Site-specific genetic engineering of the *Anopheles gambiae* Y chromosome

Federica Bernardini[a], Roberto Galizi[a], Miriam Menichelli[a], Philippos-Aris Papathanos[b], Vicky Dritsou[b], Eric Marois[c], Andrea Crisanti[a,b], and Nikolai Windbichler[a,1]

[a]Department of Life Sciences, South Kensington Campus, Imperial College London, London SW7 2AZ, United Kingdom; [b]Dipartimento di Medicina Sperimentale Via Gambuli, Centro di Genomica Funzionale, University of Perugia, 06132 Perugia, Italy; and [c]Institut de Biologie Moléculaire et Cellulaire, Institut National de la Santé et de la Recherche Médicale U963, Centre National de la Recherche Scientifique, Unité Propre de Recherche 9022, 67084 Strasbourg, France

Edited by Anthony A. James, University of California, Irvine, CA, and approved April 17, 2014 (received for review March 18, 2014)

Despite its function in sex determination and its role in driving genome evolution, the Y chromosome remains poorly understood in most species. Y chromosomes are gene-poor, repeat-rich and largely heterochromatic and therefore represent a difficult target for genetic engineering. The Y chromosome of the human malaria vector *Anopheles gambiae* appears to be involved in sex determination although very little is known about both its structure and function. Here, we characterize a transgenic strain of this mosquito species, obtained by transposon-mediated integration of a transgene construct onto the Y chromosome. Using meganuclease-induced homologous repair we introduce a site-specific recombination signal onto the Y chromosome and show that the resulting docking line can be used for secondary integration. To demonstrate its utility, we study the activity of a germ-line–specific promoter when located on the Y chromosome. We also show that Y-linked fluorescent transgenes allow automated sex separation of this important vector species, providing the means to generate large single-sex populations. Our findings will aid studies of sex chromosome function and enable the development of male-exclusive genetic traits for vector control.

biotechnology | SIT

Mosquito species of the *Anopheles gambiae* complex represent the principal vectors of human malaria, and they pose an enormous burden on global health and economies. Every year, 300–500 million people are infected by malaria and more than 1 million people die as a consequence of *Plasmodium* parasite infections (1). The malaria mosquito *A. gambiae* has two pairs of autosomes, termed 2 and 3, and a pair of heteromorphic sex chromosomes X and Y, XX in females and XY in males (2). Extensive nonpairing regions exist between the X and the degenerate Y chromosome. and evidence points to a factor located on the Y chromosome that primarily determines the sex in *Anopheles* (3). Current models suggest that the evolutionary differentiation of Y chromosomes begins with the acquisition of a male determining factor on a proto-Y chromosome (4, 5). This event is followed by a progressive suppression of recombination between the still largely homomorphic proto-sex chromosomes, a process attributed to the acquisition of sexually antagonistic mutations, which are beneficial to the heterogametic sex but detrimental to the homogametic sex (6–8). The lack of recombination, together with the male-limited transmission, leads to the degeneration of the Y chromosome, which involves accumulation of deleterious mutations, spread of transposable elements, and silencing of all or most of the genes present on the proto-Y (9–11). As a result, Y chromosomes of many species appear to be strongly heterochromatic and harbor only few genes often involved in male fertility (12–17). The accumulation of repetitive sequences, many of which are also present on other chromosomes, hampers the assembly of Y chromosome contigs following shotgun sequencing. Indeed, despite the completion of the *A. gambiae* genome project (18), and the knowledge that the primary signal is likely to be associated with the inheritance of the Y (3, 19), no assembly of the *Anopheles* Y chromosome has been achieved. At present, public databases host only a few hundred kilobases of *A. gambiae* sequences attributed to the Y, a chromosome that is estimated to comprise 10% of the genome and to be at least 20 Mb in size. None of these Y-specific scaffolds have been physically mapped, because the Y chromosome does not polytenize. The exploration of the Y chromosome will improve our understanding of the evolutionary forces involved in driving chromosome evolution and may enable the manipulation of the molecular pathways that control sex determination and sexual differentiation in mosquitoes. In a number of organisms, Y chromosome genes have been found to be essential for male fertility or sex determination. Recently, a number of excellent candidate genes potentially involved in these processes have been identified on the Y chromosome of anopheline mosquitoes (20, 21). Because interfering with male fertility is an essential part of vector control strategies such as the sterile insect technique, the identification of such genes is of particular interest to mosquito biologists. In this paper, we demonstrate the targeted molecular manipulation of the Y chromosome in *A. gambiae*, thus opening up a number of ways to explore one of the most fascinating of evolution's upshots and to harness this genetic tool for vector control.

Significance

Interfering with sex determination and male fertility are potentially powerful approaches for the genetic control of the human malaria vector *Anopheles gambiae*. Despite this fact, the male-specific Y chromosome of this mosquito has remained largely unexplored, because of its repetitive, heterochromatic structure. Little is known about its ability to support gene transcription in different tissues and during gametogenesis, yet this information is crucial for understanding the function of this chromosome. We show, using a combination of knock-in and site-specific genetic engineering steps, how transgenes can be specifically introduced onto the Y chromosome. The Y-linked strains we have created provide the means to generate large single-sex populations and to establish male-exclusive genetic traits for the control of this important vector species.

Author contributions: N.W. designed research; F.B., R.G., M.M., and E.M. performed research; P.-A.P. and N.W. contributed new reagents/analytic tools; F.B., P.-A.P., V.D., E.M., and N.W. analyzed data; and A.C. and N.W. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession nos. KJ608148–KJ608160).

