *Anopheles gambiae* s.l. in Banambani village, Mali, to determine population size and structure. *Med. Vet. Ent.* **12**: 74-83.
- Touré Y.T. (1999). The malaria challenge in the 21st Century : perspectives for Africa. *Parassitologia* **41** : 507-509.
- Tripet, F., Touré, Y.T., Taylor, C.E., Norris, D.E., Dolo, G. & Lanzaro, G.C. (2001) DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Mol. Ecol.* **10** : 1725-1731.
- Turelli M. & Hoffmann A.A. (1991). Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353**: 440-442.
- Turelli M., Hoffman A.A. & McKechnie S.W. (1992). Dynamics of cytoplasmic incompatibility and mtDNA variation in natural *Drosophila simulans* populations. *Genetics* **132**: 713-723.
- Vernick K.D., Fujioka, Seeley D.C., Tandler B., Aikawa M. & Miller L.H. (1995). *Plasmodium gallinaceum* : a refractory mechanism of ookinete killing in the mosquito, *Anopheles gambiae*. *Exp. Parasitol.* **80** : 583-595.
- Viard F. (1996). Autofécondation et migration en population subdivisées: apport des marqueurs microsatellites chez le gastéropode *Bulinus truncatus*. Thèse. Université de Montpellier II – Sciences et Techniques du Languedoc. France.
- Walton C., Thelwell N.J., Priestman A. & Butlin R.K. (1998). The use of microsatellites to study gene flow in natural populations of *Anopheles* malaria vectors in Africa: potential and pitfalls. *J. Mosq. Control Assoc.* **14**(3): 266-272.
- Wang R., Kafatos F.C. & Zheng L. (1999). Microsatellite markers and genotyping procedures for *Anopheles gambiae*. *Parasitology Today* **15**(1): 33-37.
- Wang R., Zheng L., Touré Y.T., Dandekar T. & Kafatos F.C. (2001). When genetic distances matters : measuring genetic differentiation at microsatellite locus in whole-genome scans of recent and incipient mosquito species. *Proc. of Natl. Acad. of Sci.* **98** : 10769-10774.
- Wasserman M. (1963). Cytology and phylogeny of *Drosophila*. *Am. Nat.* **896** : 333-352.
- Wattier R., Engel C.R., Saumitou-Laprade P. & Valero M. (1998). Short allele dominance as a source of heterozygote deficiency at microsatellite locus : experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Mol. Ecol.* **7** : 1569-1573.
- Weber J.L. & Wong C. (1993). Mutation at human short tandem repeats. *Hum. Mol. Genet.* **2**(8): 1123-1128.
- Weill, M., Chandre, F., Brengues, C., Manguin, S., Akogbeto, M., Pasteur, N., Guillet, P. & Raymond, M. (2000) The *kdr* mutation occurs in the Mopti form of *Anopheles gambiae* s.s. through introgression. *Insect Mol. Biol.* **9**(5): 451-455.
- Weinberg W. (1908). Über den Nachweis der Vererbung beim Menschen. *Jahresh. Verein f. vaterl. Naturk. In Wruttemberg* **64**: 368-382.
- Weir B.S. & Cockerham C.C. (1984). estimating F-statistics for the analysis of population structure. *Evolution* **38**: 1358-1370.

- White G.B. (1985). *Anopheles bwambae* sp.n., a malaria vector in the Semliki Valley, Uganda, and its relationships with other sibling species of the *An. gambiae* complex (Diptera : Culicidae). *System. Ent.* **10** : 501-522.
- Wierdl M., Dominska M. & Petes T.D. 1997. Microsatellite instability in yeast: dependence on the length of the microsatellite. *Genetics* **146**: 769-779.
- Wirtz R.A., Zavala F., Charoenvit Y., Campbell G.H., Burkot T.R., Schneider I., Esser K.M., Beaudoin E.L. & Andre R.C. (1987). Comparative testing of monoclonal antibodies against *Plasmodium falciparum* sporozoites for ELISA development. *Bull. World Health Organ.* **65** : 139-145.
- Wondji C., Simard F. & Fontenille D. (2002). Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. *Insect Mol. Biol.* **11** (1) : 11-19.
- Wright S. (1931). Evolution in Medelian populations. *Genetics* **16**: 97-159.
- Wright S. (1943). Isolation by distance. *Genetics* **28**: 114-138.
- Wright S. (1951). The genetical structure of populations. *Ann. Eugen.* **15**: 323-354.
- Wright S. (1978). Evolution and the genetics of populations. Vol. 4, Variability in and among natural populations. Chicago, University Chicago Press.
- Zheng L., Benedict M.Q., Cornel A.J., Collins F.H. & Kafatos F.C. (1996). An integrated genetic map of the african human malaria vector mosquito, *Anopheles gambiae*. *Genetics* **143**: 941-952.
- Zheng L., Cornel A.J., Wang R., Erfle H., Voss H., Ansorge W., Kafatos F.C. & Collins F.H. (1997). Quantitative trait locus for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. *Science* **276**: 425-428.



ANNEXE 1

Charles Wondji, Frédéric Simard & Didier Fontenille (2002). Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. *Insect Molecular Biology* 11(1), 11-19.

Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry

C. Wondji¹, F. Simard^{1,2} and D. Fontenille^{1,2}

¹Laboratoire de l'Institut de Recherche pour le Développement (IRD), Organisation de Coordination pour la Lutte Contre les Endémies en Afrique Centrale (OCEAC), Yaoundé, Cameroun; and ²Laboratoire de Lutte contre les Insectes Nuisibles (LIN), Institut de Recherche pour le Développement (IRD), Montpellier, France

Abstract

We studied genetic variation at ten microsatellite DNA loci in *Anopheles gambiae* populations from the Forest chromosomal form collected in four villages in Cameroon (Central Africa). Both recently described M and S molecular forms occur in sympatry in this area. Geographic differentiation within form was low ($F_{st} < 0.017$) despite geographical distance between collection sites ranging from 35 to 350 km. However, higher ($F_{st} > 0.035$) and statistically significant levels of genetic differentiation were observed between forms, being the highest between sympatric M and S populations collected within the same village. Results were consistent across all loci spread throughout the genome, therefore reflecting a genome-wide pattern. Considering previous findings of strong assortative mating within forms and general lack of hybrids in areas of sympatry, we propose that there is now sufficient direct and indirect evidence to consider both M and S molecular forms of *An. gambiae* as distinct species that have probably speciated recently.

Keywords: *Anopheles gambiae*, molecular forms, microsatellites, malaria, Africa.

Introduction

Reproductive isolation and incipient speciation within the species *Anopheles gambiae*, the most efficient malaria

vector in Sub-Saharan Africa, has been the focus of much research for more than two decades. The first evidence for genetic heterogeneity within this mosquito species came from cytological observations of the banding pattern of polytene chromosomes that revealed a complex system of polymorphic paracentric inversions leading to different chromosomal arrangements (Coluzzi & Sabatini, 1967; Coluzzi *et al.*, 1979). Frequencies of alternative arrangements, especially involving inversions on chromosome 2, were shown to correlate with ecological/climatic factors such as the degree of aridity of the environment, suggesting an adaptive potential for inversions, different combinations favouring survival under a variety of environmental conditions (Coluzzi *et al.*, 1979; Coluzzi, 1982; Touré *et al.*, 1994, 1998). Furthermore, extensive studies of karyotype distributions in natural *An. gambiae* populations often revealed strong and persistent deviations from Hardy–Weinberg equilibrium due to a deficit or even complete absence of certain heterokaryotypes, providing ample evidence against panmixia. These results led to the designation, in West Africa, of five 'chromosomal forms' named under a non-Linnean nomenclature as Forest, Savanna, Mopti, Bamako and Bissau (Coluzzi *et al.*, 1985; Touré *et al.*, 1998). These forms have been characterized based on the presence or absence of certain diagnostic inversions combinations (e.g. inversion 2Rb in Savanna, 2Rj in Bamako, 2Rbc/u in Mopti, 2Rd in Bissau), different relative frequencies of polymorphic arrangements (e.g. the standard arrangement 2R+/2L+ is almost fixed in Forest, but rare in the other forms), geographical distribution and ecological data (e.g. Mopti is adapted to dryer environments where it breeds all year long in irrigated fields, while Forest is exclusively found in more humid forested areas). They appear more or less genetically isolated in the field, presumably through prezygotic barriers because viable and fertile hybrids have been obtained in the laboratory (Di Deco *et al.*, 1980). Cytogenetic analysis alone, however, does not allow precise evaluation of the degree of hybridization between forms because of the presence of 'cryptic' heterokaryotypes impossible to identify with confidence, shared arrangements between forms (especially the standard uninverted karyotype 2R+) and the adaptive nature of inversions strongly

Received 11 July 2001, accepted after revision 4 September 2001 Correspondence Frédéric Simard, Laboratoire de Lutte contre les Insectes Nuisibles (LIN), Institut de Recherche pour le Développement (IRD), 911 Avenue Agropolis, BP 5045, 34 032 Montpellier Cedex 1, France Tel +33 (0)4 67 04 32 24; fax +33 (0)4 67 54 20 44, e-mail oceac@camnet.cm

exposed to selection, making them poor markers for studies of reproductive isolation (Touré *et al.*, 1998; della Torre *et al.*, 2001).

Based on the rationale that reproductive isolation would lead to genome-wide heterogeneity between chromosomal forms, recent efforts have focused on the pattern of variation observed with molecular markers that are not linked to inversions, in order to help resolving the taxonomic status of these forms (Favia *et al.*, 1994, 1997, 2001; Lanzaro *et al.*, 1998; Gentile *et al.*, 2001; Mukabayire *et al.*, 2001; della Torre *et al.*, 2001). All these studies reported negligible amounts of genetic differentiation between forms with the exception of loci in the ribosomal DNA (rDNA) cluster, on chromosome X. rDNA loci, however, did not support the subdivision of *An. gambiae* natural populations into five different incipient taxa, but suggested the existence of only two different entities, now referred to as molecular forms M and S (della Torre *et al.*, 2001). In Mali and Burkina Faso, exact correspondence between molecular and chromosomal forms has been found, all Mopti individuals belonging to the molecular form M while sympatric Savanna and Bamako specimens showed the S pattern (Favia *et al.*, 1997). This association breaks down in other areas of Africa, but reproductive isolation of the M and S forms is maintained (Chandre *et al.*, 1999; della Torre *et al.*, 2001). Thus, mate recognition seems to be independent of chromosomal inversions involved in ecotypic adaptation.

Accordingly, in forested areas of Central Africa, both M and S share the standard chromosome 2 arrangement, which is typical of the Forest chromosomal form (della Torre *et al.*, 2001; C. Wondji *et al.*, unpublished). Thus, the Forest form of *An. gambiae*, previously described as a single panmictic unit based on chromosomal arrangements frequencies, appears to be a variant common to both molecular forms. If M and S are effectively reproductively isolated from one another, genetic discontinuity within the Forest form should be observed in areas where both forms occur in sympatry. On the other hand, if environmental selection on inversions is the main force driving genetic variability within and between *An. gambiae* populations, genetic homogeneity is expected within the Forest chromosomal form.

To test this hypothesis, we studied genetic variation at ten microsatellite loci in five *An. gambiae* populations from the Forest chromosomal form, collected in four villages in South Cameroon. Sampling within the Forest form allowed us to assess the level of genetic differentiation between molecular forms while overcoming the effect of both geographical distance between populations (both forms exist in sympatry in this area), and differential selective pressure on inversions (both forms are homosequential in this environment, presenting only the standard, uninverted, chromosome 2 arrangement).

Results and discussion

Genotypes at ten microsatellite loci were determined for 242 *An. gambiae* specimens collected in four villages in Cameroon. In this area, the Forest chromosomal form is the only one encountered (or nearly so, Robert *et al.*, 1993; della Torre *et al.*, 2001; C. Wondji *et al.*, unpublished). All loci were highly polymorphic showing between 8 and 16 alleles per locus. Similar levels of variability were observed in all populations with a mean number of alleles per locus ranging from 8.3 to 10.0 and average unbiased estimates of heterozygosity across all loci approximately 0.80 (Table 1).

Heterogeneity of the gene pool

We applied a hierarchical design to detect population subdivision by testing for deviations from Hardy–Weinberg equilibrium (HWE) based upon a null hypothesis of heterozygote deficit.

