View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Caltech Authors - Main

Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 11187–11191, December 1991
Biochemistry

Low-resolution genome map of the malaria mosquito *Anopheles gambiae*

(PCR/microamplification/*in situ* hybridization/dot blot/polytene chromosomes)

LIANGBIAO ZHENG*, ROBERT D. C. SAUNDERS†, DANIELA FORTINI‡, ALESSANDRA DELLA TORRE‡, MARIO COLUZZI‡, DAVID M. GLOVER†, AND FOTIS C. KAFATOS*§¶

*Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138; †Department of Biochemistry, University of Dundee, DD1 4HN Dundee, United Kingdom; ‡Istituto di Parassitologia, Università di Roma "La Sapienza," Piazzale Aldo Moro 5, 00185 Rome, Italy; and §Institute of Molecular Biology and Biotechnology, and Department of Biology, University of Crete, P.O. Box 1527, 711 10 Heraklion, Crete, Greece

Contributed by Fotis C. Kafatos, October 2, 1991

ABSTRACT We have microdissected divisions of the *Anopheles gambiae* polytene chromosomes, digested the DNAs with a restriction enzyme, and PCR-amplified the DNA fragments to generate a set of pooled probes, each corresponding to ≈2% of the mosquito genome. These divisional probes were shown to have high complexity. Except for those derived from near the centromeres, they hybridize specifically with their chromosomal sites of origin. Thus, they can be used to map cloned DNAs by a dot blot procedure, which is much more convenient than *in situ* hybridization to polytene chromosomes. We discuss additional potential uses of these probes, such as easier isolation of molecular markers and genes, including those that cross-hybridize with clones available from other insects. It is expected that the probes will substantially accelerate molecular genetic analysis of this most important malaria vector.

The recombinant DNA revolution has led to previously unimaginable progress in understanding basic processes of life and the molecular basis of disease. This progress has been achieved mostly with a few favorable and extensively studied model systems, such as some microorganisms, *Caenorhabditis elegans*, *Drosophila melanogaster*, and the mouse, as well as humans. The time has now come to undertake molecular genetic analysis of important but as yet poorly studied and little understood phenomena. A case in point is the intricate parasite–vector interactions that are crucial for transmission of many debilitating diseases, especially in the tropics. In recent years considerable progress has been made in molecular parasitology. However, very little is understood on the vector side: what, for example, makes the mosquito *Anopheles gambiae* (*sensu strictu*) such an effective vector for *Plasmodium*, in contrast to the relative recalcitrance of other mosquito species? This question is not merely academic, since >100 million new cases of malaria are thought to be contracted each year, and as many as 1–2 million deaths annually, mostly of children, are due to the disease (1).

Many disease vectors, including mosquitoes, are members of the order Diptera and thus relatives of the genetically best understood higher organism *D. melanogaster*. To what extent can the tools and knowledge accumulated in 80 years of intensive genetic and, most recently, molecular genetic research on the fruit fly be used to accelerate the study of the malaria mosquito? Here we report the development of a low-resolution physical map of the mosquito genome, using an approach developed in *Drosophila*, and discuss some of the numerous uses to which this map can be put to advance the molecular genetic analysis of mosquitoes.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

MATERIALS AND METHODS

Microdissection and PCR Amplification. Polytene chromosomes from nurse cells of half-gravid females were prepared from *An. gambiae* s.s. (karyotype: Xag; 2La; 2R+; 3L+; 3R+; M.C., unpublished results) and were microdissected into 54 divisions or subdivisions (see Table 1). Each dissected chromosomal segment was digested with proteinase K and the DNA was phenol/chloroform extracted, digested with *Sau*3AI, and ligated to an adaptor, as described (2). The mixture was then digested with *Bcl* I to prevent ligation between the adaptor molecules, and the DNA was amplified by the PCR in a final volume of 100 μl using one strand of the adaptor as primer (2). Conditions were as follows: 1 cycle of 94°C, 1.5 min; 37°C, 1 min; then 30 cycles of 94°C, 1 min; 55°C, 1 min; 72°C, 3 min; and finally 1 cycle of 55°C, 1 min; 72°C, 10 min. From this original stock solution, 1% samples were reamplified in 100 μl with the same primer.