[1]To whom correspondence should be addressed. E-mail: n.windbichler@imperial.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1404996111/-/DCSupplemental.

Downloaded by guest on January 14, 2020

Results

The Transgenic Line T4 Carries a Transgene That Segregates with the Y Chromosome. We injected wild-type *A. gambiae* embryos with the plasmid pHome-T together with in vitro-transcribed piggyBac transposase helper RNA (*Materials and Methods*) (22). The pHome-T plasmid contains piggyBac inverted repeats for transposase-mediated random integration and the 3xP3-GFP and Actin5C-RFP fluorescent markers. In addition, the construct contains the 18-bp I-SceI endonuclease recognition site, located within the GFP ORF. This site is not present in the *Anopheles* genome and can be used to specifically cut chromosomes carrying the integrated plasmid exclusively at this position. We obtained a number of G_0 mosquitoes expressing the fluorescent markers transiently, and they were separately crossed to wild-type mosquitoes. One male G_0 founder gave rise to 42 transgenic offspring, all of which emerged as males. When five of these transgenic males were crossed to wild-type females in single mating experiments, we again observed that all transgenic larvae were emerging as male adults (206 of 421), whereas the non-transgenic larvae (215 of 421) emerged as females. These results indicated that the transgene construct (Fig. 1) had potentially integrated within the Y chromosome. Expression of both fluorescent marker genes in this transgenic line, termed T4, was clearly visible (Fig. 2); however, the expression of RFP driven by the *Drosophila* Actin5C promoter was patchy and visible mainly in the periphery of the larval gut. The pattern of neuronal GFP fluorescence appeared to be normal, resembling that of autosomal insertions. We performed fluorescence in situ hybridization (FISH) by using the pHome-T plasmid as a probe against mitotic chromosome spreads generated from male larval imaginal discs of line T4. As a control, the multicopy ribosomal DNA (rDNA) locus on the X chromosome was used. In all cases, a single weak signal was obtained from the smaller sex chromosome (Fig. 2*B* and Fig. S1), supporting the notion that the transformation construct had integrated into the Y chromosome.

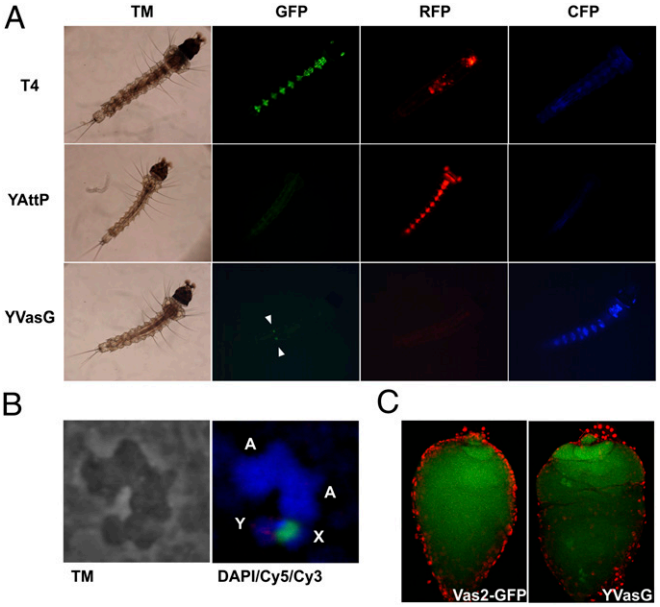

Fig. 2. Phenotype of Y-linked strains. (*A*) Expression of green (GFP), red (RFP), and blue (CFP) fluorescent markers in the tissues of L2 larvae shown as transmitted-light image (TM) in *Left*. The white arrowheads indicate GFP expression in the developing larval gonads. (*B*) FISH with probes designed against the X-linked rDNA (labeled with Cy5) and the pHome-T construct (labeled with Cy3) hybridized to the *A. gambiae* transgenic line. Mitotic chromosome slide preparations were prepared from imaginal discs of fourth instar larvae. X and Y indicate the sex chromosomes and A indicates autosomes. (*C*) Confocal analysis of GFP expression in dissected testes of transgenic adult males.