First, we considered that all specimens belonged to the same, panmictic population. In this case, eight out of ten loci showed positive *F_{is}* values, heterozygote deficit being statistically significant in six cases (Table 2, 'pooled populations'). This finding, suggesting heterogeneity of the gene pool, was strengthened by analysis of linkage disequilibrium between loci because twelve out of forty-five tests were significant ($P < 0.001$, binomial test with 0.05 'success' rate) and involved loci that were not necessarily on the same chromosome. Altogether, these results are indicative of a Wahlund effect (mixing of different gene pools).

Geographical distance between populations may lead to genetic differentiation (isolation by distance), therefore we considered each village as a basic reproductive unit. In Nkoteng, Dschang and Mbalmayo, where all individuals analysed belonged to either one of the molecular forms, most of the loci conformed to Hardy–Weinberg expectations (Table 2). Significant deficit in heterozygotes were observed at less than three loci out of ten tested in any of these populations, and linkage disequilibrium analysis within each population resulted in one to four significant tests out of forty-five ($P > 0.18$). However, in Simbock, where both molecular forms S and M occur in sympatry, strong and highly significant deficit in heterozygotes was observed at six out of ten loci suggesting further separation of the gene pool in this location. Thus, we defined two subpopulations within Simbock, according to the molecular form (M and S) and tested for goodness of fit to Hardy–Weinberg expectations. In the M and S forms, respectively, two and three loci out of ten showed significant excess of homozygotes (loci Ag3H93 and Ag3H750 in the M form, and loci Ag3H93, Ag3H170 and Ag3H750 in the S form), which falls within the range observed in the other samples. No significant linkage disequilibrium was observed in the M population ($P > 0.39$) while six out of forty-five tests were statistically significant in the S population ($P = 0.024$).

Table 1. Genetic variability at ten microsatellite loci in *An. gambiae* from Cameroon

Locus		Populations					
		Nkoteng (2n = 106)	Dschang (2n = 108)	Simbock S (2n = 66)	Simbock M (2n = 98)	Mbalmayo (2n = 106)	All (2n = 484)
AgXH80 (X 2B)*	Nall	12	16	9	9	9	16
	Hobs	0.87	0.85	0.71	0.96	0.88	0.85
	Hexp	0.83	0.89	0.82	0.82	0.82	0.84
AgXH99 (X 2C)	Nall	6	7	5	6	7	8
	Hobs	0.74	0.72	0.64	0.59	0.79	0.70
	Hexp	0.80	0.74	0.67	0.70	0.76	0.74
AgXH49 (X 1D)	Nall	10	9	7	8	4	11
	Hobs	0.89	0.80	0.74	0.57	0.69	0.74
	Hexp	0.80	0.81	0.77	0.64	0.58	0.72
Ag2H417 (II 7)	Nall	6	8	5	7	7	10
	Hobs	0.65	0.65	0.68	0.76	0.86	0.72
	Hexp	0.56	0.71	0.62	0.69	0.70	0.66
Ag2H769 (II 9B)	Nall	12	12	12	14	11	16
	Hobs	0.77	0.88	0.80	0.76	0.87	0.82
	Hexp	0.90	0.88	0.88	0.86	0.88	0.88
Ag2H325 (II 23C-D)	Nall	8	8	7	7	8	9
	Hobs	0.66	0.75	0.78	0.77	0.64	0.72
	Hexp	0.80	0.76	0.77	0.71	0.65	0.74
Ag3H93 (III 29A)	Nall	10	11	7	7	7	11
	Hobs	0.79	0.85	0.43	0.55	0.75	0.67
	Hexp	0.85	0.82	0.76	0.78	0.73	0.79
Ag3H555 (III 35B-D)	Nall	9	9	10	8	8	11
	Hobs	0.83	0.81	0.73	0.73	0.87	0.80
	Hexp	0.81	0.84	0.81	0.72	0.77	0.79
Ag3H170 (III 35)	Nall	11	10	10	15	14	16
	Hobs	0.76	0.64	0.58	0.90	0.75	0.73
	Hexp	0.84	0.84	0.88	0.92	0.88	0.87
Ag3H750 (III 39A)	Nall	13	10	11	10	11	13
	Hobs	0.59	0.69	0.42	0.70	0.69	0.62
	Hexp	0.88	0.86	0.85	0.85	0.88	0.86
Mean across all loci	Nall	9.7	10	8.3	9.1	8.6	12.1
	Hobs	0.76	0.76	0.65	0.73	0.78	0.74
	Hexp	0.81	0.81	0.78	0.77	0.77	0.79

*Cytological location of the locus; 'All' refers to populations pooled. 2n, number of chromosomes scored; Nall, number of alleles per locus; Hobs, observed heterozygosity (direct count); Hexp, expected heterozygosity under Hardy-Weinberg equilibrium (Nei, 1978)

Although each population exhibited heterozygote deficit at some loci, it is more likely the result of locus-specific constraints such as null alleles (Callen *et al.*, 1993), limited allelic range (Epplen *et al.*, 1993) or preferential amplification of one allele in heterozygotes (Wattier *et al.*, 1998), rather than population substructure or inbreeding. Deviations from HWE are a common finding in population genetics studies that have been conducted so far using microsatellite loci in *An. gambiae* and/or its sibling species, *An. arabiensis* (Lehmann *et al.*, 1996a; Lanzaro *et al.*, 1998; Simard *et al.*, 2000; Donnelly & Townson, 2000) and direct evidence for null alleles have been reported (Lehmann *et al.*, 1996b; Walton *et al.*, 1998; Lanzaro *et al.*, 1998). Several evidences indeed suggest locus-specific effects in our dataset. First, loci Ag3H750 and Ag3H170 showed positive *Fis* values in all five populations tested and heterozygote deficits reached statistical significance in four and two cases, respectively.

High and significant positive *Fis* values were also observed at locus Ag3H93 in both populations from Simbock. However, when all populations are pooled or when the M and S forms are not set apart in Simbock, these deviations are strengthened by heterozygote deficits at other loci and the lack of heterozygotes appears as a genome-wide pattern, involving most of the loci (six out of ten in the pooled population, as well as in Simbock). Second, these loci appear in a state of linkage equilibrium at every level of the analysis, suggesting that the different alleles are distributed evenly between populations and the deviation observed does not result from genetic substructure. On the basis of considerably higher levels of heterogeneity observed when populations are pooled, we regarded each of the five populations we defined as panmictic units. However, because the effect of deviations from HWE on *Fst* estimators is unclear, we apply them with caution and

Table 2. Test for deviation from Hardy–Weinberg equilibrium: *F*_{is} and significance level for goodness of fit tests to Hardy–Weinberg equilibrium, based upon a null hypothesis of heterozygote deficit, for all *An. gambiae* populations

Population	Locus											Mean across loci*
	Location	AgXH80	AgXH99	AgXH49	Ag2H417	Ag2H769	Ag2H325	Ag3H93	Ag3H555	Ag3H170	Ag3H750	
Pooled populations		-0.010	+0.091	+0.027	-0.078	+0.085	+0.044	+0.126	+0.015	+0.164	+0.279	+0.074
Mbalmayo	(2 <i>n</i> = 484)	-0.061	-0.045	-0.201	-0.237	+0.010	+0.012	-0.029	-0.122	+0.145	-0.212	-0.031
Simbock	(2 <i>n</i> = 38)	-0.163	+0.159	+0.107	-0.091	-0.115	-0.081	+0.289	-0.024	+0.024	+0.184	-0.051
All	(2 <i>n</i> = 204)	-0.111	+0.064	-0.042	-0.164	+0.066	-0.030	+0.148	-0.058	+0.088	+0.204	-0.016
Area of sympatry		-0.018	+0.163	+0.147	-0.094	-0.126	-0.058	+0.356	+0.080	+0.155	+0.340	+0.119
M + S		-0.137	+0.052	+0.031	-0.093	-0.091	-0.022	+0.438	-0.097	+0.337	+0.505	+0.157
S		-0.053	+0.082	-0.113	-0.154	+0.143	+0.179	+0.063	-0.023	+0.096	+0.319	-0.054
		-0.047	+0.033	+0.017	+0.083	-0.002	+0.017	-0.029	+0.032	+0.232	+0.195	+0.062
		-0.039	+0.072	-0.022	-0.018	+0.081	+0.068	+0.101	+0.028	+0.211	+0.313	+0.088

*F*_{is} was calculated according to Weir & Cockerham (1984). Note that when populations are pooled, *F*_{is} = *F*_{it}. Significance levels were corrected according to the Bonferroni procedure to take into account multiple tests (ten tests per population). Bolded values: *P* < 0.05; bolded underlined values: *P* < 0.01. 2*n*, number of chromosome scored. *Mean *F*_{is} across loci and its statistical significance was estimated by bootstrapping over loci (5000 replicates).

consistently pay attention to the homogeneity of the results across loci.

It is worth noting that, in our study, loci Ag2H769 and Ag2H325 are located within chromosomal inversions that are polymorphic in *An. gambiae* (Ag2H769 is within inversion 2Rj and Ag2H325 is within 2La). Although selection on inversions may influence the genotype at linked microsatellites loci by genetic hitchhiking (Lanzaro *et al.*, 1998), it is unlikely in our case because an overwhelming majority of specimens from this area belong to the Forest chromosomal form (see Experimental procedures), which is characterized by standard (uninverted) homozygous karyotypes. However, no karyotypic information was available for the samples we have tested and putative influence of the cytological location of the locus (within or outside inversions) will be further considered.

Geographic differentiation within molecular forms

Levels of genetic differentiation between populations were assessed by estimating *F*_{st} (Weir & Cockerham, 1984) and tested using the exact test of genotypic differentiation (Goudet *et al.*, 1996).

Genotype distribution was similar at all loci in the three villages within the S form and between both M populations: only two out of thirty pair-wise tests showed significant heterogeneity within the S form (between Nkoteng and Simbock at loci AgXH99 and Ag3H93), and three out of ten tests were significant between M populations from Simbock and Mbalmayo (Table 3A). Single-locus *F*_{st} estimates were very low and did not significantly differ from 0, except in the above-mentioned cases (Table 3A). However, when information from all loci is combined, mean *F*_{st} estimates reached statistical significance except between populations of the S form collected in Dschang and Simbock. Despite very low values (0.007 < *F*_{st} < 0.017), significant *F*_{st} estimates suggest that populations are indeed genetically differentiated, probably as a result of isolation by distance. This finding, however, needs further consideration and additional samples are actually being processed to ascertain this hypothesis.

