Technical Note

A Gene-Based Method for Cytogenetic Mapping of Repeat-Rich Mosquito Genomes

Reem A. Masri [1], Dmitriy A. Karagodin [2], Atashi Sharma [3] and Maria V. Sharakhova [1,2,*]

1   Department of Entomology and the Fralin Life Sciences Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; reemm7@vt.edu
2   Laboratory of Evolutionary Genomics of Insects, The Federal Research Center Institute of Cytology and Genetics, Siberian Branch of the Russian Academy of Sciences, 630090 Novosibirsk, Russia; karagodin@bionet.nsc.ru
3   Department of Biochemistry and the Fralin Life Sciences Institute, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA; atashi04@vt.edu
*   Correspondence: msharakh@vt.edu

Simple Summary: Mosquitoes from the family Culicidae are vectors of numerous human diseases such as dengue, Zika, and West Nile fevers, malaria, and lymphatic filariasis. The availability of high-quality genome assemblies for mosquitoes assists in identifying genes responsible for important epidemiological traits such as vector competence, insecticide resistance, and mosquito behavior and stimulates a better understanding of the genetic structure of mosquito populations. Thus, the development of genomic resources and tools is a necessary step for designing and implementing novel genome-based approaches to control mosquitoes. The emphasis of this study is the further development and optimization of a physical genome mapping approach. Physical mapping is placing genomic scaffolds to the chromosomes based on hybridization of the specific probes. This task is difficult for mosquitoes with large genome sizes which are enriched with repetitive DNA sequences and can be potentially misassembled. Here, we describe in detail a simple and robust technique for the physical mapping of such mosquito genomes. This method can be further used for physical genome mapping in other mosquitoes and insects.

Abstract: Long-read sequencing technologies have opened up new avenues of research on the mosquito genome biology, enabling scientists to better understand the remarkable abilities of vectors for transmitting pathogens. Although new genome mapping technologies such as Hi-C scaffolding and optical mapping may significantly improve the quality of genomes, only cytogenetic mapping, with the help of fluorescence in situ hybridization (FISH), connects genomic scaffolds to a particular chromosome and chromosome band. This mapping approach is important for creating and validating chromosome-scale genome assemblies for mosquitoes with repeat-rich genomes, which can potentially be misassembled. In this study, we describe a new gene-based physical mapping approach that was optimized using the newly assembled *Aedes albopictus* genome, which is enriched with transposable elements. To avoid amplification of the repetitive DNA, 15 protein-coding gene transcripts were used for the probe design. Instead of using genomic DNA, complementary DNA was utilized as a template for development of the PCR-amplified probes for FISH. All probes were successfully amplified and mapped to specific chromosome bands. The genome-unique probes allowed to perform unambiguous mapping of genomic scaffolds to chromosome regions. The method described in detail here can be used for physical genome mapping in other insects.

Keywords: genome; physical mapping; DNA hybridization

1. Introduction

Mosquito genomics is rapidly developing as a discipline that can answer a large spectrum of biological questions. However, its major focus is assisting in the development

of new strategies for mosquito control [1]. Mosquito genomics is employed to address new threats in the spread of infectious diseases such as malaria, dengue, Zika fever, chikungunya, and others [2,3]. With the help of novel and affordable sequencing technologies such as long-read technologies from Pacific Biosciences (PacBio) [4] and Oxford Nanopore Technology (ONT) [5,6], the number of mosquito genomes available through public databases is rapidly increasing. Nevertheless, a high number of mosquito genomes are still far from complete and are represented by multiple relatively short scaffolds with low N50 statistics; only a few mosquito genomes are assembled at the chromosome-scale level. Most of the mosquitoes with chromosome-level assemblies belong to malaria vectors from the subfamily Anophelinae: *Anopheles gambiae* [7–9], *An. albimanus* [10–12], *An. atroparvus* [12–14], *An. stephensi* [12,14], *An. funestus* [14,15], *An. merus* [12], *An. arabiensis* [16], and *An. coluzzii* [16]. Only one chromosome-scale assembly has been developed for Culicinae mosquitoes, the major vector of arboviruses, *Aedes aegypti* [17]. Recently, a significantly improved version of the Asian tiger mosquito *Ae. albopictus* genome became available [18]. This effort placed 57% of the genome to the chromosomes. However, chromosome-level genome assembly for this species is still missing.

There are two major components that affect the quality of the genome: heterozygosity and repetitive DNA sequences. The first problem, heterozygosity, can be avoided by using inbred homozygous individuals [17] or by sequencing a single mosquito genome [19]. The second problem, repetitive DNA sequences, leads to algorithmically compressed sequences [20]. Repetitive DNA sequences are a significant impediment for genome quality in mosquitoes from the Culicinae subfamily [21,22]. Unlike Anophelinae mosquito genomes, Culicinae genomes are extremely enriched by different classes of transposable elements which have a tendency to localize in gene introns and spread throughout the genome [23,24]. This problem can be resolved through additional direct mapping of the genomic scaffolds to the chromosomes. Although new physical mapping technologies such as optical mapping [25] or Hi-C scaffolding [26,27] could significantly increase contiguity of the genomic scaffolds, only cytogenetic mapping based on fluorescence in situ hybridization (FISH) provides physical evidence of the scaffold location in particular regions of the chromosomes [28]. This method allows one to relate DNA probes of a known sequence to a particular band on the chromosomes. Thus, further development of the cytogenetic physical method is crucial.

The goal of the current study is to simplify and further optimize the cytogenetic mapping technique for highly repetitive mosquito genomes. We propose a new approach based on PCR amplification of DNA probes using complementary DNA (cDNA) that does not include repetitive DNA sequences from gene introns. This method was used for the development of a physical map for *Ae. albopictus* and is based on the hybridization of fifty cDNA fragments or gene exons from twenty-four scaffolds to mitotic chromosomes from imaginal discs, generating the first physical map of the *Ae. albopictus* genome [18]. The genome of this mosquito is enriched with transposable elements, making its size between 1.190 and 1.275 Gb long across populations, which is around five times larger than the genome of the major malaria vector *An. gambiae* [7]. Here, we describe this method in all required details so that it can be applied towards physical genome mapping in other mosquitoes or in other insect species of interest.