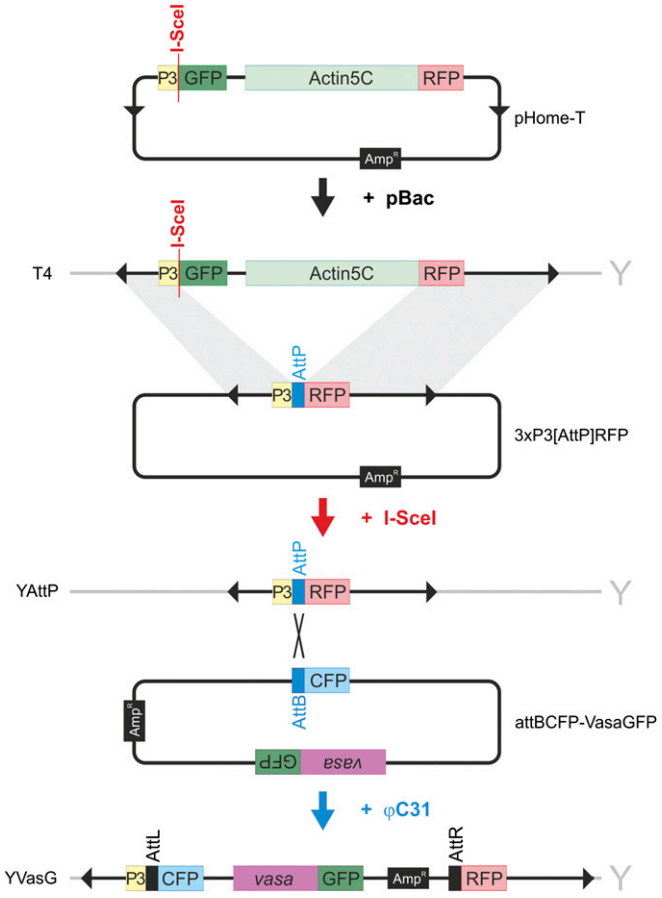

Fig. 1. Site-specific genetic engineering of the Y chromosome. Overview of constructs and transgenic lines generated in this study and the stepwise approach taken to modify the Y chromosome.

Characterization of the Genomic Neighborhood of the T4 Transgene. Expression of the fluorescent reporter transgenes suggested that the construct had inserted within a region of chromatin allowing transcriptional activity. We performed genomic mapping by inverse PCR using primers binding to the PiggyBac-inverted repeats to characterize the genomic region flanking the integrated construct. We obtained genomic sequences that aligned to AAAB01003622.1, a scaffold not assigned to a known chromosome but that was characterized as a scaffold containing fragments originating exclusively from male libraries (23). This scaffold has a length of only 1.7 kb, and we therefore generated a BAC library from genomic DNA of T4 males and used transgene-specific primers to identify a single positive clone of approximately 110 kb. This clone was sequenced and assembled into a total of 13 contigs. The sequence of all contigs was found to consist, in a large part, of short sequence repeats and sequence fragments matching known transposable elements. Using a cutoff of 100 aa as a minimum length, we selected 39 putative ORFs within these contigs to test for Y linkage. PCR performed on genomic DNA showed that 3 ORFs of the 39 were amplified from wild-type or T4 males but not from females (Table S1). We next used RNAseq data from adult males and females to identify novel transcribed regions (NTRs) mapping to the 13 contigs (Fig. S2). Of 20 identified NTRs, 12 show a male-specific expression pattern similar to recently identified Y-linked genes used as a control. We performed similarity searches with NTRs supported by putative

Downloaded by guest on January 14, 2020

ORFs, and although some may present previously unidentified Y linked genes, others match to known Y-linked repetitive elements (Dataset S1).

The Y-Linked Transgene Is Stable and Can Serve as a Marker for Automated Sexing. Recombination between the *A. gambiae* X and Y chromosomes has not been reported to occur; however, evidence is derived mainly from genetic experiments by using chromosomal translocations (23). The Y linkage of the transgene in line T4 allowed us to test experimentally whether exchange between these chromosomes occurs at a significant rate. These experiments also allowed us to test whether Y linkage of the T4 transgene was stable and whether recombination or remobilization of the transgene could break its linkage with the Y chromosome. A total of 16,750 T4 larvae were sorted in the Complex Object Parametric Analyzer and Sorter (COPAS) particle sorter according to their distinct intensities of green fluorescence, because the expression of GFP is expected to occur in male larvae only (Fig. 3*A*). The GFP-negative cloud corresponding to putative female larvae (6,415 individuals) was entered for a second run in the COPAS to check for the presence of contaminating GFP^+ larvae. No GFP-expressing male larvae were detected. The GFP-positive and -negative larvae were separately reared to adulthood. All GFP-positive adults emerged as males, whereas all GFP negatives were found to be females, thus indicating the utility of line T4 for automated sex separation. This COPAS sorting experiment was repeated multiple times at intervals of several months, encompassing at least 12 generations. Sorting was always perfectly accurate, indicating that the transgene is stable. No putative X/Y recombination events were detected in the T4 line, and we concluded that recombination was either very rare or that the T4 insertion site is located in a region of the Y chromosome that does not readily recombine with the X. Thus, automated sorting of the T4 line represents a fast and reliable method to generate pure populations of either transgenic T4 males or nontransgenic, virgin females.