Genetic differentiation between molecular forms

Strikingly higher levels of differentiation were observed when comparing populations belonging to alternative molecular forms than when geographical differentiation was investigated within each molecular form (Table 3). Except for locus Ag2H417 and, to a lower extent Ag2H325 that showed limited amounts of genetic differentiation between forms, all pair-wise *F*_{st} estimates were highly homogenous across loci and were statistically significant (*P* < 0.01 in most cases, Table 3B). Removing from the analysis either loci that deviated from HWE or loci mapped within polymorphic inversions did not change the results obtained (Table 3). Thus, the pattern observed reflects a genome-wide characteristic

Table 3. Measure of genetic differentiation (*Fst*) between *An. gambiae* populations of the Forest chromosomal form (A) Within molecular forms

Locus	Within the molecular form S			Within the molecular form M
	Nkoteng vs Dschang	Nkoteng vs Simbock S	Dschang vs Simbock S	Simbock M vs Mhalmayo
AgXH80	0.008	0.003	0.006	0.001
AgXH99	0.012	0.047	0.014	0.031
AgXH49	0.002	0.019	0.024	0.003
Chromosome X	0.007	0.023	0.014	0.012
Ag2H417	0.041	0.000	0.006	0.008
Ag2H769	0.012	0.021	0.000	0.009
Ag2H325	0.000	0.000	0.004	0.012
Chromosome 2	0.014	0.008	0.003	0.010
Ag3H93	0.000	0.008	0.000	0.047
Ag3H555	0.000	0.006	0.004	0.034
Ag3H170	0.018	0.020	0.000	0.004
Ag3h750	0.000	0.003	0.000	0.015
Chromosome 3	0.002	0.009	0.000	0.024
All loci	0.007	0.013	0.005	0.017
[95% C I]*	[0–0.016]	[0.004–0.026]	[0–0.010]	[0.008–0.026]
Mean across seven loci†	0.009	0.014	0.008	0.014
Loci outside inversions‡	0.008	0.014	0.006	0.018

(B) Between molecular forms

Locus	(S Form) (M Form)	Nkoteng vs Simbock M	Nkoteng vs Mhalmayo	Dschang vs Simbock M	Dschang vs Mhalmayo	Simbock S vs Simbock M	Simbock S vs Mhalmayo	All S vs All M
AgXH80		0.030	0.016	0.037	0.023	0.041	0.014	0.025
AgXH99		0.044	0.029	0.029	0.063	0.102	0.140	0.047
AgXH49		0.068	0.094	0.095	0.131	0.151	0.167	0.104
Chromosome X		0.047	0.046	0.054	0.072	0.096	0.106	0.058
Ag2H417		0.012	0.022	0.001	0.035	0.000	0.021	0.008
Ag2H769		0.046	0.035	0.018	0.013	0.046	0.019	0.023
Ag2H325		0.023	0.059	0.021	0.061	0.000	0.032	0.032
Chromosome 2		0.029	0.039	0.014	0.035	0.016	0.024	0.022
Ag3H93		0.037	0.021	0.040	0.008	0.039	0.000	0.013
Ag3H555		0.049	0.044	0.049	0.025	0.105	0.074	0.044
Ag3H170		0.033	0.013	0.035	0.011	0.027	0.017	0.017
Ag3h750		0.034	0.015	0.043	0.011	0.078	0.026	0.030
Chromosome 3		0.038	0.023	0.042	0.014	0.063	0.030	0.026
All loci		0.038	0.035	0.038	0.037	0.060	0.052	0.035
[95% C I]*		[0.030–0.047]	[0.022–0.050]	[0.025–0.053]	[0.018–0.062]	[0.032–0.088]	[0.020–0.087]	[0.021–0.052]
Mean across seven loci†		0.040	0.043	0.037	0.050	0.065	0.068	0.041
Loci outside inversions‡		0.040	0.032	0.042	0.038	0.069	0.058	0.037

Bolded values: $P < 0.05$. Bolded underlined values: $P < 0.01$. *95% confidence intervals (C I) around *Fst* estimates based on all loci information were built by bootstrapping over loci (1000 permutations), †excluding loci showing significant heterozygote deficits at the population level (i.e. loci Ag3H93, Ag3H170 and Ag3H750); ‡excluding loci within polymorphic inversions (i.e. loci Ag2H769 and Ag2H325).

and led to overall *Fst* estimates of 0.035–0.052 between allopatric populations and $Fst = 0.060$ between sympatric populations in Simbock ($P < 0.01$), providing evidence for limited gene exchange between the M and S molecular forms of *An. gambiae*.

It is interesting to note that sympatric populations showed the highest level of genetic differentiation. Although differences are not statistically significant (95% confidence intervals of *Fst* estimates do overlap to a large extent, Table 3B), this pattern is very similar to what has been observed with cuticular hydrocarbons profiles. Milligan *et al.* (1993) indeed reported diagnostic differences in cuticular hydrocarbons composition between *An. gambiae*

chromosomal forms when found in sympatry, yet correct assignment of individuals to their respective form (and even species within the *An. gambiae* complex) was impossible between allopatric populations. Cuticular hydrocarbons are well known to play a role in species (mate or kin) recognition in different arthropods (see Flechter & Michener, 1987) and the pattern observed by Milligan *et al.* (1993) may thus reflect reproductive character displacement as a consequence of natural selection of a premating mechanism for the reinforcement of mate recognition in areas of sympatry. Reproductive character displacement towards differential cuticular hydrocarbons profiles has indeed been demonstrated in *Drosophila serrata* when sympatric with its closely

related species *D. birchii* (Higgie *et al.*, 2000), but it is unclear how selection acting specifically on cuticular hydrocarbons genes (or their putative regulation factors) could translate into a genome-wide pattern such as observed in the present study.

Taxonomic status of the molecular forms of An. gambiae

Our results clearly demonstrated the existence of a barrier to gene flow between homokaryotypic (Forest) populations of the S and M forms of *An. gambiae* in an area of sympatry. Although values of *Fst* estimated from allelic and genotypic frequencies distributions at ten microsatellite loci do not provide clear-cut evidence for complete reproductive isolation between molecular forms, the results presented above coupled with previous evidence for complete lack of M/S hybrids in Cameroon (out of a total of 950 specimens typed throughout the country; C. Wondji *et al.*, unpublished) represents a very strong argument for the species status of the M and S forms.

Our results are consistent with those initially obtained by Favia *et al.* (1997) using PCR RFLP in 203 field specimens collected in Mali, in an area of sympatry between Mopti, Savanna and Bamako, who reported complete absence of M/S hybrid patterns despite biased sampling towards inclusion of putative Mopti/Savanna hybrid karyotypes. Female hybrids produced in the laboratory, however, showed clear M/S patterns by PCR-RFLP as well as by the recently developed PCR diagnostic (Favia *et al.*, 1997, 2001). Further support for reproductive isolation of the M and S forms was provided by Chandre *et al.* (1999) through an extensive study of the distribution of the pyrethroid resistance gene *kdr*. In an area of sympatry in Ivory Coast, the authors found the *kdr* mutation (sometimes reaching frequency as high as 0.96) in the S form only; it was not detected in sympatric and synchronous M populations.

The occurrence of M/S hybrids in the field, however, has been reported several times. Della Torre *et al.* (2001) found three hybrid patterns out of 1161 *An. gambiae* adult mosquitoes tested from throughout Africa (hybrid frequency = $3/1161 = 0.26\%$), but two of them were observed in localities where the S form only had been recorded. The authors therefore mentioned contamination as a possible cause of the hybrid patterns they reported, but recognized that cross-mating between *An. gambiae* and its sibling species *An. arabiensis* also occurs in areas of sympatry, although at a lower frequency of approximately 0.1–0.2% (Touré *et al.*, 1998). Further evidence for the viability of M/S hybrids in the wild was provided by Taylor *et al.* (2001), who reported occurrence of M/S hybrid larvae at a frequency of 0–1.29% (average over 3 years = 0.71%) in Banambani village, Mali. In a nearby village from the same area, Tripet *et al.* (2001) identified an inseminated female showing the M/S pattern, which demonstrated that M/S hybrids can be produced in

the field, survive up to the adult stage and are reproductively active. Determination of the genotype of this female at four microsatellite loci ruled out the possibility of contamination and led the authors to estimate adult hybrid frequency at $1/329 = 0.3\%$. Recent findings of the highly adaptive *kdr* mutation in M mosquitoes from Benin and subsequent molecular analysis of the DNA sequence of a large upstream intron provided direct evidence for introgression between molecular forms (Weill *et al.*, 2000). Such an event, however, must be rare and recent because no M mosquitoes displaying the *kdr* mutation were found outside Benin despite thorough sampling throughout West and Central Africa (Chandre *et al.*, 1999; Weill *et al.*, 2000; della Torre *et al.*, 2001).

Thus, moderate estimates of *Fst* as observed in the present study may reflect shared ancestral polymorphism rather than ongoing gene flow. There is indeed strong evidence for recent dramatic changes in the demography and history of these taxa, mostly due to anthropic environmental changes (Coluzzi, 1999; Powell *et al.*, 1999), which suggest that the pattern of differentiation observed may relate to a transient state of incipient speciation, with *Fst* values that have not yet reached their equilibrium. Moreover, at a global (species) scale, both M and S forms may not have spent the $4Ne$ generations (since initial splitting) necessary for distinct phylogenetic lineages to establish (Takahata, 1992). Thus, attempts to differentiate populations through classical population genetics and/or phylogenetics would undoubtedly reveal conflicting patterns, even if gene flow between molecular forms is currently extremely restricted and exclusively limited to occasional events of introgression driven by selection on adaptive genes.

Conclusion

The primary objective of our study was to estimate levels of genetic differentiation between the M and S molecular forms of *An. gambiae* while overcoming confounding effects of both geographical distance between populations and environmental selection on chromosomal inversions. This was achieved by sampling at a local scale (both forms were found in sympatry in the same village) within the Forest chromosomal form of *An. gambiae* (standard chromosome 2 karyotype). We found highly significant genetic differentiation between molecular forms, whereas geographical isolates within each form were much more homogenous. These results are consistent with previous findings of strong assortative mating within both M and S forms and general lack of hybrids in areas of sympatry. Taking into account possible limitations of equilibrium models of genetic differentiation (i.e. Rousset, 1996), we would emphasize that there is now sufficient support to consider that the molecular forms of *An. gambiae* may represent distinct species that have speciated in a recent past.

Experimental procedures

Study sites and sampling

Mosquitoes were collected either after landing on human volunteers or by indoor pyrethrum spraying, in four localities of Cameroon: Simbock, Mbalmayo, Nkoteng and Dschang

Simbock (3°51'N, 11°30'E) is located approximately 5 km from the capital city Yaoundé, in a forested area within the guineo-congolese bio-climatic domain. Mbalmayo (3°30'N, 11°30'E) is located 35 km south of Simbock in similar ecological settings. Nkoteng (4°30'N, 12°03'E) is about 140 km north-east of Simbock, in a forest-savanna transition area and Dschang (5°25'N, 10°10'E), approximately 350 km north-west from Yaoundé, is located in forest-savanna highlands

A detailed analysis of the geographical distribution and relative prevalence of *An. gambiae* molecular and chromosomal forms in this area will be given elsewhere (C Wondji *et al.*, unpublished). Briefly, in Nkoteng and Dschang, prevalence of the S form was more than 98%. Mosquitoes belonging to the M molecular form were predominant in the other two localities considered in this paper, representing up to 97% of the specimens collected in Mbalmayo. However, in Simbock, the sympatric S population accounted for 9.5% of the total number of *An. gambiae* mosquitoes tested. Chromosomal identification was reported for twenty females from Simbock ($n = 8$), Nkoteng ($n = 8$) and Dschang ($n = 4$) and additional cytogenetics data can be found in Robert *et al.* (1993) and della Torre *et al.* (2001). All females examined showed typical Forest arrangements, except one specimen from Nkoteng, chromosomally identified as Savanna. The homozygous standard karyotype (2R+/2L+) was by far the most frequent, although two 2Rb/+ heterokaryotypes were found in Dschang, and one female 2Rb/+ and two 2La/+ carriers were found in Nkoteng. In Simbock, all the specimens presented the standard uninverted homokaryotype 2R+/2L+

Collections were conducted between December 1998 and January 2000 in Simbock, between December 1998 and April 2000 in Nkoteng, in June 2000 in Mbalmayo and in August 2000 in Dschang. Mosquitoes from the *An. gambiae* complex were visually sorted from other anophelines according to morphological identification keys (Gillies & Coetzee, 1987). All specimens were stored individually and kept at -20°C until further analysis

DNA extraction and genotype scoring

Genomic DNA was extracted from wings or legs of each individual mosquito as described by Cornel & Collins (1996) and resuspended in 100 μL of TE buffer. Only female *An. gambiae* specimens were included in the analysis after species identification was carried out by the diagnostic PCR described by Scott *et al.* (1993). Determination of the molecular form (M/S) of each specimen was done by the diagnostic PCR-based assay recently described by Favia *et al.* (2001), after validation of the technique for the study of field specimens collected in Cameroon (C Wondji *et al.*, unpublished). Each specimen was unequivocally assigned to one or the other of the M or S forms and no hybrid pattern was observed.

Ten microsatellite loci were selected from published *An. gambiae* sequence data (Zheng *et al.*, 1996; Wang *et al.*, 1999) based on high polymorphism, mean allele sizes ranging from 80 to 150 bp, and cytological location spread over the three chromosome pairs. We used loci AgXH80, AgXH99 and AgXH49 on chromosome X, loci Ag2H417, Ag2H769 and Ag2H325 on chromosome 2 and loci Ag3H93, Ag3H555, Ag3H170 and Ag3H750 on chromosome 3. Microsatellite alleles were PCR amplified as previously described

(Simard *et al.*, 1999, 2000). PCR products were loaded on 15% non-denaturing polyacrylamide gels and the allelic bands were visualized after rapid silver staining (Sanguinetti *et al.*, 1994).