Biotinylated Probe Preparation and *in Situ* Hybridization. Half-gravid females of *An. gambiae* G3 from the National Institutes of Health and *An. gambiae* s.s. (karyotype: X+; 2La; 2R+; 3L+; 3R+) from Rome were prepared and stored in Carnoy's solution at −20°C (3). Chromosomal squashes were prepared essentially as described (V. Krishnamoorty and F. Collins, personal communication), except that the chromosomes were denatured by placing the slides on top of a heating block at 100°C for 2 min before use.

Divisional probes were labeled in a random-priming reaction (4) with the Klenow fragment of *Escherichia coli* DNA polymerase I and biotinylated dUTP (biotin-16-dUTP; Boehringer Mannheim). The slides were hybridized with the probes in a final volume of 20 μl [containing 4× standard saline citrate (SSC), 1× Denhardt's solution, and 1% dextran sulfate] at 58°C overnight and were washed for two 20-min periods in 2× SSC at 54°C. Then at room temperature, the slides were washed for two 5-min periods in phosphate-buffered saline (PBS) and for one 2-min period in PBS with 0.1% Triton X-100 and were rinsed with PBS (2). The rinsed slides were stained with diaminobenzidine tetrahydrochloride with the ABC elite kit (Vector Laboratories). In most cases, the signals were distinct under phase-contrast microscopy. To maximize sensitivity and search for ectopic hybridization, sometimes the signals were further enhanced with the diaminobenzidine tetrahydrochloride enhancement kit (Amersham). Phase-contrast micrographs of the preparations were made using an Olympus BH2 microscope and Kodak Gold 100 color print film. The banding pattern was read following a map currently utilized at the Institute of Parasitology of the University of Rome "La Sapienza" for the

¶To whom reprint requests should be addressed at Harvard University.

Downloaded at California Institute of Technology on July 31, 2020

cytotaxonomy and cytogenetics of the *An. gambiae* complex (M.C., unpublished data).

Dot Blot Hybridization. About 10 μg of each reamplified divisional probe was denatured at 100°C and quickly cooled on ice. The samples were mixed with an equal volume of 20× SSC and loaded in a dot blot apparatus (BRL) onto an Amersham Hybond-N nylon membrane that had been prewetted with 10× SSC. The filters were then treated as recommended by the manufacturer. Hybridization with random-primed radioactive probes was carried out at 65°C in 10× Denhardt's solution/1× SSC/0.05% sodium pyrophosphate/0.1% $NaDodSO_4$ overnight. The filter was washed for two 15-min periods at 65°C with 1× SSC/0.05% sodium pyrophosphate/0.5% $NaDodSO_4$ and for two 15-min periods with 0.1× SSC/0.05% sodium pyphosphate/0.5% $NaDodSO_4$. The washed filter was exposed to Kodak XAR-5 film.

Southern Blot. DNA of recombinant λEMBL3 phages carrying DNA sequences of the vitellogenin locus was digested with *Xho* I and *Sal* I and then displayed in a 0.6% agarose gel. Southern blot hybridization (5) was carried out by standard procedures.

RESULTS

Preparation and Complexity of Divisional Probes. The *An. gambiae* polytene karyotype consists of five chromosome arms: one arm of the X chromosome and the right and left arms of chromosomes 2 and 3 (2R, 2L, 3R, and 3L, respectively). All the arms converge into a common chromocenter, consisting of underreplicated centromeric chromatin. A second heterochromatic arm of the X chromosome is also underreplicated and is not distinguished from the chromocenter. Each polytene arm is divided into numbered divisions: 1–6 for the X chromosome, 7–19 for 2R, 20–28 for 2L, 29–37 for 3R, and 38–46 for 3L. The divisions are not equal in length and each is partitioned into up to six numbered subdivisions (e.g., 22A-22F; ref. 6 and M.C., unpublished data).

Chromosomal segments were microdissected from nurse-cell polytene preparations. In most cases these segments corresponded to a numbered division, but the longest divisions were further split into two or three segments. Thus, a total of 54 chromosomal segments were obtained; they are listed consecutively in Table 1, which also indicates the proximal (to the centromere) end of each chromosomal arm. Assuming a genomic size of 0.25 pg (7) and roughly discounting a 20% fraction as nonbanded centromeric heterochromatin, as in *Drosophila* (8, 9), we calculate that each microdissected segment should encompass an average of 4 megabases of DNA sequence, slightly <2% of the genome.