2. Materials and Methods

In this study, we used the Foshan strain, which was utilized for sequencing of the first and the last versions of the *Ae. albopictus* genome [18,29]. Mosquito eggs were hatched at 28 °C, and after several days, 2nd or 3rd instar larvae were transferred to 20 °C to obtain a high number of mitotic divisions in the imaginal discs of the 4th instar larvae.

At the time when the experiments were planned, the *Ae. albopictus* Aalbo_primary.1 genome assembly was still unavailable [19]. Therefore, we used a transcriptome from the C6/36 cell line of *Ae. albopictus* [30] to design target probes for FISH (Figure 1A). A total of 15 transcripts which represent protein-coding genes with unique locations in the

newly assembled genome, with sizes bigger than 5000 nucleotides (nt), were selected (Table 1). Primers were developed using the primer3plus program [31] with the following criteria: optimal primer size of 23 nt, melting temperature (Tm of 58 °C with maximum Tm difference of 5 °C, and GC content of 45%. The melting temperatures of the oligonucleotides were determined using the New England Biolab's Tm calculator (NEB, New England Biolabs, Ipswich, MA, USA). To visually see the FISH signal on mitotic chromosomes using microscopy, the final sizes of the PCR products were at least 3800 nt.

Table 1. Transcript locations in the *Aedes albopictus* genome assembly and chromosome map. Probes used for fluorescence in situ hybridization (FISH) are shown by transcript and gene IDs. Probe locations are indicated by scaffold number, chromosome band in the genome, or in the chromosomes.

Transcript ID	Scaffold ID/Number	Chromosome Band	Forward (F) and Reverse (R) Primers
18 S rDNA		1q22	F: ATGCAAAATGCAGGAACCTC R: GGTAATAGCAGCTGGGCTTG
XM_019696654	NW_021838798.1/6	1p25	F: TCGTTCGTGTAGATAAAGTCCAG R: ATGGTTATGAGGTTCCAACAACT
XM_019681895	NW_021838576.1/4	2p12	F: GAAGTAACGGGCTCAGTTCTGGTTTC R: CTTCGAGTAGTTGGACCAGTTCGAGA
XM_019677471	NW_021838576.1/4	2p25	F: AGCTCAACCAAAGGAAGGATTTA R: TGATTGTTCACCTTGTTTTCCAC
XM_019673127	NW_021838576.1/4	2p25	F: AAGTGTGCGGAAAATTGTTTACA R: TAATGGATTGAAGCTGCTTTTCG
XM_019675272	NW_021838153.1/2	2p32	F: TCCCTCTTTTATGAACAGCTGTT R: ACAAAACATTCATGCAGTTGTCA
XM_019703202	NW_021838687.1/5	2q34	F: AAACGACAAGAAATGTTCTGCAA R: TTGTGCCGCATTATCATTCATTT
XM_019682530	NW_021838153.1/2	2q44	F: AACAGAGCGGTATCTACAAAGAG R: GTAGAACACGAAGGCATTAGGTA
XM_019675517	NW_021838832.1/63	3p13	F: ACTTCGGTTATGGGTAAGGTTTT R: CAAAGACATGGGATTTTCTCGTC
XM_019697550	NW_021838154.1/20	3p14	F: GGAAGTTTTGTGTCGAAGGAAAA R: GTCGTCCAGATTGTACAGATCTT
XM_019682463	NW_021837045.1/1	3p32	F: GAGACCAACGCAGAGTACGTCTTCAC R: CGCATAGGCTCTGATGAACTTAGTCG
XR_002127435	NW_017857621.1/18	3q12	F: CCATTAAAAACGCCATCTAGCAA R: TATGAGTGTAGTGTGCTAGCAAG
XM_019682067	NW_021838465.1/3	3q23	F: AGGTACCGTACAAAAGAAGTGAA R: ACGGAACTAAGAAACAAAGTCCT
XM_019699815	NW_021837489.1/14	3q23	F: TTCGGTGGAAAAATCGTTTTGAA R: CGAGTTTCTCTATTGCCACTTTG
XM_019701500	NW_021838465.1/3	3q43	F: CGATGGACTGTTCTTCCAATCTA R: TTTGATTTGTGTTTGTCCCAGAC
XM_019701398	NW_021838465.1/3	3q44	F: TCCCGTTACTTCTACGAAATGAG R: CCATCTTCTGGTTTGCATAACAG

For cDNA synthesis, RNA was first extracted from ten 24 h old gravid female mosquito ovaries (Figure 1B). Female mosquitoes were placed on ice for immobilization; then, the ovaries were dissected in distilled water using a binocular microscope and immediately placed in TRIzol reagent (Thermo Fisher Scientific Inc., Waltham, MA, USA). Extraction was then performed using the Direct-zol RNA Miniprep kit (Zymoresearch, Irvine, CA, USA) following the manufacturer's protocol. cDNA was synthesized using ~200 ng RNA