A set of approximately fifty specimens per population was investigated but, due to the strong predominance of one molecular form to the other in each of the locations sampled, only specimens from the S form were further considered from the villages of Dschang and Nkoteng, while only the M population from Mbalmayo could be investigated. Intensive sampling in Simbock, however, allowed us to consider one sample of each molecular form, although only thirty-three specimens of the S form were available.

Data analysis

Microsatellite-based genotype frequencies were tested against Hardy-Weinberg expectations for each locus in the pooled population, in each location, and within M and S populations from Simbock. Statistical significance was assessed by the exact probability test available in GENEPOP 3.2 (Raymond & Rousset, 1995). Exact tests of linkage disequilibrium between pairs of loci were computed in the pooled population and within each population, using GENEPOP 3.2. The sequential Bonferroni procedure (Holm, 1979) was applied to evaluate significance when multiple tests were performed.

Differentiation between populations was assessed by *F*-statistics (Wright, 1978), calculated according to Weir & Cockerham (1984). Significance of *F_{is}* and *F_{st}* was assessed using the exact probability test (Rousset & Raymond, 1995) and the G-based exact test of genotypic differentiation (Goudet *et al.*, 1996), respectively (tests available in GENEPOP 3.2). Where mentioned, 95% confidence intervals around *F_{st}* estimates were built by bootstrapping over loci (1000 randomizations).

Acknowledgements

This work was supported in part by grant N°980101 from MIM/TDR/WHO and by the French Institut de Recherche pour le Développement (IRD) – Research unit R016. We are indebted to T. Tchuinkam, C. A. Nkondjio, P. Awono-Ambene, J.-C. Toto, J.-Y. Meunier, S. Zebaze-Kemleu and R. Nyambam and for valuable help in the field and useful technical support in the lab. We thank F. Rousset, M. Raymond and three anonymous reviewers for critical comments that greatly improved earlier versions of the manuscript.

References

- Callen, D.F., Thompson, A.D., Shen, Y., Phillips, H.A., Richards, R.I., Mulley, J.C. and Sutherland, G.R. (1993) Incidence and origin of 'null' alleles in the (AC)_n microsatellite markers. *Am J Hum Gen* **22**: 1–10
- Chandre, F., Manguin, S., Brengues, C., Dossou-Yovo, J., Darriet, F., Diabaté, A., Carnevale, P. and Guillet, P. (1999) Current distribution of pyrethroid resistance gene (*ldr*) in *Anopheles gambiae* complex from West Africa and further evidence for reproductive isolation of the Mopti form. *Parassitologia* **41**: 319–322.
- Coluzzi, M. (1982) Spatial distribution of chromosomal inversions and speciation in anopheline mosquitoes. In: *Mechanisms of Speciation* (ed. C. Barigozzi), pp. 143–153. A.R. Liss, Inc., New York.

- Coluzzi, M. (1999) The clay feet of the malaria giant and its African roots: hypotheses and inferences about origin, spread and control of *Plasmodium falciparum*. *Parassitologia* **41**: 277–283.
- Coluzzi, M., Petrarca, V. and Di Deco, M.A. (1985) Chromosomal inversion intergradation an incipient speciation in *Anopheles gambiae*. *Boll Zool* **52**: 45–63.
- Coluzzi, M. and Sabatini, A. (1967) Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. *Parassitologia* **9**: 73–88.
- Coluzzi, M., Sabatini, A., Petrarca, V. and Di Deco, M.A. (1979) Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Trans R Soc Trop Med Hyg* **73**: 483–497.
- Cornel, A. and Collins, F.H. (1996) PCR of the Ribosomal DNA Intergenic Spacer Regions as a method for identifying mosquitoes in the *Anopheles gambiae* complex. *Meth Mol Biol* **50**: 321–332.
- Di Deco, M.A., Petrarca, V., Villani, F. and Coluzzi, M. (1980) Polimorfismo cromosomico da inversioni paracentriche ed eccesso degli eterocariotipi in ceppi di *Anopheles* allevati in laboratorio. *Parassitologia* **22**: 304–306.
- Donnelly, M.J. and Townson, H. (2000) Evidence for extensive genetic differentiation among populations of the malaria vector *Anopheles arabiensis* in Eastern Africa. *Insect Mol Biol* **9**: 357–367.
- Epplen, C., Melmer, G., Siedlaczek, I., Schwaiger, F.W., Mäueler, W. and Epplen, J.T. (1993) On the essence of 'meaningless' simple repetitive DNA in Eucaryotes genomes. In: *DNA Fingerprints: State of the Science* (eds S. D. J. Pena, R. Chakraborty, J. T. Epplen and A. J. Jeffreys), pp. 29–45. Birkhauser Verlag, Basel, Switzerland.
- Favia, G., della Torre, A., Bagayoko, M., Lanfrancotti, A., Sagnon, N.F., Touré, Y.T. and Coluzzi, M. (1997) Molecular identification of sympatric chromosomal forms of *Anopheles gambiae* and further evidence of their reproductive isolation. *Insect Mol Biol* **6**: 377–383.
- Favia, G., Dimopoulos, G., della Torre, A., Touré, Y.T., Coluzzi, M. and Louis, C. (1994) Polymorphism detected by random PCR distinguish between different chromosomal forms of *Anopheles gambiae*. *Proc Natl Acad Sci USA* **91**: 10315–10319.
- Favia, G., Lanfrancotti, A., Spanos, L., Siden-Kiamos, I. and Louis, C. (2001) Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. *Insect Mol Biol* **10**: 19–23.
- Flechter, D.J.C. and Michener, C.D. (1987) *Kin Recognition in Animals*. John Wiley & Sons, New York.
- Gentile, G., Slotman, M., Ketmaier, V., Powell, J.R. and Caccone, A. (2001) Attempts to molecularly distinguish cryptic taxa in *Anopheles gambiae* s.s. *Insect Mol Biol* **10**: 25–32.
- Gillies, M.T. and Coetzee, M. (1987) *A supplement to the Anophelinae of Africa South of the Sahara*, 2nd edn. Pub South Afr Inst Med Res, Johannesburg, South Africa. 55, p. 143.
- Goudet, J., Raymond, M., De Meeüs, T. and Rousset, F. (1996) Testing differentiation in diploid populations. *Genetics* **144**: 1933–1940.
- Higgie, M., Chenoweth, S. and Blows, M.W. (2000) Natural selection and the reinforcement of mate recognition. *Science* **290**: 519–521.
- Holm, S. (1979) A simple sequentially rejective multiple test procedure. *Scand J Statistics* **6**: 65–70.
- Lanzaro, G.C., Touré, Y.T., Carnahan, J., Zheng, L., Dolo, G., Traoré, S., Petrarca, V., Vernick, K.D. and Taylor, C.E. (1998) Complexities in the genetic structure of *Anopheles gambiae* populations in West Africa as revealed by microsatellite DNA analysis. *Proc Natl Acad Sci USA* **95**: 14260–14265.
- Lehmann, T., Hawley, W.A. and Collins, F.H. (1996b) An evaluation of evolutionary constraints on microsatellite loci using null alleles. *Genetics* **144**: 1155–1163.
- Lehmann, T., Hawley, W.A., Kamau, L., Fontenille, D., Simard, F. and Collins, F.H. (1996a) Genetic differentiation of *Anopheles gambiae* populations from East and West Africa: comparison of microsatellite and allozyme loci. *Heredity* **77**: 192–200.
- Milligan, P.J.M., Phillips, A., Broomfield, G., Molyneux, D.H., Touré, Y.T. and Coluzzi, M. (1993) A Study of the use of gas chromatography of cuticular hydrocarbons for identifying members of the *Anopheles gambiae* (Diptera: Culicidae) complex. *Bull Entomol Res* **83**: 613–624.
- Mukabayire, O., Caridi, J., Wang, X., Touré, Y.T., Coluzzi, M. and Besansky, N.J. (2001) Patterns of DNA sequence variation in chromosomally recognized taxa of *Anopheles gambiae*: evidence from rDNA and single-copy loci. *Insect Mol Biol* **10**: 33–46.
- Nei, M. (1978) Estimation of the average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
- Powell, J.R., Petrarca, V., della Torre, A., Caccone, A. and Coluzzi, M. (1999) Population structure, speciation and introgression in the *Anopheles gambiae* complex. *Parassitologia* **41**: 101–113.
- Raymond, M. and Rousset, F. (1995) Genepop (version 1.2), a population genetics software for exact tests and ecumenicism. *J Hered* **86**: 248–249.
- Robert, V., Petrarca, V., Le Goff, G. and Manga, L. (1993) Quelques données cytogénétiques sur le complexe *Anopheles gambiae* au Sud-Cameroun. *Bulletin de liaison et de documentation de l'OCEAC* **26**: 99–101.
- Rousset, F. (1996) Equilibrium values of measures of population subdivision for stepwise mutation processes. *Genetics* **142**: 1357–1362.
- Rousset, F. and Raymond, M. (1995) Testing heterozygote excess and deficiency. *Genetics* **140**: 1413–1419.
- Sanguinetti, C.J., Neto, E.D. and Simpson, A.J.G. (1994) Rapid silver staining and recovery of PCR products separated on polyacrylamide gels. *Biotechniques* **17**: 915–919.
- Scott, J.A., Brogdon, W.G. and Collins, F.H. (1993) Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. *Am J Trop Med Hyg* **49**: 520–529.
- Simard, F., Fontenille, D., Lehmann, T., Girod, R., Brutus, L., Gopaul, R., Dournon, C. and Collins, F.H. (1999) High amounts of genetic differentiation between populations of the malaria vector *Anopheles arabiensis* from West Africa and eastern outer islands. *Am J Trop Med Hyg* **60**: 1000–1009.
- Simard, F., Lehmann, T., Lemasson, J.J., Diatta, M. and Fontenille, D. (2000) Persistence of *Anopheles arabiensis* during the severe dry season conditions in Senegal: an indirect approach using microsatellite loci. *Insect Mol Biol* **9**: 467–479.
- Takahata, N. (1992) Evolutionary genetics of human paleo-populations. In: *Mechanisms of Molecular Evolution: Introduction to Molecular Biology* (eds N. Takahata and A. G. Clark), pp. 1–21. Sinauer Associates Inc, Sunderland, MA.
- Taylor, C., Touré, Y.T., Carnahan, J., Norris, D.E., Dolo, G., Traoré, S., Edillo, F.E. and Lanzaro, G.C. (2001) Gene flow among populations of the malaria vector, *Anopheles gambiae*, in Mali, West Africa. *Genetics* **157**: 743–759.
- della Torre, A., Fanello, C., Akogbeto, M., Dossou-Yovo, J., Favia, G., Petrarca, V. and Coluzzi, M. (2001) Molecular evidence of

- incipient speciation within *Anopheles gambiae* s.s. in West Africa. *Insect Mol Biol* **10**: 9–18
- Touré, Y.T., Petrarca, V., Traoré, S.F., Coulibaly, A., Maiga, H.M., Sankaré, O., Sow, M., Di Deco, M.A. and Coluzzi, M. (1994) Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae* s.str. Mali, West Africa. *Genetica* **94**: 213–223.
- Touré, Y.T., Petrarca, V., Traoré, S.F., Coulibaly, A., Maiga, H.M., Sankaré, O., Sow, M., Di Deco, M.A. and Coluzzi, M. (1998) Distribution and inversion polymorphism of chromosomally recognised taxa of the *Anopheles gambiae* complex in Mali, West Africa. *Parassitologia* **40**: 477–511
- Tripet, F., Touré, Y.T., Taylor, C.E., Norris, D.E., Dolo, G. and Lanzaro, G. (2001) DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. *Mol Ecol* **10**: 1725–1732.
- Walton, C., Thelwell, N.J., Priestman, A. and Butlin, R.K. (1998) The use of microsatellites to study gene flow in natural populations of *Anopheles* malaria vectors in Africa: potential and pitfalls. *J Am Mosq Control Assoc* **14**: 266–272
- Wang, R., Kafatos, F.C. and Zheng, L. (1999) Microsatellite markers and genotyping procedures for *Anopheles gambiae*. *Parasitol Today* **15**: 33–37.
- Wattier, R., Engel, C.R., Saumitou-Laprade, P. and Valero, M. (1998) Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Mol Ecol* **7**: 1569–1573.
- Weill, M., Chandre, F., Brengues, C., Manguin, S., Akogbeto, M., Pasteur, N., Guillet, P. and Raymond, M. (2000) The *kdr* mutation occurs in the Mopti form of *Anopheles gambiae* s.s. through introgression. *Insect Mol Biol* **9**: 451–455.
- Weir, B.S. and Cockerham, C.C. (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**: 1358–1370
- Wright, S. (1978) *Evolution and Genetics of Populations. Vol 4 Variability Within and Among Natural Populations*. University of Chicago Press, Chicago, IL.
- Zheng, L., Benedict, M.Q., Cornel, A.J., Collins, F.H. and Kafatos, F.C. (1996) An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* **143**: 941–952

ANNEXE 2

Antonio Nkondjio C., Awono-Ambene P., Toto J.C., Nyamban R., Kemleu-Zébazé S. Meunier J.Y., **Wondji C.**, Tchuinkam T & Fontenille D (2002). High malaria transmission in a village close to Yaoundé, the capital city of Cameroon. *Journal of Medical Entomology*. **39 (2)** : 350-355.