The microdissected chromosomal DNA segments were digested with *Sau*3AI and ligated to a 20-base-pair double-stranded oligonucleotide adaptor with a 4-nucleotide 5′ overhang, as described (2). The adaptor-linked fragments were amplified by PCR, using one strand of the synthetic oligonucleotide as primer, thereby creating 54 "microamplified" DNA pools (divisional probes). Considering that *Sau*3AI has a 4-nucleotide recognition sequence and the average pool complexity is calculated as 4 megabases, we expect that each probe should consist of ≈16,000 distinct fragments, averaging 0.25 kilobase in length.

Table 1. Specificity of microamplified chromosomal probes of *An. gambiae*

Division	Signal	Specificity	Division	Signal	Specificity
Chromosome X			Chromosome 2L		
1	+	+	20 (proximal)	+	−
2	+	+	21A	+	−
3	+	+	21BC	+	+
4	+	+	21DF	+	+
5	+	+	22AC	+	+
6 (proximal)	+	−	22DF	+	+
Chromosome 2R			23	+	+
7	+	+	24	+	+
8AC	+	+	25	+	+
8DE	+	+	26	+	+
9	+	+	27	+	+
10	+	+	28	+	+
11	+	+	Chromosome 3R		
12AC	+	+	29		
12DE	+	+	30	−	
13AC	+	+	31	−	
13DE	+	+	32	−	
14AB	+	+	33	−	
14CE	+	+	34	−	
14AB	+	+	35	+	+
15	+	+	36	−	
16	+	+	37 (proximal)	+	−
17	+	+	Chromosome 3L		
18	+	+	38 (proximal)	+	−
19A	+	+	39	+	+
19BE (proximal)	+	−	40	+	+
			41	+	+
			42	+	+
			43	+	+
			44	+	+
			45	−	
			46	+	+

In situ polytene hybridization signal and specificity are indicated. +, Presence; −, absence. The divisions proximal to the centromeres are noted.

Downloaded at California Institute of Technology on July 31, 2020

As predicted, gel electrophoresis of typical probes showed a smear, without discrete bands; the majority of the fragments were in the range of 150–500 base pairs, but fragments several kilobases long were also present (Fig. 1*A*). To assess probe complexity, we digested (with *Xho* I and *Sal* I) DNAs of three overlapping genomic clones, encompassing 38 kilobases of the vitellogenin locus (a kind gift of P. Romans, Department of Zoology, University of Toronto). We blot-hybridized by the Southern procedure the resulting digest, using as probe the division 18 pool, which should contain the vitellogenin DNA region (ref. 3 and see below). Of the 20 distinct bands visible by ethidium bromide (Fig. 1*B*), 17 were evident in autoradiograms of the blot (Fig. 1*C*), albeit at intensities that were not always proportional to the DNA content; only three small bands totaling 1.2 kilobases appeared not to be represented in the probe. Thus, the tested probe appears to have a high complexity (see also ref. 2).

Specificity of the Probes: *In situ* Hybridization. The specificity and the complexity of the probes were confirmed by testing their hybridization to polytene chromosome preparations. Results are summarized in Table 1 and typical examples are shown in Fig. 2. Only 8 probes failed to give a detectable signal, while 46 probes gave strong signals. In particular, hybridization of 40 probes was detected only to their division of origin and was thus considered specific. With these probes, the polytene bands of the pertinent division hybridized with a signal strength roughly proportional to their intensity of Giemsa staining (DNA content; e.g., see the large photograph of division 7 in Fig. 2). Only in a few cases were there minor deviations at the borders because of slight inaccuracies in microdissection: examples are the underrepresentation of a terminal band of the segment (e.g., the proximal band of 14B in probe 14AB; Fig. 2) or the presence of a band just outside the intended segment (e.g., the first band of division 8A in probe 7; Fig. 2). Interestingly, the pericentromeric probes (divisions 6, 19BE, 20, 37, and 38) hybridized to the base of all the chromosomes, presumably because of the presence of repeats that are shared by the