and Superscript III Reverse Transcriptase following the manufacturer's protocol (Thermo Fisher Scientific Inc., Waltham, MA, USA). cDNA synthesis was performed using either oligo(dT), random hexamer, or gene-specific primers. However, after multiple attempts and troubleshooting steps, cDNA synthesized using oligo(dT) was determined to be the only successful method to give the desired band on agarose gels. RNA residuals were digested following the manufacturer's protocol by adding 1 μL of RNase H to each tube and incubating for 20 min at 37 °C. For later use, cDNA was diluted in Fisher DNAse/RNAse free water (Thermo Fisher Scientific Inc., Waltham, MA, USA) to 100 ng/μL and stored at −20 °C.
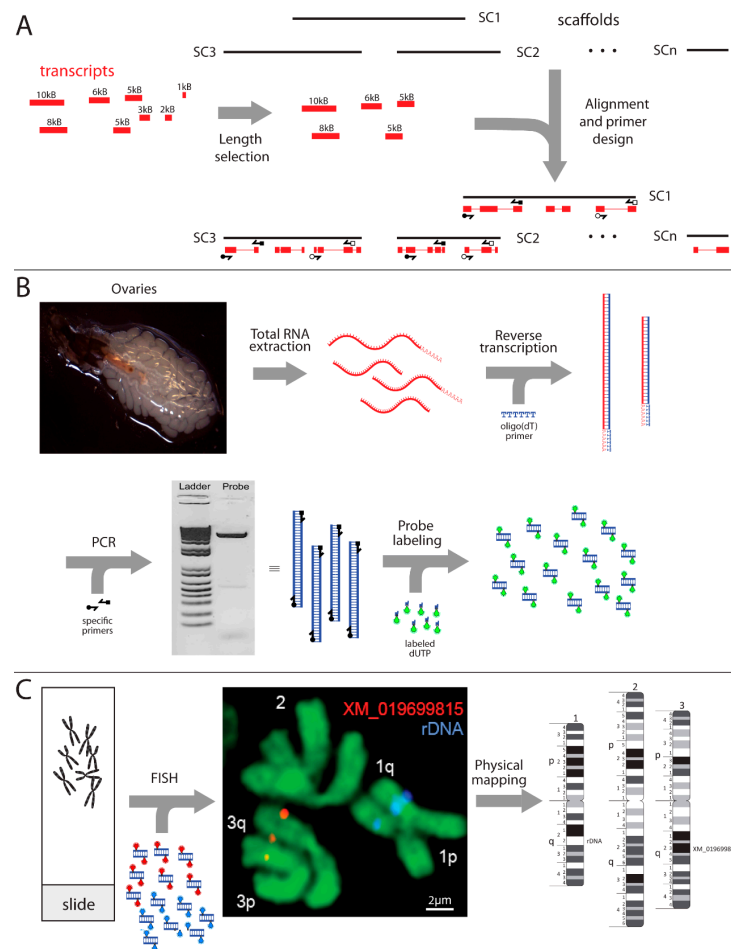
Figure 1. Genome mapping approach based on complementary DNA (cDNA). Procedure includes: (**A**) primer design using transcripts from the *Ae. albopictus* genome; (**B**) extraction of RNA from female ovaries, obtaining complementary DNA based on RNA, amplifying the probe by PCR and probe labeling; (**C**) obtaining the chromosome preparations from mosquito imaginal discs, fluorescent in situ hybridization (FISH), and physical mapping of the probes onto chromosome idiograms. Chromosomes and chromosome arms are indicated by the numbers 1, 2, and 3 and the letters p (petite) and q (long), respectively. A transcript number XM_019699815 is indicated by the same color as signal on FISH image. rDNA stands for ribosomal DNA.

To amplify DNA fragments, PCR was performed with the specific primers mentioned above (Table 1) using NEB Q5 high-fidelity DNA polymerase following the manufacturer's protocol (New England Biolabs, Ipswich, MA, USA). Briefly, 10 μL of 5× Q5 reaction buffer, 1 μL 10mM dNTPs, 2.5 μL of 10 μM forward primer, 2.5 μL of 10 μM reverse primer, 800 ng of cDNA as a DNA template, 0.5 μL 2 u/μL Q5 high-fidelity DNA polymerase, 10 μL 5× Q5 high GC enhancer, and nuclease-free water for a 50 μL reaction were used. PCR cycle conditions following the protocol were: initial denaturation, 98 °C for 10 min; denaturation

at 98 °C for 1 min; annealing at 50–72 °C for 30 s (optimal annealing temperature was 5 °C degrees below our lowest primer Tm); and 72 °C for 30 s per 1000 nt (2 min 30 s for an average 5000 nt probe). The optimal cycle number was 32. However, this number can be manipulated from 25 to 35 to obtain an optimal gel band. The final extension was at 72 °C for 2 min then held at 4 °C. Briefly, 5 µL of PCR product was run on 1% agarose to check for the desired DNA fragment (Figure 1B), and then, DNA was purified from the PCR mixture using Promega Wizard SV gel and a PCR clean-up system protocol (Promega Corporation, Madison, WI, USA).

For FISH, slides were prepared from imaginal discs of 4th instar larvae of *Ae. albopictus*. Before dissection, larvae were placed on ice for several minutes to immobilize them. Then, larvae were transferred to a slide along with a drop of a cold, hypotonic solution of 0.5% sodium citrate (Fisher scientific, Waltham, MA, USA) and dissected under an Olympus SZ microscope (Olympus America, Inc., Melville, NY, USA). The dissection procedure has been thoroughly explained in the previously published protocol [32]. Slides were analyzed under an Olympus CX41 microscope (Olympus America, Inc., Melville, NY, USA) at 400× magnification (40× objective and 10× eyepieces) (Figure 1C).