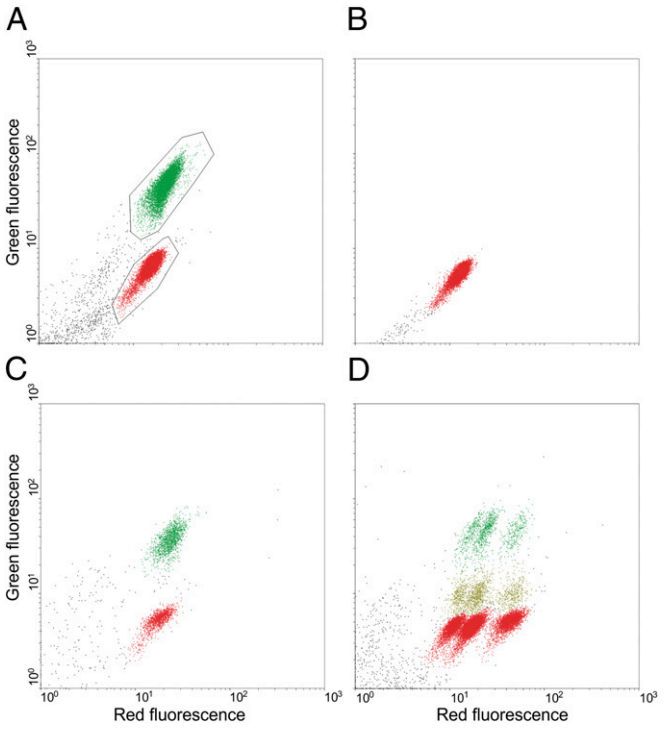

Fig. 3. Flow cytometry analysis and automated sex separation of the T4 strain. (*A*) A total of 16,750 larvae of the T4 line were analyzed according to their level of green and red fluorescence (green and red clouds) and sorted via the gates indicated (black lines). (*B*) Purified larvae (6,415 individuals) gated in *A* as having only background fluorescence were subjected to a second sorting run in the COPAS to check for the absence of contaminating GFP positive larvae. (*C*) Analysis of the fluorescence profile of ~2,500 F_1 larvae carrying the T4 transgene and expressing the I-SceI nuclease from which 1,246 GFP-positive larvae (upper compact cloud) and 902 GFP-negative larvae (lower compact cloud) were COPAS purified. (*D*) Analysis of approximately 9,100 F_2 larvae from an intercross of the F_1. Three classes of red fluorescent larvae are seen along the *x* axis, indicating normal segregation of the red-marked I-SceI transgene. Three classes of green fluorescent larvae are also seen along the *y* axis: high green fluorescence (6.3% of larvae), low green fluorescence (6.8% of larvae), and GFP-negative larvae (87%).

The Y Chromosome Is Accessible to Modification in the Male Germ Line. The presence of the unique I-SceI site within the integrated transformation construct opened the possibility to specifically modify the Y chromosome, if it was accessible to modification in the male germ line. To test for this hypothesis, we crossed T4 males to females of the transgenic line VFS1. This line transcribes the I-SceI homing endonuclease in the germ-line cells under the control of the germ line-specific *vasa* promoter and it also carries a 3xP3-RFP marker gene. As expected, no indication of I-SceI activity was found in the F_1 generation as COPAS sorted GFP-positive larvae and GFP-negative larvae gave rise to male and female adults (>500 individuals were analyzed for each sex), respectively (Fig. 3*C*). Transgenic male and female F_1 individuals were intercrossed, and the F_2 was analyzed by fluorescent sorting. Three classes of red fluorescent larvae are seen along the *x* axis, indicating normal segregation of the RFP-marked I-SceI transgene (Fig. 3*D*). In addition, we observed varying levels of green fluorescence in F_2 larvae on the *y* axis: High GFP fluorescence (seen in 6.3% of all larvae) corresponds to male larvae in which the GFP transgene had not been affected by I-SceI activity, or in which the I-SceI–caused DNA break was repaired in a way that preserved GFP integrity. Lower GFP fluorescence (6.8% of larvae) indicates male larvae in which I-SceI caused a DNA break within GFP followed by imprecise DNA repair that lowered GFP fluorescence. GFP-negative individuals (87%) are expected to include all female larvae, and male larvae in which I-SceI caused a mutation that completely abolished GFP activity. Of 1,134 adult mosquitoes arising from these GFP-negative larvae, 697 were females and 437 (38.5%) were males, indicating a high level of I-SceI activity. Judging by the proportion of male larvae that carried a detectable modification in their Y chromosome, the minimal rate of I-SceI activity can be estimated to be approximately 80%. However, this number is likely to be an underestimate, because some Y chromosomes may have been cut and repaired precisely, or repaired without causing a loss of GFP fluorescence. We concluded that the transgene construct on the Y chromosome was accessible and could be efficiently modified by I-SceI.

Generation of the Y-Linked φC31 Docking Line YAttP. We have shown that I-SceI cleavage can trigger homologous repair between homologous chromosomes and from plasmid repair templates (22). We designed such a "knock-in" strategy to place an AttP-specific integration site onto the Y chromosome. We generated the plasmid 3xP3[AttP]RFP (Fig. 1) by site-directed mutagenesis of the pHome-T construct. It carries the RFP (monomeric DsRed) marker under the control of the 3xP3 promoter. The plasmid was designed to contain, between the promoter and the RFP coding sequence, an AttP recombination site that replaces both the GFP coding sequence and the Actin5C promoter, which are present in pHome-T. The plasmid retains regions of homology to the T4 locus of 0.6 kb and 2.1 kb, 5′ and 3′ of the AttP site, respectively. The AttP recombination signal is recognized by the phage φC31 integrase and would allow for the subsequent site-specific integration of constructs carrying a corresponding AttB signal (24, 25). We hypothesized that I-SceI induced DNA double-strand breaks followed by homologous repair from 3xP3[AttP]RFP in the germ-line progenitor cells could

Downloaded by guest on January 14, 2020

be observed as male larvae switching fluorescent reporters, i.e., showing red but no longer any green fluorescence. This approach allows for the identification of transgenics, and, because the AttP site is located between the 3xP3 promoter and DsRed sequences, would also allow the identification of subsequent φC31 integrations by the loss or conversion of red fluorescence. We coinjected 3xP3 [AttP]RFP into embryos of transgenic line T4 together with a source of I-SceI. We used either in vitro-transcribed I-SceI mRNA or a helper plasmid in which the I-SceI coding sequence is placed under the control of the *A. gambiae vasa* promoter. Table 1 shows the outcome of these experiments. We observed 11 3xP3-RFP–positive progeny of 6,160 larvae screened (0.18%) when using helper plasmid. No RFP-positive larvae were obtained from the progeny of males injected with I-SceI RNA. Genomic PCR and sequencing confirmed that cassette replacement had occurred in this transgenic line now termed YAttP (Fig. 2). This result shows that induced sequence-specific chromosomal breaks can be repaired by homologous recombination with a circular synthetic repair template and, therefore, that *A. gambiae* is amenable to knock-in gene engineering procedures.