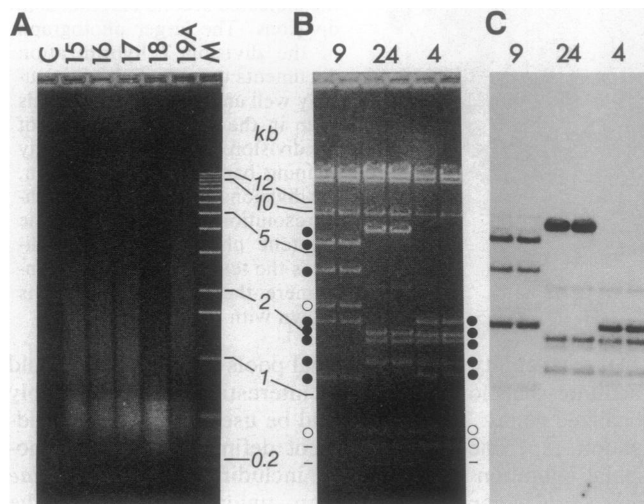

FIG. 1. Complexity of a PCR-amplified divisional pool. (*A*) PCR reamplified divisional pool DNAs (5 μl) were separated in a 1.5% agarose gel. Lane C shows a control PCR, in which no pool DNA was added, and lane M includes molecular size markers. Other lanes display the division 15, 16, 17, 18, and 19A pools, respectively. (*B*) Duplicate 5-μg samples of DNA from the phages λAgvg9, -24, and -4 were digested with *Xho* I and *Sal* I and separated in a 0.6% agarose gel. (*C*) The autoradiogram of *B* after hybridization with radioactively labeled division 18 probe. In *B*, solid dots indicate strongly hybridizing bands of phages 9 and 24 (on the left) or 4 (on the right), open dots indicate bands that are positive only after long exposure, and tick marks indicate bands that do not detectably hybridize.

corresponding regions of all the chromosomes. The pericentromeric repeats of 2R did not reach into 19A, but the repeats of 2L extended from divisions 20 to 21A (Table 1).

In summary, the set of effective probes covered all of the polytenized X chromosome, the 2R and 2L chromosomes, and all but division 45 of 3L chromosome. Of the nine divisions of 3R only two, 34 and 37, were represented in our collection of effective probes.

Specificity of the Probes: Hybridization with Defined Clones. Specificity was confirmed in a more sensitive manner, by testing how previously cloned sequences are distributed among the various probes. Dot blot filters were prepared, on which all the division-specific DNA pools were arranged in order. Characterized DNA clones were radioactively labeled and hybridized to these filters. Unique hybridization signals indeed were identified, corresponding to the cytogenetic origin of the cloned fragment (Fig. 3). Thus, a vitellogenin gene probe hybridized at high stringency uniquely to the division 18 pool, in agreement with the cytological location determined by *in situ* hybridization to polytene chromosomes (3). Similarly, two random cDNA clones, cDNA51 and cDNA9 (a kind gift from F. Collins, Malaria Branch, Centers for Disease Control, Atlanta), hybridized only to divisions 2 and 14CE, respectively, where they also map by polytene chromosome hybridization (F. Collins and V. Krishnamoorty, personal communication). We conclude that the specificity of the division-specific probes is adequate to permit the mapping of cloned DNA fragments by this blot hybridization technique, with a resolution of $\approx$2% of the genome.

DISCUSSION

We have developed a set of 46 probes covering $\approx$80% of the *An. gambiae* chromosomes, each probe corresponding to $\approx$2% of the genome. When probes for the remaining chromosomal segments are obtained and all the microamplified pools are converted into libraries, effectively, these probes and libraries will constitute a low-resolution physical map of the mosquito genome. The map should prove a valuable tool for rapidly accelerating molecular genetic analysis of the important anopheline disease vectors.

We have shown that the microamplified probes have reasonably high complexity and specificity. The complexity is ensured by the microamplification procedure itself (2), whereas the specificity is subject to features of the particular genome (prevalence and distribution of repetitive sequences, ectopic pairing of chromosomes, etc.). Forty of the available probes are highly specific at the level of *in situ* hybridization to chromosomes. This does not imply that they totally lack repetitive sequences or sequences derived from elsewhere in the genome but it does indicate that any such sequences are not quantitatively major. We envisage four major uses for these probes.