Purified DNA was labeled with Cy3- or Cy5-dUTP (Enzo Life Sciences Inc., Farmingdale, NY, USA) by nick translation using published protocols (Figure 1B), but probe incubation was performed for 1 h 45 min instead of 2 h 30 min [32]. The duration time was decreased to obtain the optimal fragment sizes between 300 and 500 nt. The increased time resulted in over-digestion of the probe, which reduced the effectiveness of FISH. The hybridization buffer which was used to dissolve the probes contained 50% formamide, 10% dextran sulfate sodium (DSS), and 0.1% Tween 20 in 2× Saline-Sodium Citrate, SSC (Thermo Fisher Scientific, Waltham, MA, USA). In situ hybridization (Figure 1C) was performed using a modified protocol [32]. Briefly, 1 µL of 10 mg/µL RNase A in 100 µL of 2× SSC solution was placed on each slide for 30 min. Slides were then placed in the following solutions: 2× SSC for 5 min at 37 °C, and pretreated in pepsin solution (50 mL sterile water, 50 µL 1M HCl, 0.1 mg/mL of pepsin) for 5 min at 37 °C. After that, slides were washed in 1× Phosphor Buffered Saline, PBS (Thermo Fisher Scientific, Waltham, MA, USA) and formaldehyde fixation solution (1.5 mL 37% formaldehyde to 50 mL 1× PBS) for 10 min each at room temperature. They were then rinsed in 1× PBS for 5 min and dehydrated using an ethanol series (70%, 80%, and 100%) for 5 min each and dried. Then, 250 ng of probe in hybridization buffer was added to the slide, covered with a 22 × 22-mm glass coverslip, and sealed with Elmer's rubber cement. The slides were then placed on a Thermobrite (Abbott Molecular Inc., Chicago, IL, USA) programmer at 73 °C for 5 min, followed by 37 °C overnight. The next day, slides were washed using the following solutions: in 1× SSC, pre-warmed at 60 °C, for 5 min and, at 37 °C, 4× SSC/NP40 (40 mL of water, 10 mL of 20× SSC, and 50 µL of NP40) for 10 min. After that, slides were rinsed in 1× PBS and stained in 50 µL of Oxazol Yellow, YOYO-1 stain (Thermo Fisher Scientific, Waltham, MA, USA) at a 1:1000 dilution in 1× PBS under parafilm in a dark box for 15 min. Slides were then rinsed in 1× PBS for 3 min and a drop of prolong gold antifade (Invitrogen Corporation, Carlsbad, CA, USA) was added to each slide. Slides were covered with glass coverslips and analyzed using a Zeiss LSM 880 Confocal Microscope (Carl Zeiss Microimaging, Inc., Thornwood, NY, USA) at 1000× magnification (100× objective and 10× eyepieces) (Figure 1C).

3. Results

In this study, we describe in detail a new physical approach for mapping repeat-rich mosquito genomes to their mitotic chromosomes. This method was used, for the first time, to physically map the *Ae. albopictus* genome [18]. Here, we describe this technique with all the methodological details required for further utilization of this method. Figure 1 summarizes the entire procedure. For the probe design, we utilized transcripts from the *Ae. albopictus* cell line genome (Figure 1A) [30]. Based on our genome analyses, these transcripts have unique locations in the genome and, thus, can be used for further physical

genome mapping (Table 1). To simplify primer design, the original sizes of the transcripts varied from 5012 to 10,567 bases. The primers were designed to obtain a final product of the DNA probe between 3800 and 5000 nt. The size of the DNA probe was chosen so as to obtain a clear signal on small mitotic chromosomes. Mosquito ovaries were utilized to obtain RNA for further development of a large amount of cDNA (Figure 1B). cDNA was produced from mRNA that mostly represents gene exons and does not contain gene introns, which are enriched with different classes of transposable elements in repeat-rich Culicinae genomes [23,33]. Repetitive DNA from gene introns produces multiple unspecific signals on the chromosomes that make it impossible to detect the correct position of a gene probe.