Secondary Integration of a Germ Line-Specific Marker Gene by Site-Specific Recombination. To study the activity of germ line-specific regulatory elements on the Y and to demonstrate the utility of line YAttP as a tool for the site-specific genetic engineering of the Y chromosome, we generated the construct attBCFP-VasGFP (Fig. 1). In this vector, an AttB site was placed upstream of a promoterless eCFP coding sequence. In addition, it carries a cassette in which GFP is placed under the control of the *vasa* promoter (26). Recombination between AttP and AttB sites and integration of the plasmid was designed to replace the RFP CDS with the CFP ORF and could be detected by screening for the shift to blue fluorescence in the progeny. Table 1 shows the outcome of experiments in which we injected line YAttP with attBCFP-VasGFP and a helper plasmid expressing the φC31 integrase from a *vasa* promoter sequence. We injected a total of 111 embryos from which 5 male survivors hatched. Because the plasmid attBCFP-VasGFP does not allow the detection of transient fluorescence in G_0 individuals, all male survivors were outcrossed to wild-type females. Of 791 progeny screened, we obtained 6 CFP-positive larvae (0.76%). We confirmed the identity of the newly generated Y-linked strain YVasG by genomic PCR and sequencing.

The *vasa* Regulatory Region Drives GFP Expression During Early Male Spermatogenesis from the Y Chromosome. All transgenic G_2 larvae of strain YVasG showed a gonad-specific GFP signal (Fig. 2*A*) from the L1 larval stage onwards and developed into adult males as expected. We dissected testes from male adults of the YVasG strain and analyzed GFP fluorescence in the gonads by using confocal microscopy (Fig. 2*C*). GFP expression was detectable in all stages of spermatogenesis including the germ-line stem cells (GSCs) in the apical tip and developing spermatocytes but not in supporting somatic cells. DAPI staining of the hub region demonstrated that cells in the anterior tip of the testis were expressing GFP, thus indicating that the regulatory regions of *vasa* were active in GSCs when located on the Y chromosome. The pattern of GFP expression is identical to testes from line Vas2GFP that expresses *vasa*-driven GFP from position 2465559 of chromosome 3L. This finding suggests that strain YAttP in combination with the previously characterized non–sex-specific *vasa* promoter allows the male-exclusive expression of transgenes starting at the earliest stages of spermatogenesis.

Discussion

Although hundreds of transgenic *Anopheles* strains have been generated since the inception of transposase-mediated germ-line transformation, so far no strain harboring a construct inserted into the Y chromosome has been described, despite this chromosome representing an estimated 10% of the genome. In higher eukaryotes, gene expression and the accessibility of DNA is determined, to a large extent, by chromatin structure. These facts suggested that, because of its heterochromatic nature, the Y chromosome was refractory to the random integration catalyzed by transposases or, alternatively, that marker genes commonly used to identify transformation events would undergo complete silencing when located on the Y chromosome. Here, we characterized a Y-linked transgenic line obtained by random integration. Expression of fluorescent maker genes in two different tissues suggested that the transgene construct had inserted within a region of chromatin, allowing transcriptional activity. Such a euchromatic region is of interest because it might host male-specific genes. We identified a number of Y-specific PCR markers that may aid in characterizing *anopheline* population history and geographic structure, which is critical for understanding genetic structure and speciation and for the effective implementation of malaria control strategies (27, 28). The majority of selected putative ORFs analyzed by genomic PCR were also amplified from females, indicating that these sequences are likely to be present also on other chromosomes. We also identified a number of transcribed regions in the neighborhood of the integrated construct. Male-specific NTRs could represent putative Y-linked genes and lend themselves for future functional

Table 1. Embryo microinjections and fluorescent phenotypes observed in the outcross progeny

Source	Embryos injected	Transient males/male survivors outcrossed	Larvae screened	Fluorescent phenotypes observed	
I-SceI					
RNA	~2,000	6	964	Nonfluorescent	473 (♀)
				Actin5C-RFP/3xP3-GFP	461 (♂)
				Actin5C-RFP	30 (♂)
				3xP3-RFP	**0**
Plasmid	~4,000	15	6,160	Nonfluorescent	3,113 (♀)
				Actin5C-RFP/3xP3-GFP	2,820 (♂)
				Actin5C-RFP	216 (♂)
				3xP3-RFP	**11** (♂)
Integrase					
Plasmid	111	5	791	Nonfluorescent	390 (♀)
				3xP3-RFP	395 (♂)
				3xP3-CFP/Vas-GFP	**6** (♂)

T4 males injected with 3xP3[AttP]RFP plasmid and either in-vitro transcribed I-SceI RNA or the pVas2-I-SceI helper plasmid were outcrossed to wild-type females. The resulting progeny was screened for fluorescence. YAttP males injected with attBCFP-VasaGFP plasmid and the integrase helper plasmid were outcrossed to wild-type females. The resulting progeny was screened for fluorescence. Phenotypes indicating successful gene conversion or integration are shown in bold.

Downloaded by guest on January 14, 2020

analysis. Many of the ORFs selected for genomic PCR and those underlying NTRs largely match known repetitive elements and transposons. Interestingly, putative paralogues of a number of the identified transcribed regions map to scaffold AAAB01008885.1 close to the centromere of the X chromosome (Dataset S1).