Mapping Cloned DNA Fragments. We have developed a dot blot mapping procedure, using these probes in a filter-immobilized array ("polytene chromosomes on a filter"). The technique was validated by mapping three cloned DNA fragments. It is likely to be widely useful, as it is much easier than *in situ* hybridization to polytene chromosomes (for example, it does not depend on a supply of half-gravid females) and as it permits variations in stringency of hybridization. It may be subject to occasional mistakes because of ectopic contaminants in the divisional pools (not detected thus far), and it is not applicable to cloned fragments containing a high proportion of repetitive DNA; however, even a DNA fragment derived from a pericentromeric region rich in repetitive DNA could be mapped by this procedure, provided it does not itself contain a high proportion of repetitive sequences.

Downloaded at California Institute of Technology on July 31, 2020

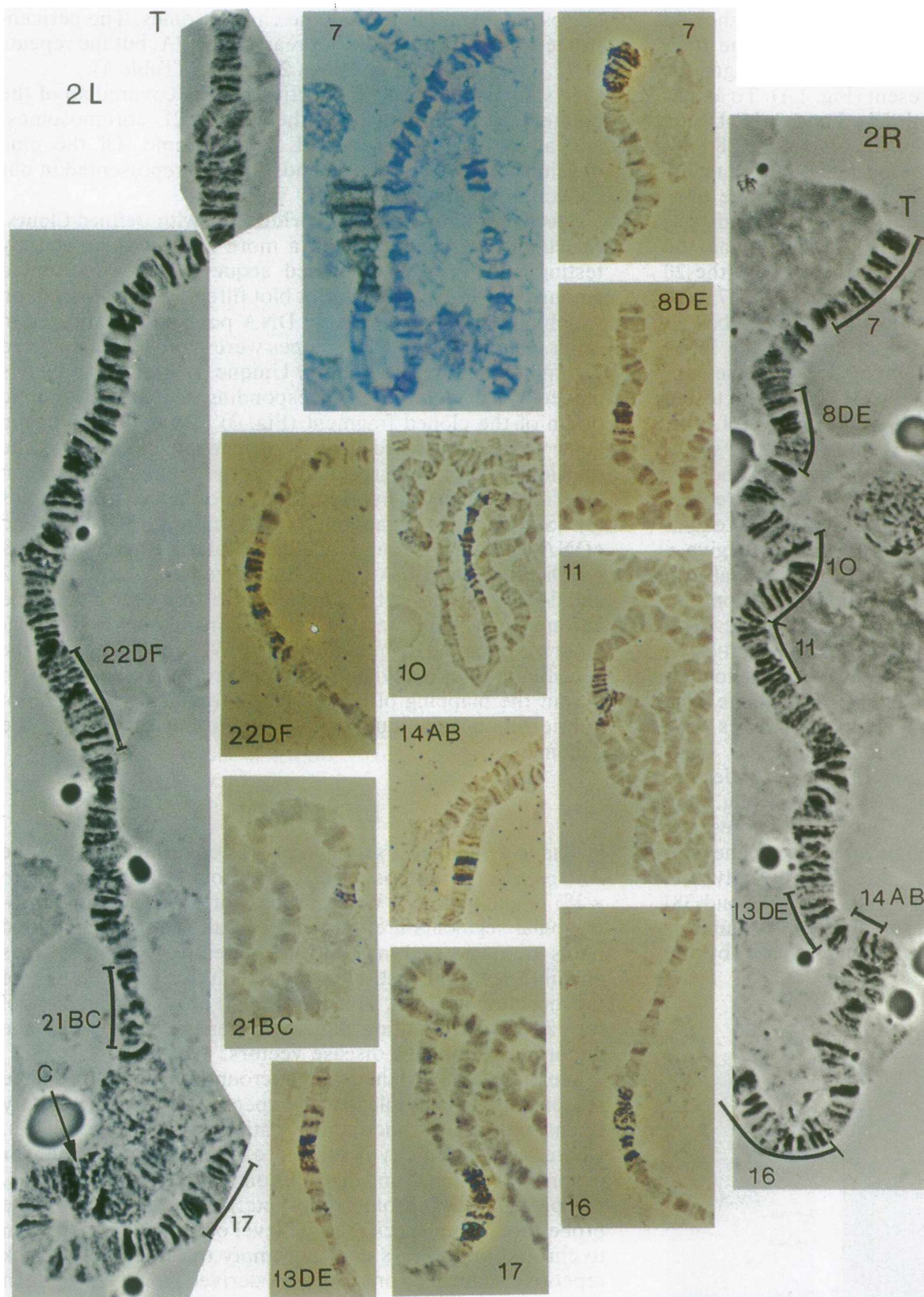

FIG. 2. Specificity of divisional pools. Each DNA pool was labeled with biotin and hybridized to nurse-cell polytene chromosomes by *in situ* hybridization. Typical *in situ* hybridization signals are shown with DNAs from the indicated second chromosome divisions. The larger photograph of the division 7 hybridization documents the specificity particularly well and shows weak signals even in the low-DNA regions of the division, between the strongly staining bands. For comparison, the divisions are marked in a higher-resolution photograph of the polytene chromosomes. T indicates the telomere and C the centromere; the base of the 2R arm is shown with the 2L arm.