After the probes were designed, we used standard procedures [28] to obtain chromosome preparations from 4th instar imaginal discs of *Ae. albopictus* (Figure 1C) and proceeded with probe labeling and FISH (Figures 1C and 2). Similar to other mosquitoes, the *Ae. albopictus* karyotype includes three pairs of chromosomes. In correspondence with *Ae. aegypti* chromosomes [34], the shortest chromosome was described as chromosome 1, the longest chromosome was identified as chromosome 2, and the mid-length chromosome was designated as chromosome 3 [18]. Unlike the anophelines, the sex-determining chromosomes in all culicine mosquitoes, including *Ae. albopictus*, are homomorphic [35]. For the physical mapping of the DNA probes (Figures 1C and 3), we utilized previously developed chromosome idiograms [18]. Chromosomes of *Ae. albopictus* can easily be distinguished using chromosome and arm length differences [18]. Furthermore, *Ae. albopictus* chromosomes have a clear banding pattern that aids in physical mapping. In this study, 15 PCR-amplified probes and rDNA were mapped to the bands on an idiogram determined by FISH (Table 1; Figures 2 and 3). An 18 S ribosomal DNA probe that hybridized to the secondary constriction on 1q22 was used as an additional marker for chromosome 1 (Figures 1C and 3). The positions of the mapped transcripts were consistent with their positions in the *Ae. aegypti* genome, supporting the accuracy of the physical mapping.
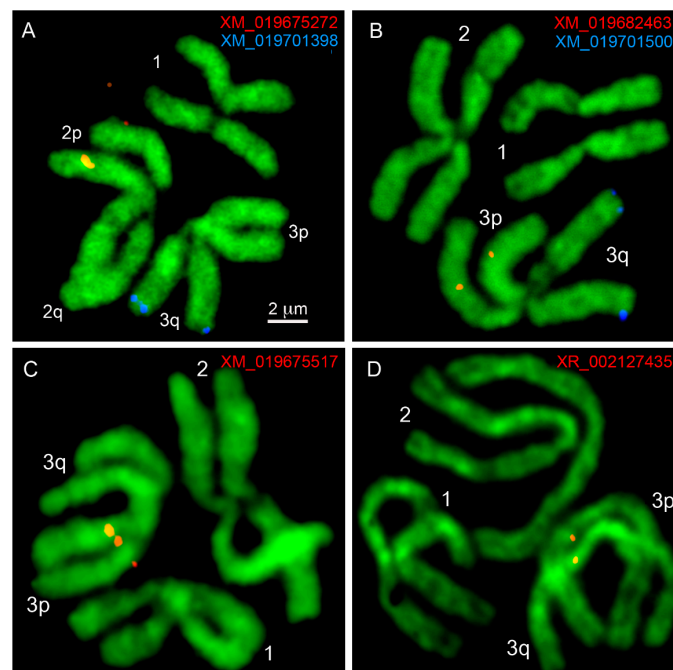
Figure 2. Examples of FISH in mitotic chromosomes of *Ae. albopictus.* The locations of transcripts XM_019675272 and XM_019701398 (**A**); XM_019682463 and XM_019701500 (**B**); XM_019675517 (**C**); and XR_002127435 (**D**) are shown. Chromosomes and chromosome arms are indicated by the numbers 1, 2, and 3 and the letters p and q, respectively. Transcripts numbers on the figure are shown with the same color as the probe signals on the chromosomes.
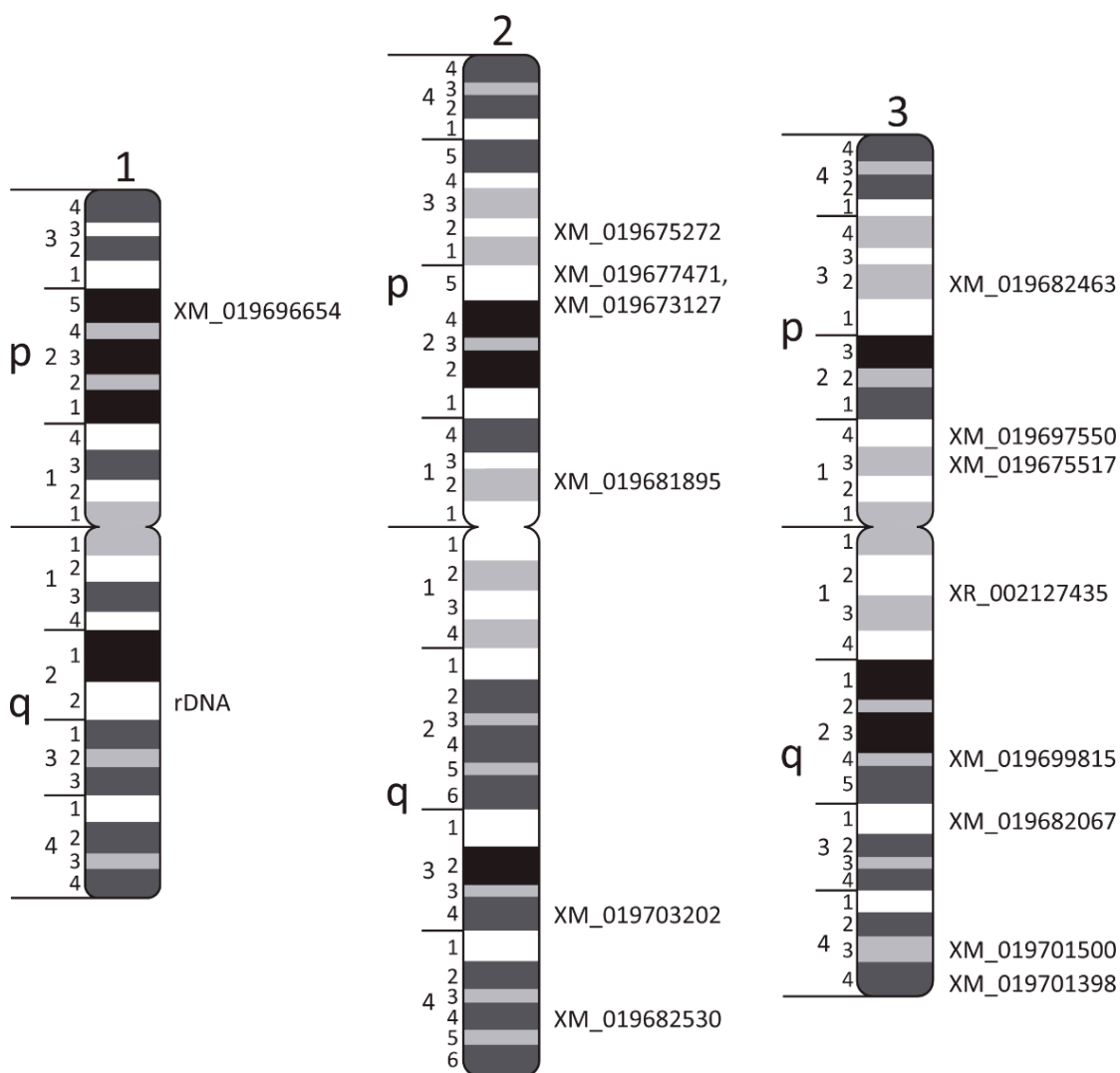
Figure 3. The locations of 15 DNA probes in the chromosome idiograms of *Ae. albopictus*. Chromosomes and chromosome arms are indicated by the numbers 1, 2, and 3 and the letters p and q, respectively. Chromosome divisions and subdivisions are shown on the left side of the idiograms. Transcripts are indicated on the right side of the idiograms; rDNA stands for ribosomal locus.

4. Discussion

Cytogenetic genome mapping could utilize different sources of DNA probes and chromosomes, depending on the size of the genomes and the quality of the chromosomes. This study focuses on the further development of cytogenetic physical mapping for application in highly repetitive genome research in mosquitoes such as *Ae. albopictus*. A detailed comparison of the new gene-based method and the conventional Bacterial Artificial Chromosome (BAC)-based approach for physical genome mapping is summarized in Table 2. Although BAC clones have been successfully used for physical genome mapping for *An. gambiae* [7], *Ae. aegypti* [17,23], and *Cx. quinquefasciatus* [36], utilizing BAC clones for highly repetitive genomes is challenging. First, BAC library development and BAC-end sequencing are expensive and unproductive in repeat-rich genomes because a large portion of the BAC-ends are represented by repeats located in multiple places in the genome. Second, blocking unspecific BAC clone hybridization to chromosomes by unlabeled repeats, so-called Cot-1 DNA, is also an inefficient and extremely time-consuming process as it requires extraction of large amounts of genomic DNA. Utilizing cDNA as a template for the probe development, which we proposed in this study, allows for specific hybridization of the probes to the chromosomes without additional DNA sequencing and repeat blocking

during FISH. Although both methods are labor-intensive, the gene-based method allows us to avoid long procedures such as extraction of multiple BAC clones. Using a PCR-based approach simplifies the probe design because it requires only a primer design and DNA amplification. The duration of the entire FISH procedure is reduced since a separate probe denaturation step is not required for the PCR-amplified probe.