The fact that the T4 locus was transcriptionally active made it an exciting target for gene engineering. We used line T4 in combination with a plasmid template to stimulate homologous recombination and introduce a site-specific recombination signal onto the Y chromosome. Our approach relied on the well characterized endonuclease I-SceI and the presence of an I-SceI site within the GFP ORF of the transgene. The availability of novel programmable endonucleases based on the transcription activator-like effector nuclease (TALEN) or clustered regularly interspaced short palindromic repeats (CRISPR/Cas) architectures as well as a growing number of sequences uniquely present on the Y chromosome such as those described in this study suggests that this powerful approach could be extended to introduce foreign sequences (such as docking sites or desired mutations) also into other regions of the Y or, indeed, into any locus of interest on other chromosomes. Although homologous recombination was inefficient (occurring in 0.18% of the progeny), it helped to successfully introduce the AttP recombination site, paving the way for the more efficient site-specific integration method. We found that secondary integration mediated by the φC31 integrase occurred at least at a 4× higher rate.

Although the use of COPAS automated fluorescent larval sorting is limited to research, it has been suggested that future iterations of this technology could be scaled up to develop robust protocols for the mass production of sorted insects. The strains carrying Y-linked fluorescent markers we have developed are, as we show, ideally suited for such automated fluorescence sexing. This property may become useful in applications requiring large numbers of pure males, such as SIT (29), and in experiments that require genetic crosses to a large number of virgin wild-type females, readily obtained here from the T4 line purified at the early larval stage. The fact that transgenic males give rise to male progeny that all inherit the transgene represents an additional advantage of these strains because, unlike with autosomal markers for sexing (30), no segregation of the sexing marker occurs so that a pure-breeding stock can be maintained. This feature allows combining the sexing trait with other autosomally encoded traits, for example male sterility. Alternatively, genetic sexing traits such as insecticide resistance markers can now also be easily introduced to the Y chromosome. The ability to introduce transgenes to the Y chromosome also allows the generation of male-exclusive genetic traits. We have shown that the regulatory regions of the *Anopheles* gene *vasa* drive expression of transgenes in the male and female germ line throughout larval and adult gonad development (26). We then attempted to generate a male-specific version of the *vasa* promoter by using a shortened 5′ UTR. However, this approach not only reduced overall activity of the promoter, but also lead to the loss of expression in the GSCs of the testes. As demonstrated here, the ability to express transgenes from the Y chromosome allows for the generation of male-exclusive traits by using non–sex-specific components and to expressed tagged or mutant forms of Y-linked genes.

We have recently generated a synthetic male sex ratio distortion system that operates by destroying the X chromosome during male spermatogenesis and creates progeny consisting almost exclusively of males (Galizi et al., under review). The induction of extreme male-biased sex ratios could be a powerful strategy to supress or eliminate pest populations. However, transgenes when expressed from autosomal locations are transmitted to only half of the male progeny. In its current form, this technology would require continuous mass releases to achieve population suppression. In contrast, if the distorter trait could be successfully linked to the Y chromosome, this chromosome would be converted into a selfish sex chromosome that, in theory, could spread through and eliminate a natural vector population even when seeded by a very low number of transgenic males (31). Alternative uses of Y-linked meiotic drive have also been described (32). The YattP line described here paves the way to establishing such vector control strategies and, thus, represents a promising tool in the fight against malaria.

Materials and Methods

Mosquito Rearing. The wild-type *A. gambiae* strain (G_3) and the transgenic mosquito line YAttP and YVasG were reared under standard condition at 28 °C and 80% relative humidity with access to fish food as larvae and 5% (wt/vol) glucose solution as adults. For egg production, young adult mosquitoes (3–5 d after emergence) were allowed to mate for at least 6 d and then fed on mice. Three days later, an egg bowl containing rearing water (dH_2O supplemented with 0.1% pure salt) was placed in the cage. One to two days after hatching, the larvae (L1 stage) were placed into rearing water containing trays. The protocols and procedures used in the study were approved by the Animal Ethics Committee of Imperial College in compliance with UK Home Office regulations.

Analysis of Y-Specific BAC Clone. High molecular weight DNA was extracted from transgenic T4 individuals. A BAC library was constructed in vector pIndigoBAC5-HindIII from Epicentre by using the standard HindIII cloning site. The library was pooled in 1 × 96-well plate with each well containing approximately 125 independent primary clones with an average insert of 120 kb. Primers actin5rev (5′TGGAAATGAGAAGTAAGGTGCATCTGCA3′) and exfp3sEq. (5′ATCCGCCACAACATCGAGGACG3′) were used to screen the clones for the presence of a Y-specific insert. Size determination of the insert was performed by NotI digestion and PFGE gel separation. Single Molecule Real Time (SMRT) technology was chosen to sequence one positive BAC clone, and PacBio SMRT Portal assembler was used for sequence quality filtering and assembly. The contigs were uploaded to GenBank (accession nos. KJ608148–KJ608160).