Cloning Sequences Within a Target Chromosomal Region. These probes will be valuable for cloning DNA sequences from interesting chromosomal locations. Studies of polytene chromosomes have revealed conservation and differentiation (mostly paracentric inversions) among and within the six sibling species of the *An. gambiae* complex (10, 11). For example, the two salt water species, *Anopheles merus* and *Anopheles melas*, which have low vectorial capacity, are karyotypically distinct. Similarly, various ecotypes of *An. gambiae* s.s. and *Anopheles arabiensis* differ in vectorial capacity and in characteristic inversions. Microamplified probes could be used to clone the inversion breakpoints, helping to elucidate the mechanisms that generate the inversions (distinct transposable elements?). The inversions, which are expected to reduce recombination locally, may harbor genes involved in vector refractoriness, or protect polygenic systems adaptively advantageous to the mosquito (11). Availability of the divisional pools and libraries should facilitate the cloning of these interesting and presumably localized genes. The pools could be used for *in situ* hybridization experiments to document definitively the chromosomal evolution of *Anopheles*, including the *An. gambiae sensu latu* complex. In addition, unique probes from the corresponding libraries, both within and outside the inversions, could be sequenced in various species and ecotypes, to explore at a molecular level the population biology and evolution of this interesting genus.

Construction of a Molecular Genetic Map. The paucity of classical genetic markers for *An. gambiae* makes imperative the development of a molecular genetic map, using restriction fragment length polymorphisms or other DNA markers, such as those identified by microsatellites or the random amplified polymorphic DNA (RAPD) procedure (12). The microamplified probes should facilitate obtaining DNA markers evenly

Downloaded at California Institute of Technology on July 31, 2020

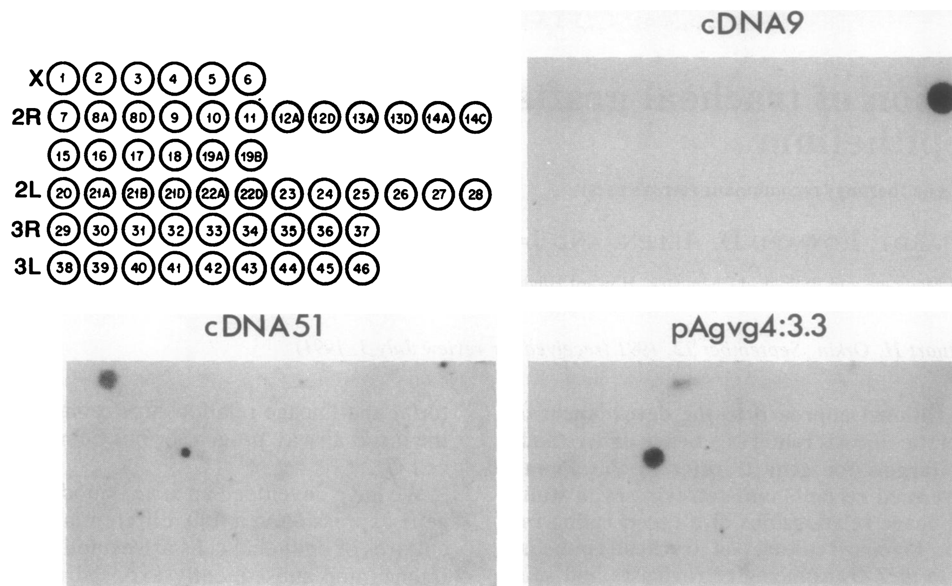

FIG. 3. Mapping of cloned DNA fragments with divisional pools. Approximately 10 μg of each divisional pool DNA was loaded in a dot blot apparatus onto a nylon membrane as diagrammed (*Upper Left*). The other three panels show hybridization of the indicated radioactively labeled clones, which permit their mapping to a single division (14CE for cDNA9, 2 for cDNA51, and 18 for a vitellogenin gene contained in subclone pAgvg4:3.3).