High Malaria Transmission Intensity in a Village Close to Yaounde, the Capital City of Cameroon

CHRISTOPHE ANTONIO-NKONDJIO,¹ PARFAIT AWONO-AMBENE, JEAN-CLAUDE TOTO, JEAN-YVES MEUNIER,¹ SYLVIE ZEBAZE-KEMLEU,¹ ROSE NYAMBAM,¹ CHARLES S. WONDJI,¹ TIMOLÉON TCHUINKAM,² AND DIDIER FONTENILLE^{1,3}

Organisation de Coordination pour la Lutte Contre les Endémies en Afrique Centrale, P.O. Box 288, Yaoundé, Cameroon

J. Med. Entomol. 39(2): 350-355 (2002)

ABSTRACT A 2-yr longitudinal malaria study was undertaken in a suburb of Yaounde, the capital city of Cameroon, in the village of Simbock, ≈2 km from the city limits. This study allowed assessment of malaria transmission intensity and dynamics in this region before implementation of pyrethroid impregnated bed nets through the national vector control program. Anophelines were captured on human volunteers by pyrethrum spray collections and in resting sites outdoors. Malaria vectors were *Anopheles funestus* Giles, *Anopheles gambiae* s.s. Giles (M and S forms), *Anopheles moucheti* Evans, and *Anopheles nili* Theobald. *An. moucheti* was the most abundant mosquito captured during the study, accounting for >54% of total anophelines caught. The annual *Plasmodium falciparum* Welch entomological inoculation rates measured by enzyme-linked immunosorbent assay were 277 infected bites per human for the first year and 368 for the second year. *An. gambiae* s.s., *An. funestus*, *An. moucheti*, and *An. nili* were responsible for 23.8%, 26.8%, 39.2%, and 10.2% of malaria transmission, respectively. Malaria transmission is perennial throughout the year. All these vectors were highly anthropophilic because only two out of 566 mosquitoes blood-meal tested were not taken on humans.

KEY WORDS *Anopheles gambiae*, *Anopheles funestus*, *Anopheles moucheti*, *Anopheles nili*, malaria, anopheline

IN CAMEROON, AS in many other countries of the central African region, major vectors of malaria are *Anopheles gambiae* s. s. Giles and *Anopheles funestus* Giles present all over the country, *Anopheles nili* Theobald in borders of rivers and streams, *Anopheles moucheti* Evans in forest areas, and *Anopheles arabiensis* Patton in the northern part of the country (Mouchet and Gariou 1966, Carnevale et al. 1992, Njan Nloga et al. 1993). Some vectors considered as minor take a locally active part in malaria transmission. *Anopheles melas* Theobald is probably a good vector in the coastal region. *Anopheles paludis* Theobald, *Anopheles pharoensis* Theobald, and *Anopheles hancocki* Edwards have been found infected by *Plasmodium falciparum* Welch in Cameroon (Pajot and Segers 1964, Robert et al. 1992, Fontenille et al. 2000). *Anopheles wellcomei* Theobald and *Anopheles marshallii* Theobald (or/and *Anopheles hargreavesi* Evans) might occasionally transmit malaria to humans. Many studies conducted in the forest area of south Cameroon in the 1990s proved that malaria vectors relative abundance and transmission intensity differ greatly from one village to

another (Fondjo et al. 1992, Manga et al. 1992, Le Goff et al. 1993, Meunier et al. 1999). Significant differences can also be found from year to year in the same locality, especially in rapidly changing environments as is the case in the suburbs of Yaounde, where deforestation is increasing and urbanisation is rising (Manga et al. 1995). Since the Cameroon government is engaging itself in the fight against malaria by implementing vector control programs under the Roll back malaria initiative, we have conducted a longitudinal study of malaria transmission in the village of Simbock, nearby Yaounde, the capital city of Cameroon.

Materials and Methods

Study Area. This study was carried out in a quarter called "block 6" of the village of Simbock (3° 50' N, 11° 30' E). About 300 inhabitants were living in this area situated in a rural forest region of Cameroon, only 2 km from Yaounde city. Most of the houses are built in the traditional style with mud walls and roofs of corrugated iron. As time goes by, modern urbanized dwellings come closer and closer to the village. The inhabitants stay inside their houses from ≈2100 hours, and rarely use mosquito netting. Most of the people living in Simbock work in Yaounde town. They go to

¹ Institut de Recherche pour le Développement (IRD), P.O. Box 1857, Yaoundé, Cameroon.

² University of Dschang, P.O. Box 67, Dschang, Cameroon.

³ E-mail: didier.fontenille@mpl.ird.fr.

work early in the morning every day, and are back in the evening. The animal life in the village consists of few dogs, occasional goats and pigs, and chickens. There are no bovinds. The village is built on the slope of a hill between 700 and 750 m above sea level. The river Mefou, which is ≈ 100 m away, creates a permanent swamp. Ditches of sand are dug in the valley. The climate is equatorial. There are two rainy seasons, extending from March to June and from September to November. Average annual rainfall during the last 18 yr was 1,466 mm. During the study, rainfalls varied annually: 1,480 mm in 1998, 1,792 mm in 1999, and 1,554 mm in 2000. The average minimum and maximum monthly temperatures, recorded by the national meteorological service, were constantly high, ranging from 18–25°C in July to 20–29°C in March.

Mosquito Collections. Adult mosquitoes were captured monthly from November 1998 to September 1999 and every 2 mo from November 1999 to September 2000. The following four collection techniques were used: (1) adult volunteers captured mosquitoes landing on their legs from 1900 to 0600 hours at the same sites, in three different indoor locations in the village, for two consecutive nights every survey (two teams of volunteers collecting the mosquitoes during each night: one from 1900 to 0100 hours, the other from 0100 to 0600 hours). In total, 132 indoor human-nights were conducted during the study. The human biting rate was expressed as the number of mosquito bites per person per night during each survey sample. (2) From November 1999 to September 2000, collectors were also positioned outdoors at each of the three sites. All volunteers gave their consent for capturing mosquitoes and were given malaria prophylaxis. (3) Pyrethrum spray collections were made in the afternoon inside three to four bedrooms, which were different from the houses used for human bait collections. (4) Outside, resting mosquitoes were collected from a pit shelter, a deserted animal shelter, and in an empty 200-liter barrel.

Field Processing of Anophelines. Anophelines were identified using morphological characteristics according to the identification key of Gillies and De Meillon (1968), and Gillies and Coetzee (1987). Ovaries from a portion of female anophelines captured on human volunteers were dissected to determine parity (Detinova 1962). All the anophelines were stored individually in numbered tubes with desiccant for laboratory processing in Yaounde.

Laboratory Processing of Anophelines. Bloodmeal sources of a sample of females captured by pyrethrum spray were identified by an enzyme-linked immunosorbent assay (ELISA) (Beier et al. 1988). The technique identified human, bovine, ovine (sheep and goat), equine (horse and donkey), pig, or chicken host. The head and thorax of female anophelines were tested for the presence of circumsporozoite protein of *Plasmodium falciparum* Welch, *Plasmodium malariae* Laveran, and *Plasmodium ovale* Stephens by ELISA, as described by Burkot et al. (1984) and modified by Wirtz et al. (1987). *Plasmodium vivax* Grassi & Feletti is not present in this region of Africa. The circum-

Table 1. Number of anophelines captured from November 1998 to October 2000 in Simbock

Species	Feeding		Resting		Total
	Indoors	Outdoors	Bedrooms	Shelters	
<i>An. coustani</i>	1	0	0	1	2
<i>An. funestus</i>	611	92	595	8	1306
<i>An. gambiae</i>	330	58	65	1	454
<i>An. hancocki</i>	2	0	1	0	3
<i>An. mouchetti</i>	2,204	440	284	7	2,935
<i>An. nili</i>	493	65	53	1	612
<i>An. paludis</i>	38	6	1	0	45
<i>An. ziemanni</i>	28	2	0	1	31
<i>An. smithii</i>	0	0	0	1	1
Total	3,707	663	999	20	5,389

sporozoite protein rate and the 95% CI were calculated. The entomological inoculation rate was calculated by multiplying the human biting rate by the circumsporozoite protein rate for each sampling period. A previous work conducted in the same village had shown that CSP Elisa is a valuable method for measuring malaria transmission (Fontenille et al. 2001). Females belonging to the *An. gambiae* complex were identified to species using the polymerase chain reaction (PCR) technique described by Scott et al. (1993). A leg or a wing was placed directly into the reaction mixture containing the species-specific primers, dNTPs, buffer, and polymerase. The length of the amplified sequences was 315 nucleotides for *An. arabiensis*, 390 for *An. gambiae*, and 464 for *An. melas*. Specimens identified as *An. gambiae* were then tested for M and S molecular forms following the recently described diagnostic PCR-based assay (Favia et al. 2001).

Results

From November 1998 to September 2000, 5,389 *Anopheles* specimens were captured (Table 1). Of the *An. gambiae* complex females captured, 324 were identified by PCR, and all were *An. gambiae*, 90.5% being from the M molecular form, the remainder being from the S molecular form (Wondji et al. 2002).

The human indoor biting rate for each species varied temporally, depending on the seasons. *An. mouchetti*, the main vector in the area, was present throughout the year with a biting rate almost always >10 bites/night (Fig. 1).

Night-biting cycles were similar for every species with an increase in the second part of the night (Fig. 2), and did not vary during the year (data not shown). In total, 80% of the bites of *An. gambiae*, 77% of *An. funestus*, 68% of *An. nili*, and 65% of *An. mouchetti* occurred after midnight.

When capturing mosquitoes, both indoors and outdoors, during the same night, from March to September 2000, it was shown that 60.1% of *An. gambiae*, 61.5% of *An. funestus*, 51% of *An. mouchetti*, and 56.1% of *An. nili* were captured indoors, suggesting these anopheline populations were mainly endophagic. There was no difference between species, with the exception of *An.*

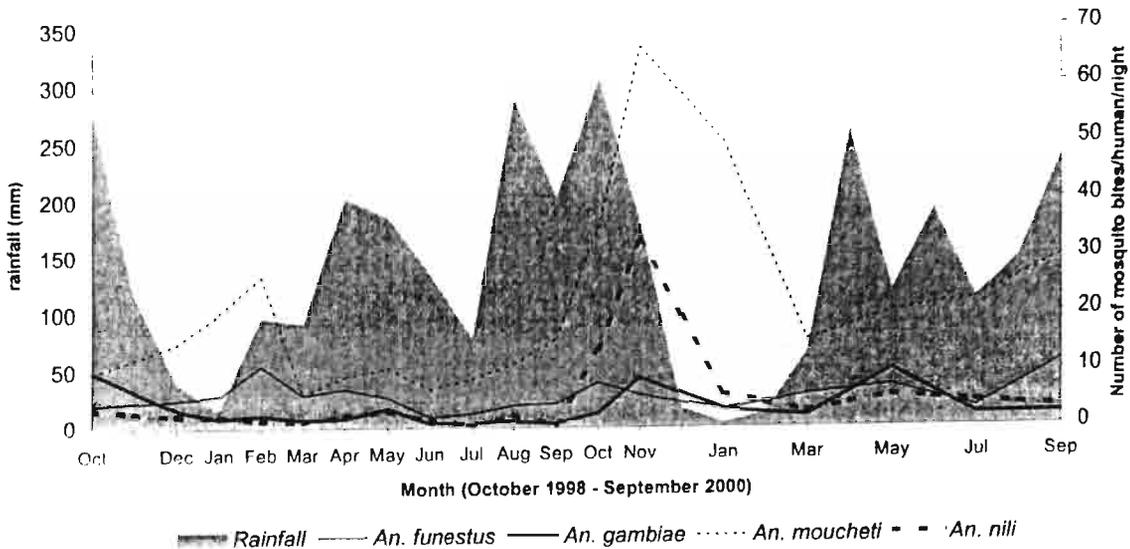


Fig. 1. Rainfall and indoor human biting rates for each vector species in Simbcock, from October 1998 to September 2000.

funestus being significantly more endophagic than *An. moucheti* ($\chi^2 = 8.77$, $df = 1$, $P = 0.003$).