RNAseq Analysis. Illumina paired-end 101-bp reads derived from libraries of whole adult males and females (unfed) were aligned against the assembled 13 contigs by using STAR (33). To control for reads that map to repetitive regions, only reads that mapped uniquely within all 13 contigs were considered. We used Cufflinks (34) to predict transcripts, and the aligned reads were used to generate updated consensus sequences of the contigs, which generally increased ORF size. Sequences encompassing transcribed regions that displayed male-specific expression were BLAT searched (BLAST-like alignment tool searched; http://genome.cshlp.org/content/12/4/656) against the *A. gambiae* PEST assembly (AgamP3) to identify whether these may already be present in the genome assembly, or to identify putative paralogues on other chromosomes. As a control, we mapped male and female reads against sequences of two recently described Y-linked genes, gYG1 and gYG2 (20), and the α-tubulin (AGAP001219) gene expressed ubiquitously in both sexes.

Generation of I-SceI RNA. The mMESSAGEmachine kit (Ambion) was used to obtain a 5′capped mRNA (I-SceI RNA) coding for the homing endonuclease I-SceI. The *I-SceI* gene was amplified from the target plasmid pP[v+,70I-SceI] by using primers ForI-SceI 5′CGCGTAATACGACTCACTATAGGGCGAATTGGGTACCGGGCCCCCCCTCGAGGTCGACGGTATCGATAAGCTTGATATCATGGGATCATCATCAGACGACGAAGCTACGGCCGATGCGCA3′ containing a T7 RNA polymerase promoter and RevI-SceI (5′TTTTTTTTTTTTTTTTTTTTTTTATTTCAGGAAAGTTTCGGAGGAGATAGTGTTCGGCAGTTTGTACA3′).

Generation of Transformation Vectors. The Phusion Site-Directed Mutagenesis Kit was used to generate the 3xP3[AttP]RFP plasmid. For this cloning strategy, we used primers, F1 (5′-CTTTGAGTTCTCTCAGTTGGGGGCACGCAACATGGTGCGCTCCTCCAAGAACG-3′) and R2 (5′-GTTACCCCAGTTGGGGCGTAGGGTCTAGTCGACTCTAGCGGTACCCCGATTG-3′), containing one-half attP site each at the 5′. The plasmid attBCFP-VasaGFP was generated by cloning of a blunt synthetic fragment, attB-fragment, containing eCFP preceded by an attB site and followed by the SV40 terminator, into the StuI site of pfVasIntGFP carrying the VasGFP (26).

Embryo Microinjections. Embryos were injected by using a Femtojet Express injector and a Narishige 202ND micromanipulator mounted on an inverted microscope (Nikon TE-DH100W) with a mixture of 0.2 μg/μL 3xP3[AttP]RFP plasmid or 0.8 μg/μL in vitro-transcribed I-SceI RNA, and 0.4 μg/μL pVas2-I-SceI

Downloaded by guest on January 14, 2020

helper plasmid for the generation of the transgenic mosquito line YAttP. The hatched larvae were screened for transient expression of the DsRed marker, and the 0.18% positive was grown up and crossed to wild-type mosquitoes. The progeny of these crosses was analyzed for DsRed fluorescence. Individual larvae showing expression of the selectable marker were then separated, and the adults that emerged were crossed individually with wild-type mosquitoes to obtain transgenic lines. Embryos were injected with a mixture of 0.2 μg/μL attBCFP-VasaGFP plasmid and 0.5 μg/μL Integrase helper plasmid to generate the transgenic mosquito line YVasG. The hatched larvae were grown up, and the surviving males were crossed to wild-type mosquitoes. The progeny of these crosses was analyzed for eCFP fluorescence. Individual larvae showing expression of the selectable marker were then separated, and the adults that emerged were crossed individually with wild-type mosquitoes to obtain transgenic lines. Transgenic mosquitoes at different developmental stages were analyzed on a Nikon inverted microscope (Eclipse TE200) at a wavelength of 488 nm to detect eGFP expression (Filter 535/20 nm emission, 505 nm dichroic) and 563 nm (Filter 630/30 nm emission, 595 nm dichroic) to detect DsRed expression.

COPAS-Assisted Analysis and Sorting of Larval Populations. COPAS sorting was performed as described in Marois et al. (30). Briefly, newly hatched larvae were transferred to the reservoir of the large-particles flow cytometry COPAS SELECT instrument (Union Biometrica) equipped with a multiline argon laser (488, 514 nm) and a diode laser (670 nm), analyzed, and sorted with the Biosort5281 software with a 488-nm filter and the following acquisition parameters: Green PMT 500, Red PMT 600, Delay 8; Width 6, pure mode selection with superdrops. The flow rate was kept below 15 detected objects per second. Larvae identified as male and female were dispensed sequentially in separate Petri dishes. The LMD files generated by the COPAS software were imported in the WinMDI freeware for data analysis and figure preparation.

Confocal Microscopy. Dissected testes from transgenic lines YVasG and Vas2GFP were fixed in methanol-free 4% formaldehyde (Pierce) in PBS for 30 min and washed three times for 15 min in 0.1% Tween-20 PBS. Testes were then transferred on fresh slides containing Vectashield mounting medium with DAPI (Vectorlabs) with coverslips. Testes images were taken by using a Zeiss LSM 510 Laser scanning confocal microscope and a 20× objective in two cannels (DAPI and GFP) sequentially.

FISH. FISH of the sexual chromosomes, X and Y, in the transgenic line T4 was performed according to a previously established protocol (35). The X chromosome was marked by using as a probe an intergenic spacer region of rDNA generated by PCR using primers UN (GTGTGCCCCTTCCTCGATGT) and GA (CTGGTTTGGTCGGCACGTTT) and labeled with the cyanine dye Cy5. The Y chromosome was marked by using as a probe the p-Home T plasmid Cy3-labeled by nick translation (Roche Nick Translation Kit). Chromosomes where counterstained with DAPI.