Table 2. Comparison of the Bacterial Artificial Chromosome (BAC)-based approach and the gene-based method for physical genome mapping.

Comparison Parameters	BAC-Based Genome Mapping	Gene-Based Genome Mapping
BAC library development and BAC-end sequencing	Required	Not required
Primer design	Not required	Required
Blocking of unspecific hybridization with C_0t_1 DNA	Required	Not required
Multiple FISH signals are produced in the chromosomes	Often	Never
Additional DNA denaturation step in 70% formamide at 72 °C	Required	Not required
Expensive	Yes	No

Another challenge associated with physical mapping in mosquitoes with highly repetitive genomes is the poor quality of their polytene chromosomes. Thus, mitotic chromosomes became a major resource for cytogenetic mapping in Culicinae mosquitoes [35,36]. However, using mitotic chromosomes requires large DNA probes for FISH. For example, ~500 nt probes, which are clearly visible in FISH on large polytene chromosomes in malaria mosquito chromosomes [10,13], are not visible on small mitotic chromosomes. Thus, in our study, only probes with sizes larger than 3800 nt were utilized for FISH. Although an alternative method based on amplification of the gene exons for the FISH probes was also used for physical mapping of the *Ae. albopictus* genome [18], this method is not efficient for Culicinae mosquitoes because most of the exons in these mosquito genomes are smaller than 38 nt and would be invisible on mitotic chromosomes after FISH.

In addition to the cytogenetic approaches that we describe here, anchoring the genome assembly to the chromosomes can be performed by using different methods such as classical genetic linkage mapping [37–39] or by using modern technologies such as optical mapping [25] and Hi-C scaffolding [26,27]. Each of these mapping methods has its own advantages and limitations. For example, genetic linkage maps, which are based on recombination frequencies, offer the opportunity to connect genomic assemblies with epidemiologically important traits in mosquitoes, including their ability to transmit pathogens [17,40]. However, due to low recombination rates around the centromeres, genetic linkage maps have very low resolution around the centromeres [37,39]. New approaches such as optical mapping based on high-resolution restriction maps from single, stained molecules of DNA [25] or the Hi-C method, which uses the frequency of chromosome contacts to estimate the distance between DNA fragments [27], are highly efficient for linear ordering of genomic scaffolds at the molecular level. However, both methods are useless in connecting the genome assembly with particular chromosomes or chromosome structures such as telomeres, centromeres, and chromosome bands. Overall, we think that the best strategy for the development of high-quality genomes is to use a combination of different genome mapping techniques in one study. For example, one of the best chromosome-scale assemblies developed for mosquitoes, the *Ae. aegypti* genome, is based on a combination of modern technologies, Hi-C and optical mapping, with classical chromosome mapping technologies and linkage mapping [17].

5. Conclusions

High-resolution cytogenetic maps are fundamental for anchoring genome sequences onto chromosomes. Using complementary DNA as a probe template for FISH, we developed a new approach for the construction of physical maps and described this method in detail using the newly assembled *Ae. albopictus* genome as an example. This method is especially beneficial for the physical mapping of the highly repetitive Culicinae mosquito genome but can also be used for development of physical genome maps for other mosquitoes or insects to better aid in their genome assembly.

Author Contributions: Conceptualization, M.V.S. and D.A.K.; data curation, R.A.M. and A.S.; funding acquisition, M.V.S.; investigation, D.A.K., R.A.M. and A.S.; supervision, M.V.S.; writing—original draft, D.A.K., R.A.M. and M.V.S. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Transcript sequences are avlible at National Center for Biotechnology Information, www.ncbi.nlm.nih.gov.

Conflicts of Interest: The authors declare no competing interest.