Blood meals from 39 resting *An. gambiae*, 299 *An. funestus*, 192 *An. moucheti*, and 36 *An. nili* were tested by ELISA. All fed on human. Two of them also fed on pig or bovine.

In total, 1,548 anophelines were classified into parous or nulliparous females. The parous rates were the following: *An. gambiae* 78.6% (70.9–85.1 CI), *An. funestus* 77.4% (72.9–81.5 CI), *An. moucheti* 76.4% (73.3–79.3 CI), and *An. nili* 74.2% (67.7–79.9 CI).

The circumsporozoite protein rate was calculated monthly for each species. Overall, 97.8% of identified *Plasmodium* were *P. falciparum* and 2.2% were *P. malariae*. No *P. ovale* infection was found. Among mos-

quitoes captured landing on human, 5.6% of *An. gambiae*, 4.1% of *An. funestus*, 1.5% of *An. moucheti*, and 1.6% of *An. nili* were tested positive for *P. falciparum*. All four species taken together, these differences were highly significant ($\chi^2 = 35.6$, $df = 3$, $P < 10^{-6}$), whereas infection rates were similar between *An. gambiae* and *An. funestus* ($P = 0.27$) and between *An. nili* and *An. moucheti* ($P = 0.84$). The CSP rates were not significantly different when calculated only on mosquitoes captured on human, or on resting mosquitoes (Table 2). No other mosquito species were found positive in Simbcock.

The annual entomological inoculation rate was 277 (infected bites/humans/yr) from October 1998 to September 1999 and 368 from October 1999 to Sep-

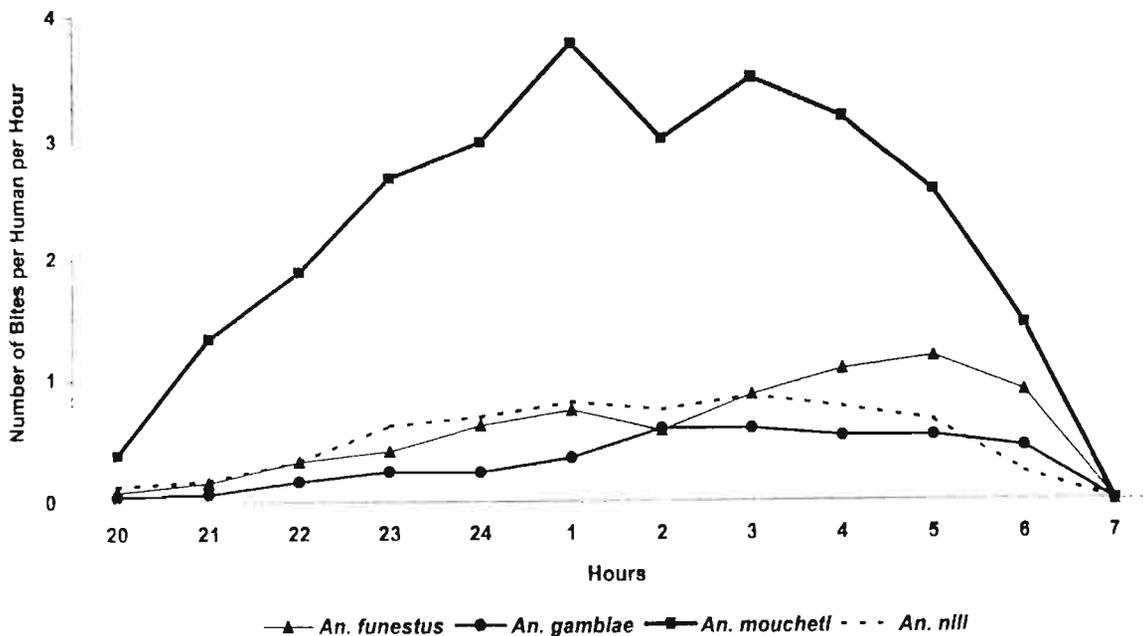


Fig. 2. Indoor night biting cycle (number of bites per human per hour) for each vector species in Simbcock, calculated from the 2 yr of captures.

Table 2. Infection rate for *Plasmodium falciparum*, calculated by ELISA from the head-thoraces of mosquitoes captured resting in houses or after landing on human

Species	Resting mosquitoes			Landing mosquitoes		
	No. tested	Positive	CSP rate (95% CI)	No. tested	Positive	CSP rate (95% CI)
<i>An. funestus</i>	593	36	6.07% (4.3-8.3)	682	28	4.11% (2.7-5.9)
<i>An. gambiae</i>	63	2	3.17% (0.38-11.0)	356	20	5.62% (3.5-8.5)
<i>An. moucheti</i>	284	4	1.41% (0.38-3.57)	2425	36	1.48% (1.04-2.04)
<i>An. nili</i>	53	1	1.89% (0.05-10.1)	564	9	1.6% (0.73-3.01)

tember 2000 (Table 3). *An. gambiae*, *An. funestus*, *An. moucheti*, and *An. nili* were responsible for 23.8, 26.8, 39.2, and 10.2% of *P. falciparum* transmission, respectively. Transmission was continuous, and reached its peak in January 1999 and May 2000 when an average of 1.65 infected bites per human per night was observed indoors. Peaks of transmission of the four vector species did not occur at the same time as shown in Fig. 3.

Discussion

Rainfall almost every month during the 2 yr of our survey and the presence of different kinds of mosquito larval development sites (e.g., a permanent river, a marsh, borrow pits, rain pools) explain the sympatric

presence of the four main African malaria vectors in Simbock. Density cycles of *An. gambiae*, *An. funestus*, *An. moucheti*, and *An. nili* depend partly on rainfall. For each species, the lower human indoor biting rates were observed at the end of the dry season, around March.

In Simbock, as in most of central Africa, sporozoite rates were three times higher in *An. gambiae* and *An. funestus* than in *An. moucheti* and *An. nili* (Fontenille and Locheuarn 1999). However, the proportion of females fed on humans among mosquitoes resting inside as well as parity rates were similar between the different vector species, suggesting a higher vectorial competence for *An. gambiae* and *An. funestus*.

The total entomological inoculation rate was around 300 infected bites per human per year during both years surveyed. Even if CSP Elisa overestimates the true transmission level by a factor 1.12, as was shown recently (Fontenille et al. 2001), the level of the transmission remains very high for a village located in the suburbs of the capital city. *An. moucheti*, which is often considered a minor vector as compared with *An. gambiae* or *An. funestus*, was the main vector in Simbock. Previous studies conducted in Yaounde or in nearby closely suburbs reported annual entomological inoculation rate ranging from 3 to 33, generally only

Table 3. Annual entomological inoculation rates (number of infected bites per human per year) per vector in Simbock

Species	First year	Second year	% (over 2 yr)
<i>An. funestus</i>	88.1	85	26.8%
<i>An. gambiae</i>	58.4	95.1	23.8%
<i>An. moucheti</i>	112.2	140.4	39.2%
<i>An. nili</i>	18.1	47.4	10.2%
Total	276.8	367.9	

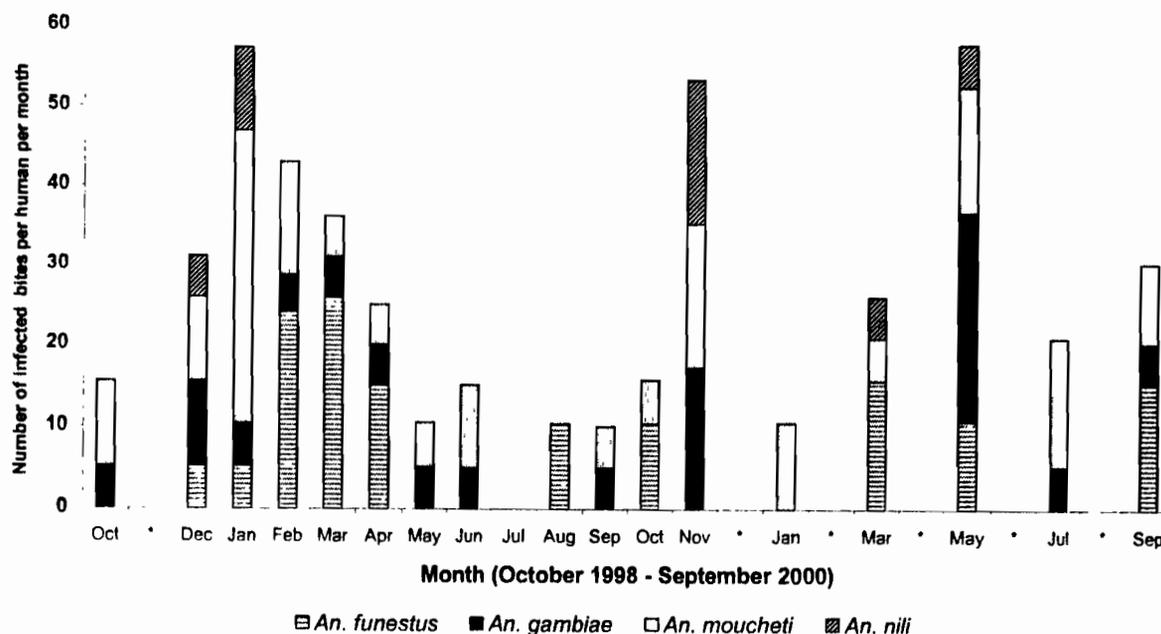


Fig. 3. Monthly entomological inoculation rate for each vector species in Simbock, from October 1998 to September 2000. (*No survey this month).

due to *An. gambiae*, the other being very rare vector species (Fondjo et al. 1992, Manga et al. 1992, Nimpaye et al. 2001).

Although transmission intensity is usually low in most African cities (Robert et al. 1986, Lindsay et al. 1990), such a high level of transmission in suburban areas has been observed in several countries. In 1991 and 1992 in the town of Bouake, Cote d'Ivoire, Dossou-Yovo et al. (1998) observed an annual entomological inoculation rate reaching 155 in some areas. In Kinshasa (DRC former Zaire), Coene (1993) observed an annual entomological inoculation rate of 30 within the city, whereas the entomological inoculation rate reached 455 in a suburban village located 15 km away from the city center. In Brazzaville (Congo), Trape and Zoulani (1987) showed a high heterogeneity in malaria transmission, with an annual entomological inoculation rate >100 infective bites per human per year in some areas.

In areas like Simbock, because plasmodia are transmitted at a high rate and continuously throughout the year by four different vector species, implementation of an efficient vector control strategy could not be easily attainable, despite the high anthropophily and endophily of the vectors. The village, located ≈2 km from classic urban areas, constitutes a "reservoir" of parasites and vectors for the more urbanized neighborhoods (Quakyi et al. 2000). However, the rapid extension of typical urban settings around Yaounde could, very probably, modify the malaria pattern in Simbock in the near future, decreasing the transmission and reducing density of species such as *An. moucheti* and *An. nili*.

Acknowledgments

The authors are grateful to the villagers in Simbock for their cooperation throughout the survey, and to Frédéric Simard for helpful suggestions. Part of this study was funded by a WHO-MIM grant No. 980101 to T.T., a WHO-RTG/TDR grant N° A00942 to C.A.-N. by the VihPal research program from the French Ministry of Research, and by the French Institut de Recherche pour le Développement.