spaced throughout the genome or saturating regions of special interest (e.g., the inversion-rich 2R chromosome arm). For example, RAPD markers could be mapped conveniently on the dot blots, permitting the selection of a maximally informative panel of markers. Alternatively, a small sample of a genomic or cDNA library could be screened with the microamplified probe of a division of special interest, in the presence of competitor total mosquito DNA (to suppress hybridization of repeats); a small proportion of positive clones would be expected, which could be confirmed by hybridization to the dot blots and then tested as restriction fragment length polymorphism markers.

Cloning Genes by Homology. To advance rapidly the molecular genetics of *Anopheles*, it is hoped that many of the presently cloned *Drosophila* genes can be used to recover the corresponding mosquito loci. A limitation in attempting cloning by cross-hybridization across relatively distant species, such as mosquitoes and *Drosophila*, is that false-positive signals are frequent under the necessary low-stringency conditions of library screening. The libraries derived from divisional pools may prove useful in this respect. In a first step, Southern blots of total DNA from these libraries could be hybridized with the *Drosophila* probe to determine where in the genome the gene of interest may reside. After adjusting the stringency of hybridization as required, distinct bands should appear in the correct library, over the expected smear of falsely cross-hybridizing sequences in the same library and others. Only then would the more difficult second step be undertaken, screening the appropriate divisional library with the same probe at the indicated stringency. The homologue should be easier to recover from a divisional rather than from a total genomic library, since it should be accompanied by 50 times fewer false-positive signals. Furthermore, the insert size of the microamplified clones should be small (average 250 base pairs), permitting direct sequencing and verification of each putative positive clone, without the time-consuming step (mapping and subcloning the cross-hybridizing region), which are required for genomic clones of standard longer length. We expect that this two-step strategy for cloning mosquito genes by homology will be generally more effective than the notoriously unreliable procedure of screening standard libraries at low stringency in a single step.

We are grateful to Drs. Douglas Seely, Louis Miller, Patricia Romans, and Frank Collins for generous gifts of half-gravid females and cloned DNA fragments and for useful suggestions. We thank Dr. Jeff Powell for helpful discussions, Bianca Klumpar and Marie Yuk-See for help with the figures, and Esther Fenerjian and Zoë M. Kafatou for secretarial assistance. This work has been supported by grants from the John D. and Catherine T. MacArthur Foundation to F.C.K., from the European Community (Stimulation Action Programme) to F.C.K. and D.M.G., and from the Rockefeller Foundation to M.C.

1. World Health Organization (1990) *World Health Stat. Q.* **43**, 68–78.
2. Saunders, R. D. C., Glover, D. M., Ashburner, M., Siden-Kiamos, I., Louis, C., Monastirioti, M., Savakis, C. & Kafatos, F. C. (1989) *Nucleic Acids Res.* **17**, 9027–9037.
3. Graziosi, C., Sakai, R. K., Romans, P., Miller, L. H. & Wellems, T. E. (1990) *J. Med. Entomol.* **27**, 905–912.
4. Feinberg, A. P. & Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6–13.
5. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
6. Coluzzi, M. & Sabatini, A. (1967) *Parassitologia* **9**, 73–88.
7. Besansky, N. J. (1990) *Mol. Cell. Biol.* **10**, 863–871.
8. Beermann, W. (1972) in *Results and Problems in Cell Differentiation*, ed. Beermann, W. (Springer, Berlin), Vol. 4, pp. 1–33.
9. Rudkin, G. T. (1969) *Genetics (Suppl.)* **61**, 227–238.
10. Coluzzi, M., Sabatini, A., Petrarca, V. & Deco, M. A. D. (1979) *Trans. R. Soc. Trop. Med. Hyg.* **73**, 483–497.
11. Coluzzi, M., Petrarca, V. & Deco, M. A. D. (1985) *Boll. Zool.* **52**, 45–63.
12. Williams, J. G. K., Kubelik, A. R., Livak, K. J., Rafalski, J. A. & Tingey, S. V. (1990) *Nucleic Acids Res.* **18**, 6531–6535.

Downloaded at California Institute of Technology on July 31, 2020