References

1. Ruzzante, L.; Reijnders, M.; Waterhouse, R.M. Of Genes and Genomes: Mosquito Evolution and Diversity. *Trends Parasitol.* **2019**, *35*, 32–51. [CrossRef]
2. Lounibos, L.P.; Kramer, L.D. Invasiveness of *Aedes aegypti* and *Aedes albopictus* and Vectorial Capacity for Chikungunya Virus. *J. Infect. Dis.* **2016**, *214*, S453–S458. [CrossRef]
3. Vazeille, M.; Madec, Y.; Mousson, L.; Bellone, R.; Barré-Cardi, H.; Sousa, C.A.; Jiolle, D.; Yébakima, A.; de Lamballerie, X.; Failloux, A.B. Zika virus threshold determines transmission by European *Aedes albopictus* mosquitoes. *Emerg. Microbes Infect.* **2019**, *8*, 1668–1678. [CrossRef] [PubMed]
4. Berlin, K.; Koren, S.; Chin, C.S.; Drake, J.P.; Landolin, J.M.; Phillippy, A.M. Assembling large genomes with single-molecule sequencing and locality-sensitive hashing. *Nat. Biotechnol.* **2015**, *33*, 623–630. [CrossRef] [PubMed]
5. Goodwin, S.; Gurtowski, J.; Ethe-Sayers, S.; Deshpande, P.; Schatz, M.C.; McCombie, W.R. Oxford Nanopore sequencing, hybrid error correction, and de novo assembly of a eukaryotic genome. *Genome Res.* **2015**, *25*, 1750–1756. [CrossRef] [PubMed]
6. Miller, D.E.; Staber, C.; Zeitlinger, J.; Hawley, R.S. GENOME REPORT: Highly Contiguous Genome Assemblies of 15 *Drosophila* Species Generated Using Nanopore Sequencing. *G3 (Bethesda)* **2018**, *8*, 3131–3141. [CrossRef] [PubMed]
7. Holt, R.A.; Subramanian, G.M.; Halpern, A.; Sutton, G.G.; Charlab, R.; Nusskern, D.R.; Wincker, P.; Clark, A.G.; Ribeiro, J.M.; Wides, R.; et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **2002**, *298*, 129–149. [CrossRef]
8. Sharakhova, M.V.; Hammond, M.P.; Lobo, N.F.; Krzywinski, J.; Unger, M.F.; Hillenmeyer, M.E.; Bruggner, R.V.; Birney, E.; Collins, F.H. Update of the *Anopheles gambiae* PEST genome assembly. *Genome Biol.* **2007**, *8*, R5. [CrossRef] [PubMed]
9. George, P.; Sharakhova, M.V.; Sharakhov, I.V. High-resolution cytogenetic map for the African malaria vector *Anopheles gambiae*. *Insect Mol. Biol.* **2010**, *19*, 675–682. [CrossRef]
10. Artemov, G.N.; Peery, A.N.; Jiang, X.; Tu, Z.; Stegniy, V.N.; Sharakhova, M.V.; Sharakhov, I.V. The Physical Genome Mapping of *Anopheles albimanus* Corrected Scaffold Misassemblies and Identified Interarm Rearrangements in Genus *Anopheles*. *G3 (Bethesda)* **2017**, *7*, 155–164. [CrossRef]
11. Compton, A.; Liang, J.; Chen, C.; Lukyanchikova, V.; Qi, Y.; Potters, M.; Settlage, R.; Miller, D.; Deschamps, S.; Mao, C.; et al. The beginning of the end: A chromosomal assembly of the New World malaria mosquito ends with a novel telomere. *bioRxiv* **2020**. [CrossRef]
12. Lukyanchikova, V.; Nuriddinov, M.; Belokopytova, P.; Liang, J.; Reijnders, M.; Ruzzante, L.; Waterhouse, R.M.; Tu, Z.; Sharakhov, I.V.; Fishman, V. *Anopheles* mosquitoes revealed new principles of 3D genome organization in insects. *bioRxiv* **2020**. [CrossRef]
13. Artemov, G.N.; Bondarenko, S.M.; Naumenko, A.N.; Stegniy, V.N.; Sharakhova, M.V.; Sharakhov, I.V. Partial-arm translocations in evolution of malaria mosquitoes revealed by high-coverage physical mapping of the *Anopheles atroparvus* genome. *BMC Genom.* **2018**, *19*, 278. [CrossRef] [PubMed]

14. Waterhouse, R.M.; Aganezov, S.; Anselmetti, Y.; Lee, J.; Ruzzante, L.; Reijnders, M.; Feron, R.; Berard, S.; George, P.; Hahn, M.W.; et al. Evolutionary superscaffolding and chromosome anchoring to improve *Anopheles* genome assemblies. *BMC Biol.* **2020**, *18*, 1. [CrossRef] [PubMed]

15. Ghurye, J.; Koren, S.; Small, S.T.; Redmond, S.; Howell, P.; Phillippy, A.M.; Besansky, N.J. A chromosome-scale assembly of the major African malaria vector *Anopheles funestus*. *Gigascience* **2019**, *8*, giz063. [CrossRef] [PubMed]

16. Zamyatin, A.; Avdeev, P.; Liang, J.; Sharma, A.; Chen, C.; Lukyanchikova, V.; Alexeev, N.; Tu, Z.; Alekseyev, M.A.; Sharakhov, I. Chromosome-level genome assemblies of the malaria vectors *Anopheles coluzzii* and *Anopheles arabiensis*. *bioRxiv* **2020**. [CrossRef]

17. Matthews, B.J.; Dudchenko, O.; Kingan, S.B.; Koren, S.; Antoshechkin, I.; Crawford, J.E.; Glassford, W.J.; Herre, M.; Redmond, S.N.; Rose, N.H.; et al. Improved reference genome of *Aedes aegypti* informs arbovirus vector control. *Nature* **2018**, *563*, 501–507. [CrossRef]

18. Palatini, U.; Masri, R.A.; Cosme, L.V.; Koren, S.; Thibaud-Nissen, F.; Biedler, J.K.; Krsticevic, F.; Johnston, J.S.; Halbach, R.; Crawford, J.E.; et al. Improved reference genome of the arboviral vector *Aedes albopictus*. *Genome Biol.* **2020**, *21*, 215. [CrossRef]

19. Kingan, S.B.; Heaton, H.; Cudini, J.; Lambert, C.C.; Baybayan, P.; Galvin, B.D.; Durbin, R.; Korlach, J.; Lawniczak, M.K.N. A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing. *Genes (Basel)* **2019**, *10*, 62. [CrossRef]

20. Baptista, R.P.; Reis-Cunha, J.L.; DeBarry, J.D.; Chiari, E.; Kissinger, J.C.; Bartholomeu, D.C.; Macedo, A.M. Assembly of highly repetitive genomes using short reads: The genome of discrete typing unit III Trypanosoma cruzi strain 231. *Microb. Genom.* **2018**, *4*, e000156. [CrossRef]

21. Arensburger, P.; Megy, K.; Waterhouse, R.M.; Abrudan, J.; Amedeo, P.; Antelo, B.; Bartholomay, L.; Bidwell, S.; Caler, E.; Camara, F.; et al. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science* **2010**, *330*, 86–88. [CrossRef] [PubMed]