References Cited

- Beier, J. C., P. V. Perkins, R. A. Wirtz, J. Koros, D. Diggs, T.P.I.I. Gargam, and D. K. Koech. 1988. Bloodmeal identification by direct enzyme-linked immunosorbent assay (ELISA) tested on *Anopheles* (Diptera: Culicidae) in Kenya. *J. Med. Entomol.* 25: 9-16.
- Burkot, T. R., J. L. Williams, and I. Schneider. 1984. Identification of *Plasmodium falciparum* infected mosquitoes by a double antibody enzyme-linked immunosorbent assay. *Am. J. Trop. Med. Hyg.* 33: 783-788.
- Carnevale, P., G. Le Goff, J. C. Toto, and V. Robert. 1992. *Anopheles nili* as the main vector of human malaria in villages of southern Cameroon. *Med. Vet. Entomol.* 6: 135-138.
- Coene, J. 1993. Malaria in urban and rural Kinshasa: the entomological input. *Med. Vet. Entomol.* 7: 127-137.
- Detinova, T. S. 1962. Age grouping methods in Diptera of medical importance, with special reference to some vectors of malaria. World Health Organization, Geneva.
- Dossou-Yovo, J., J.M.C. Doannio, S. Diarrassouba, and G. Chauvancy. 1998. Impact d'aménagements de rizières sur la transmission du paludisme dans la ville de Bouaké, Côte d'Ivoire. *Bull. Soc. Pathol. Exot.* 91: 327-333.
- Favia, G., A. Lanfrancotti, L. Spanos, I. Siden-Kiamos, and C. Louis. 2001. Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. *Insect Mol. Biol.* 10: 19-23.
- Fondjo, E., V. Robert, G. Le Goff, J. C. Toto, and P. Carnevale. 1992. Le paludisme urbain à Yaoundé (Cameroun). 2- Etude entomologique dans deux quartiers peu urbanisés. *Bull. Soc. Pathol. Ex.* 85: 57-63.
- Fontenille, D., and L. Lochouarn. 1999. The complexity of the malaria vectorial system in Africa. *Parassitologia* 41: 267-271.
- Fontenille, D., S. Wanji, R. Djouaka, and H. P. Awono-Ambene. 2000. *Anopheles hancocki*, vecteur secondaire du paludisme au Cameroun. *Bull. Liais. Doc. OCEAC* 33: 23-26.
- Fontenille, D., J. Y. Meunier, C. Antonio Nkondjio, and T. Tchuinkam. 2001. Use of circumsporozoite protein enzyme-linked immunosorbent assay compared with microscopic examination of salivary glands for calculation of malaria infectivity rates in mosquitoes (Diptera: Culicidae) from Cameroon. *J. Med. Entomol.* 38: 451-454.
- Gillies, M. T., and B. De Meillon. 1968. The Anophelinae of Africa south of the Sahara, 2nd ed. South African Institute of Medical Research, Johannesburg.
- Gillies, M. T., and M. Coetzee. 1987. A supplement to the Anophelinae of Africa south of the Sahara. South African Institute of Medical Research, Johannesburg.
- Le Goff, G., J. C. Toto, I. Nzeyimana, L. C. Gouagna, and V. Robert. 1993. Les moustiques et la transmission du paludisme dans un village traditionnel du bloc forestier sud-Cameroun. *Bull. Liais. Doc. OCEAC* 26: 133-137.
- Lindsay, S. W., H. Campbell, J. H. Adiamah, A. M. Greenwood, J. E. Bangali, and B., M. Greenwood. 1990. Malaria in a peri-urban area of the Gambia. *Ann. Trop. Med. Parasitol.* 84: 553-562.
- Manga, L., V. Robert, J. Messi, M. Desfontaine, and P. Carnevale. 1992. Le paludisme urbain à Yaoundé, Cameroun. 1- Etude entomologique dans deux quartiers centraux. *Mém. Soc. R. Belge Entomol.* 35: 155-162.
- Manga, L., J. C. Toto, and P. Carnevale. 1995. Malaria vectors and transmission in an area deforested for a new international airport in southern Cameroon. *Ann. Soc. Belge Med. Trop.* 75: 43-49.
- Meunier, J. Y., I. Safeukui, D. Fontenille, and C. Boudin. 1999. Étude de la transmission du paludisme dans une future zone d'essai vaccinal en forêt équatoriale du sud Cameroun. *Bull. Soc. Pathol. Exot.* 92: 309-312.
- Mouchet, J., and J. Gariou. 1966. *Anopheles moucheti* au Cameroun. *Cah. ORSTOM Ser. Entomol. Med.* 4: 71-81.
- Nimpaye, H., M. Van Der Kolk, D. Fontenille, and C. Boudin. 2001. Le paludisme urbain à Yaoundé (Cameroun) en 2000. Etude entomologique dans le quartier central "Dakar." *Bull. Liais. Doc. OCEAC* 34: 11-14.
- Njan Nloga, A., V. Robert, J. C. Toto, and P. Carnevale. 1993. *Anopheles moucheti*, vecteur principal du paludisme au sud-Cameroun. *Bull. Liais. Doc. OCEAC* 26: 63-67.
- Pajot, F. X., and L. G. Segers. 1964. Notes sur la biologie d'*Anopheles hargreavesi* Evans, 1927 et d'*Anopheles paludis* Theobald, 1900 dans le sud de la zone d'entraînement de Yaoundé (Cameroun), le long du fleuve Nyong. *Cah. ORSTOM. Ser. Entomol. Med. Parasitol.* 2: 3-15.
- Quakyi, I. A., R. G. Leke, R. Befidi-Mengue, M. Tsafack, D. Bomba-Nkolo, L. Manga, V. Tchinda, E. Njeungue, S.

- Kouontchou, J. Fogako, and others. 2000. The epidemiology of *Plasmodium falciparum* malaria in two Cameroonian villages: Simbok and Etoa. *Am. J. Trop. Med. Hyg.* 63: 222-230.
- Robert, V., P. Gazin, V. Ouédraogo, and P. Carnevale. 1986. Le paludisme urbain à Bobo-Dioulasso (Burkina Faso): Etude entomologique de la transmission. *Cah. ORSTOM Ser. Entomol. Med. Parasitol.* 24: 121-128.
- Robert, V., A. Van den broek, P. Stevens, R. Sloomweg, V. Petrarca, M. Coluzzi, G. Le Goff, M. A. Di Deco, and P. Carnevale. 1992. Mosquitoes and malaria transmission in irrigated rice-fields in the Benoue valley of northern Cameroon. *Acta Trop.* 52: 201-204.
- Scott, J. A., W. G. Brogdon, and F. H. Collins. 1993. Identification of single specimens of the *Anopheles gambiae* complex by polymerase chain reaction. *Am. J. Trop. Med. Hyg.* 49: 520-529.
- Trape, J. F., and A. Zoulani. 1987. Malaria and urbanization in central Africa: the example of Brazzaville. Part II: results of entomological surveys and epidemiological analysis. *Trans. R. Soc. Trop. Med. Hyg.* 81: 10-18.
- Wirtz, R. A., F. Zavala, Y. Charoenvit, G. H. Campbell, T. R. Burkot, J. Schneider, K. M. Esser, R. L. Beaudoin, and R. G. Andre. 1987. Comparative testing of monoclonal antibodies against *Plasmodium falciparum* sporozoites for ELISA development. *Bull. W.H.O.* 65: 39-45.
- Wondji, C., F. Simard, and D. Fontenille. 2002. Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of *Anopheles gambiae* in an area of sympatry. *Insect Mol. Biol.* (in press).

Received for publication 17 July 2001; accepted 31 October 2001.

ANNEXE 3

Protocoles détaillés des différentes PCR

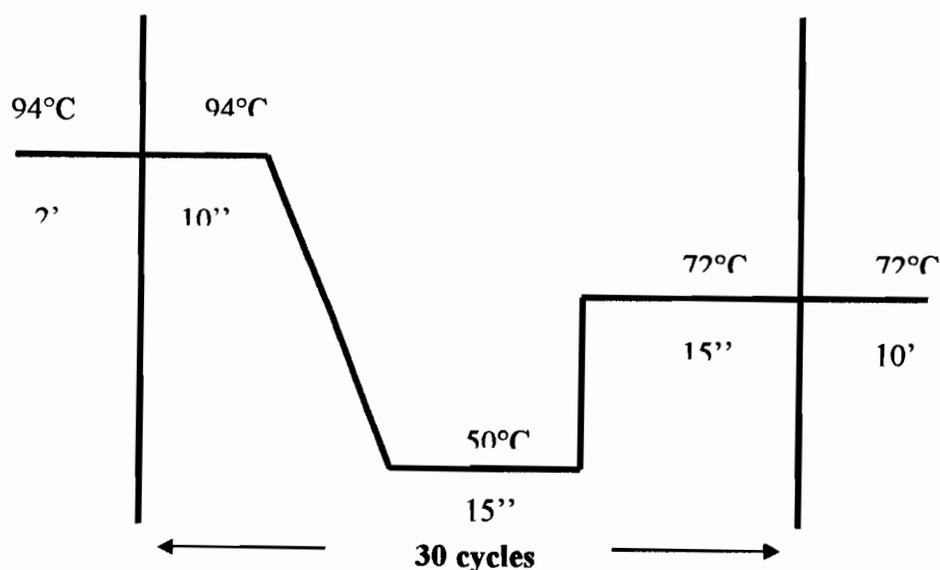
Protocole de la PCR complexe *An. gambiae*

1- Préparation du milieu réactionnel de 12,5 μ l de volume en plaçant les réactifs suivants dans un tube 1,5ml.

Tampon Taq 10X (contenant 15mM de MgCl ₂)	1,25 μ l
DNTP (25mM chacun)	0,1 μ l
Amorce AG (100ng/ μ l)	0,5 μ l
Amorce AR (100ng/ μ l)	0,5 μ l
Amorce ML (100ng/ μ l)	0,5 μ l
Amorce UN (100ng/ μ l)	0,5 μ l
Taq Polymérase (5U/ μ l)	0,063 μ l
Eau stérile	8,087 μ l

Le mélange est homogénéisé en pipetant plusieurs fois

- 2- Distribution de 11,5 μ l de la solution du milieu réactionnel dans chaque tube de PCR de 0,2ml fixé sur une plaque de PCR.
- 3- 1 μ l d'ADN de chaque spécimen à identifier est introduit dans le tube de PCR correspondant.
- 4- La PCR se déroule selon le programme suivant :



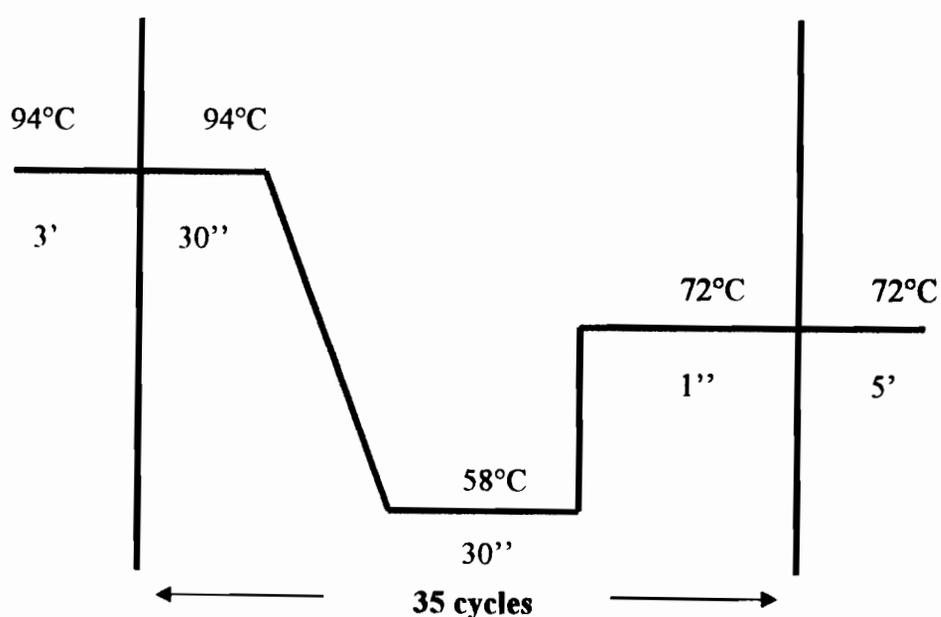
Déroulement de la PCR-RFLP pour la détermination des formes moléculaires d'*An. gambiae* selon Favia et al. (1997)

1- Préparation du milieu réactionnel de 25 μ l de volume en plaçant les réactifs suivants dans un tube 1,5ml.