22. Nene, V.; Wortman, J.R.; Lawson, D.; Haas, B.; Kodira, C.; Tu, Z.J.; Loftus, B.; Xi, Z.; Megy, K.; Grabherr, M.; et al. Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **2007**, *316*, 1718–1723. [CrossRef]

23. Timoshevskiy, V.A.; Kinney, N.A.; de Bruyn, B.S.; Mao, C.; Tu, Z.; Severson, D.W.; Sharakhov, I.V.; Sharakhova, M.V. Genomic composition and evolution of *Aedes aegypti* chromosomes revealed by the analysis of physically mapped supercontigs. *BMC Biol.* **2014**, *12*, 27. [CrossRef] [PubMed]

24. Sharakhov, I.V.; Sharakhova, M.V. Heterochromatin, histone modifications, and nuclear architecture in disease vectors. *Curr. Opin. Insect Sci.* **2015**, *10*, 110–117. [CrossRef] [PubMed]

25. Deschamps, S.; Zhang, Y.; Llaca, V.; Ye, L.; Sanyal, A.; King, M.; May, G.; Lin, H. A chromosome-scale assembly of the sorghum genome using nanopore sequencing and optical mapping. *Nat. Commun.* **2018**, *9*, 4844. [CrossRef]

26. Dudchenko, O.; Batra, S.S.; Omer, A.D.; Nyquist, S.K.; Hoeger, M.; Durand, N.C.; Shamim, M.S.; Machol, I.; Lander, E.S.; Aiden, A.P.; et al. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science* **2017**, *356*, 92–95. [CrossRef]

27. Belton, J.M.; McCord, R.P.; Gibcus, J.H.; Naumova, N.; Zhan, Y.; Dekker, J. Hi-C: A comprehensive technique to capture the conformation of genomes. *Methods* **2012**, *58*, 268–276. [CrossRef]

28. Sharakhova, M.V.; Artemov, G.N.; Timoshevskiy, V.A.; Sharakhov, I.V. Physical Genome Mapping Using Fluorescence In Situ Hybridization with Mosquito Chromosomes. *Methods Mol. Biol.* **2019**, *1858*, 177–194. [CrossRef]

29. Chen, X.G.; Jiang, X.; Gu, J.; Xu, M.; Wu, Y.; Deng, Y.; Zhang, C.; Bonizzoni, M.; Dermauw, W.; Vontas, J.; et al. Genome sequence of the Asian Tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, E5907–E5915. [CrossRef]

30. Miller, J.R.; Koren, S.; Dilley, K.A.; Puri, V.; Brown, D.M.; Harkins, D.M.; Thibaud-Nissen, F.; Rosen, B.; Chen, X.G.; Tu, Z.; et al. Analysis of the *Aedes albopictus* C6/36 genome provides insight into cell line utility for viral propagation. *Gigascience* **2018**, *7*, 1–13. [CrossRef]

31. Untergasser, A.; Cutcutache, I.; Koressaar, T.; Ye, J.; Faircloth, B.C.; Remm, M.; Rozen, S.G. Primer3–new capabilities and interfaces. *Nucleic Acids Res.* **2012**, *40*, e115. [CrossRef]

32. Timoshevskiy, V.A.; Sharma, A.; Sharakhov, I.V.; Sharakhova, M.V. Fluorescent in situ hybridization on mitotic chromosomes of mosquitoes. *J. Vis. Exp.* **2012**, *67*, e4215. [CrossRef]

33. Biessmann, H.; Walter, M.F.; Dimitratos, S.; Woods, D. Isolation of cDNA clones encoding putative odourant binding proteins from the antennae of the malaria-transmitting mosquito, *Anopheles gambiae*. *Insect Mol. Biol.* **2002**, *11*, 123–132. [CrossRef]

34. Sharakhova, M.V.; Timoshevskiy, V.A.; Yang, F.; Demin, S.; Severson, D.W.; Sharakhov, I.V. Imaginal discs: A new source of chromosomes for genome mapping of the yellow fever mosquito *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **2011**, *5*, e1335. [CrossRef]

35. Rai, K.S. A Comparative Study of Mosquito Karyotypes1. *Ann. Entomol. Soc. Am.* **1963**, *56*, 160–170. [CrossRef]

36. Naumenko, A.N.; Timoshevskiy, V.A.; Kinney, N.A.; Kokhanenko, A.A.; deBruyn, B.S.; Lovin, D.D.; Stegniy, V.N.; Severson, D.W.; Sharakhov, I.V.; Sharakhova, M.V. Mitotic-chromosome-based physical mapping of the *Culex quinquefasciatus* genome. *PLoS ONE* **2015**, *10*, e0115737. [CrossRef]

37. Severson, D.W.; Meece, J.K.; Lovin, D.D.; Saha, G.; Morlais, I. Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. *Insect Mol. Biol.* **2002**, *11*, 371–378. [CrossRef] [PubMed]

38. Hickner, P.V.; Mori, A.; Chadee, D.D.; Severson, D.W. Composite linkage map and enhanced genome map for *Culex pipiens* complex mosquitoes. *J. Hered.* **2013**, *104*, 649–655. [CrossRef] [PubMed]

39. Juneja, P.; Osei-Poku, J.; Ho, Y.S.; Ariani, C.V.; Palmer, W.J.; Pain, A.; Jiggins, F.M. Assembly of the Genome of the Disease Vector *Aedes aegypti* onto a Genetic Linkage Map Allows Mapping of Genes Affecting Disease Transmission. *PLoS Negl. Trop. Dis.* **2014**, *8*, e2652. [CrossRef]
40. Timoshevskiy, V.A.; Severson, D.W.; Debruyn, B.S.; Black, W.C.; Sharakhov, I.V.; Sharakhova, M.V. An integrated linkage, chromosome, and genome map for the yellow fever mosquito *Aedes aegypti*. *PLoS Negl. Trop. Dis.* **2013**, *7*, e2052. [CrossRef]