Tampon Taq 10X (contenant 15mM de MgCl ₂)	2,5 μ l
DNTP (25mM chacun)	0,2 μ l
Amorce A.0 (100ng/ μ l)	2,5 μ l
Amorce AR (100ng/ μ l)	2,5 μ l
Taq Polymérase (5U/ μ l)	0,2 μ l
Eau stérile	q.s.p 25 μ l

Le mélange est homogénéisé en pipetant plusieurs fois

- 2- Distribution de 21 μ l de la solution du milieu réactionnel dans chaque tube de PCR de 0,2ml fixé sur une plaque de PCR.
- 3- 4 μ l d'ADN de chaque spécimen à identifier est introduit dans le tube de PCR correspondant.
- 4- La PCR se déroule selon le programme suivant :



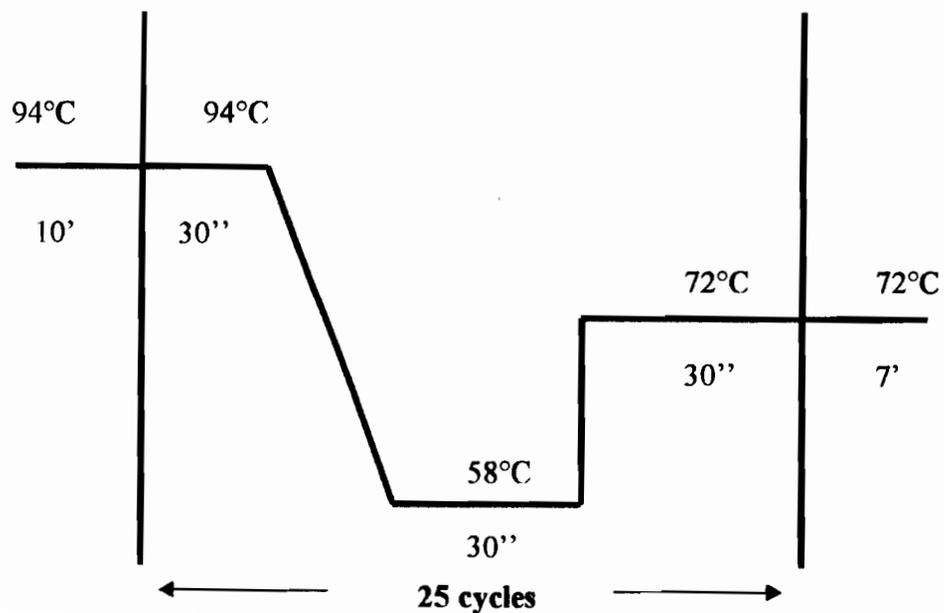
**Déroulement de la PCR pour la détermination des formes moléculaires d'*An. gambiae*
selon Favia et al. (2001)**

1- Préparation du milieu réactionnel de 12,5µl de volume en plaçant les réactifs suivants dans un tube 1,5ml.

Tampon Taq 10X	1,25 µl
(contenant 15mM de MgCl ₂)	
DNTP (25mM chacun)	0,1µl
Amorce R3 (100ng/µl)	1,55µl
Amorce R5 (100ng/µl)	1,55µl
Amorce Mopint (100ng/µl)	1,2µl
Amorce BSint (100ng/µl)	0,7µl
Taq Polymérase (5U/µl)	1µl
MgCl ₂	0,25µl
Eau stérile	4,8µl

Le mélange est homogénéisé en pipetant plusieurs fois

- 2- Distribution de 11,5µl de la solution du milieu réactionnel dans chaque tube de PCR de 0,2ml fixé sur une plaque de PCR.
- 3- 1µl d'ADN de chaque spécimen à identifier est introduit dans le tube de PCR correspondant.
- 4- La PCR se déroule selon le programme suivant :



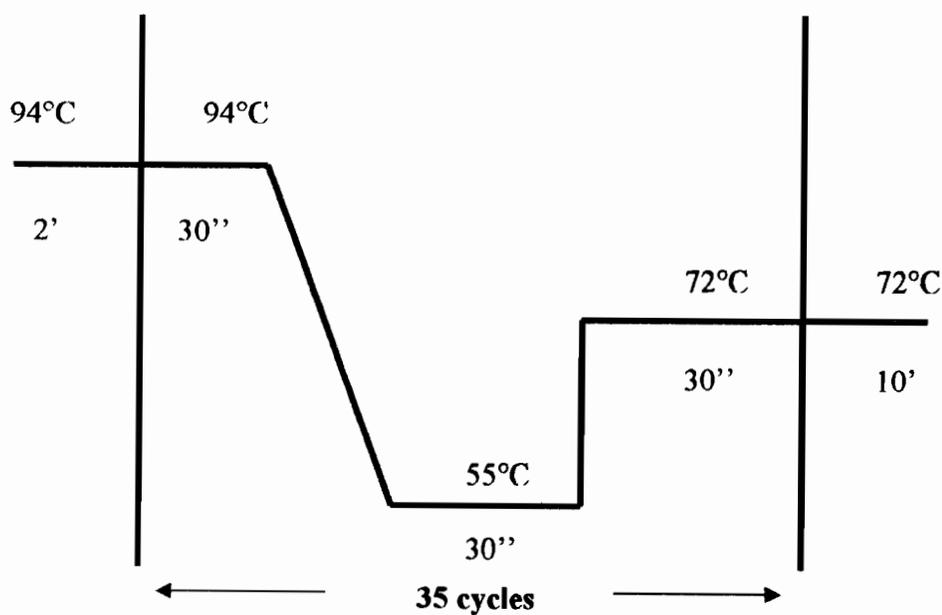
Génotypage des spécimens aux loci ADN microsatellites : déroulement de la PCR

- 1- Préparation du milieu réactionnel de 12,5 μ l de volume en plaçant les réactifs suivants dans un tube 1,5ml.

Tampon Taq 10X (contenant 15mM de MgCl ₂)	1,25 μ l
DNTP (25mM chacun)	0,1 μ l
Amorce R (5'→3') (100ng/ μ l)	0,15 μ l
Amorce L (3'→5')(100ng/ μ l)	0,15 μ l
Taq Polymérase (5U/ μ l)	0,063 μ l
MgCl ₂	0,25 μ l
Eau stérile	8,787 μ l

Le mélange est homogénéisé en pipetant plusieurs fois

- 2- Distribution de 11,5 μ l de la solution du milieu réactionnel dans chaque tube de PCR de 0,2ml fixé sur une plaque de PCR.
- 3- 1 μ l d'ADN de chaque spécimen à génotyper est introduit dans le tube de PCR correspondant.
- 4- La PCR se déroule selon le programme suivant :



ANNEXE 4

Préparation des tampons et solutions stock

– PREPARATION DES TAMPONS

1) Tampon DEB (DNA extraction buffer): Tampon de broyage des moustiques pour 100ml

Concentration finale	Quantité
0,08 M NaCl	1,6 ml (5 M NaCl) (peser 14,61 g dans 50 ml ED)
0,16 M Sucrose	5,48 g
0,06 M EDTA	12 ml (0,5 M EDTA pH 8)
0,10 M Tris-HCl	10 ml (1 M Tris-HCl, pH 9)
0,05% SDS	5 ml (10% SDS)
d'eau distillée stérile	q.s.p. (soit 71,4 ml)

Stérilisation à travers un filtre de 0,2µm de diamètre

Chauffage du tampon à 70°C pour inactiver les DNAses

- **Conservation du mélange à 4°C** au réfrigérateur pour des courtes périodes.
- **Conservation du mélange à -20°C** au congélateur pour de longues périodes
(faire des **aliquotes** : environ 2,5 ml, suffisant pour 25 moustiques).

Réchauffage du mélange avant utilisation pour resolubiliser le SDS.

AUTRES REACTIFS

- **Acétate de potassium (8,0 M)** à conserver à la température ambiante (15,7 g dans 20 ml eau distillée).
- **Ethanol (70%)** à conserver à -20°C.
- **Ethanol (100%)** à conserver à -20°C.
- Solution de **protéinase K** : le tube sigma contient 25 mg, mettre 5mg de protéinase K dans 1,5 ml. (On va l'utiliser à une concentration de **25 µg/ml** dans le tampon de broyage . Pour broyer 24 moustiques on prend 2,5 ml de tampon de broyage et on ajoute 19 µl de solution, soit 62 µg)

Ajout de la solution de protéinase K pour une concentration de **25 µg/ml** (=62 µg/2,5 ml) juste avant utilisation.

2) TBE 10X : tampon de stockage 10X Tris-borate-EDTA (pH \approx 8.2)

109,0 g	Tris base
55,6 g	acide borique
40,0 ml	0,5 M Na ₂ EDTA, pH 8,0

qsp H₂O pour 1,0 L (pas besoin d'ajuster le pH), mélanger.
conservation à température ambiante (T.A.)

3) Solutions stock**EDTA 0.5 M, pH 8.0 (ethylene diamine tetra-acetic acid)**

186,1 g	Na ₂ EDTA-2H ₂ O
---------	--

dissoudre dans 500 ml d'H₂O, mélanger sous agitateur magnétique

ajuster le pH à 8,0 avec du NaOH 10 M, ou en cristaux

compléter l'H₂O à 1 litre

conservation à T.A.

Bromure d'éthidium (10 mg/ml)

1 g	de bromure d'éthidium
100 ml	d'H ₂ O

mélanger sous agitateur magnétique pendant plusieurs heures

transférer dans une bouteille opaque ou recouvrir le récipient de papier aluminium

conservation à T.A.

NaCl 5 M

292,2 g	NaCl
800 ml	H ₂ O

ajuster l'H₂O à 1 litre

Tris 1 M

121,1 g Tris base
800 ml H₂O

ajuster le pH à la valeur voulue avec de l'HCl concentré
ajuster l'H₂O à 1 litre

MgCl₂ 1 M

203,3 g MgCl₂·6H₂O
800 ml H₂O

ajuster l'H₂O à 1 litre

APS 10%

peser 1 g d'APS
compléter avec de l'eau pour 10 ml.

Aliquoter en tubes de 1.5 ml. Conserver à 4°C.

CARNOY

pour conservation des ovaires des femelles de moustiques pour étude cytogénétique

Acide acétique glacial 1 volume (25 ml)
Ethanol absolu : 3 volumes (75 ml)

A conserver à -20° C

Réactifs et quantités pour ELISA - CSP

- **PBS** (Phosphate saline buffer)

Reconstituer le PBS en poudre Sigma :

9,7 g dans 1 L d'eau distillée

- **NP 40/BB** = Nonidet P40 pour une plaque 2 ml = 25 µl NP 40 + 2 ml BB (blocking buffer), agiter (préparer pour 5 plaques 10 ml + 125 µl NP 40)

- **BB** (blocking buffer) = Dans un litre PBS, ajouter :

- 1) 5 g de Caseine
- 2) 0,1 g de Thiomérosal
- 3) 0,04 g de Phenol red
- 4) 10 g de BSA

AGITER 2H

- **PBS/TWEEN 20** - 500 µl de Tween 20 dans 1l de PBS, agiter

Substrat de Peroxydase = Pour 3 plaques :

* 5 mg d'**Ortho-tolidine** dans 0,25 ml de **N,N-diméthyl formamide**

* 30 ml de **Tampon citrate**

* 12 µl de **H₂O₂** à 10% (ou 4 µl à 30% , ou 6 µl à 20%).

Tampon Citrate pH4

- Préparer le **Tampon Citrate pH4** : Pour 1 litre :

Acide citrique, 1 H₂O 11,77 g

Hydroxyde de Sodium 4,48 g

- Dissoudre la soude dans 200 ml d'eau bidistillée, puis l'acide citrique dans cette solution. Ajouter 400 ml d'eau bidistillée. Ajuster le pH à 4 avec de l'acide Chlorhydrique 1N. Compléter à 1 l avec de l'eau bidistillée.