ACKNOWLEDGMENTS. We thank Cecilia Cheung and Stephen M. Rothery from the Facility for Imaging by Light Microscopy, Tony Nolan, and Maria Sharakhova. This work is funded by a grant from the Foundation for the National Institutes of Health through the Vector-Based Control of Transmission: Discovery Research program of the Grand Challenges in Global Health initiative of the Bill & Melinda Gates Foundation and the European Research Council under European Union's Seventh Framework Programme ERC Grant 335724 (to N.W.).

1. World Health Organization (2012) World Malaria Report 2012. (World Health Organization, Geneva).
2. Clements AN (1992) *The Biology of Mosquitoes: Development, Nutrition, and Reproduction* (Chapman & Hall, London).
3. Baker RH, Sakai RK (1979) Triploids and male determination in the mosquito, Anopheles culicifacies. *J Hered* 70(5):345–346.
4. Charlesworth B (1996) The evolution of chromosomal sex determination and dosage compensation. *Curr Biol* 6(2):149–162.
5. Muller HJ (1932) Some genetic aspects of sex. *Am Nat* 66(703):118–138.
6. Lahn BT, Pearson NM, Jegalian K (2001) The human Y chromosome, in the light of evolution. *Nat Rev Genet* 2(3):207–216.
7. Fisher RA (1931) The evolution of dominance. *Biol Rev Camb Philos Soc* 6(4):345–368.
8. Rice WR (1984) Sex chromosomes and the evolution of sexual dimorphism. *Evolution* 38(4):735–742.
9. Steinemann M, Steinemann S (2000) Common mechanisms of Y chromosome evolution. *Genetica* 109(1-2):105–111.
10. Burgoyne PS (1998) The mammalian Y chromosome: A new perspective. *Bioessays* 20(5):363–366.
11. Castillo ER, Bidau CJ, Martí DA (2010) Neo-sex chromosome diversity in Neotropical melanopline grasshoppers (Melanoplinae, Acrididae). *Genetica* 138(7):775–786.
12. Rice WR (1994) Degeneration of a nonrecombining chromosome. *Science* 263(5144): 230–232.
13. Bachtrog D (2003) Adaptation shapes patterns of genome evolution on sexual and asexual chromosomes in Drosophila. *Nat Genet* 34(2):215–219.
14. Carvalho AB, Lazzaro BP, Clark AG (2000) Y chromosomal fertility factors kl-2 and kl-3 of Drosophila melanogaster encode dynein heavy chain polypeptides. *Proc Natl Acad Sci USA* 97(24):13239–13244.
15. Carvalho AB, Dobo BA, Vibranovski MD, Clark AG (2001) Identification of five new genes on the Y chromosome of Drosophila melanogaster. *Proc Natl Acad Sci USA* 98(23):13225–13230.
16. Steinemann M, Steinemann S (1998) Enigma of Y chromosome degeneration: Neo-Y and neo-X chromosomes of Drosophila miranda a model for sex chromosome evolution. *Genetica* 102–103(1-6):409–420.
17. Maggert KA, Golic KG (2002) The Y chromosome of Drosophila melanogaster exhibits chromosome-wide imprinting. *Genetics* 162(3):1245–1258.
18. Holt RA, et al. (2002) The genome sequence of the malaria mosquito Anopheles gambiae. *Science* 298(5591):129–149.
19. Mason GF (1967) Genetic studies on mutations in species A and B of the Anopheles gambiae complex. *Genet Res* 10(3):205–217.
20. Hall AB, et al. (2013) Six novel Y chromosome genes in Anopheles mosquitoes discovered by independently sequencing males and females. *BMC Genomics* 14:273.
21. Criscione F, Qi Y, Saunders R, Hall B, Tu Z (2013) A unique Y gene in the Asian malaria mosquito Anopheles stephensi encodes a small lysine-rich protein and is transcribed at the onset of embryonic development. *Insect Mol Biol* 22(4):433–441.
22. Windbichler N, et al. (2011) A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature* 473(7346):212–215.
23. Krzywinski J, Nusskern DR, Kern MK, Besansky NJ (2004) Isolation and characterization of Y chromosome sequences from the African malaria mosquito Anopheles gambiae. *Genetics* 166(3):1291–1302.
24. Thorpe HM, Smith MC (1998) In vitro site-specific integration of bacteriophage DNA catalyzed by a recombinase of the resolvase/invertase family. *Proc Natl Acad Sci USA* 95(10):5505–5510.
25. Meredith JM, et al. (2011) Site-specific integration and expression of an anti-malarial gene in transgenic Anopheles gambiae significantly reduces Plasmodium infections. *PLoS ONE* 6(1):e14587.
26. Papathanos PA, Windbichler N, Menichelli M, Burt A, Crisanti A (2009) The vasa regulatory region mediates germline expression and maternal transmission of proteins in the malaria mosquito Anopheles gambiae: A versatile tool for genetic control strategies. *BMC Mol Biol* 10:65.
27. Hammer MF, Zegura SL (1996) The role of the Y chromosome in human evolutionary studies. Evolutionary Anthropology: Issues. *News Rev (Melb)* 5(4):116–134.
28. Hurles ME, Jobling MA (2001) Haploid chromosomes in molecular ecology: Lessons from the human Y. *Mol Ecol* 10(7):1599–1613.
29. Condon KC, et al. (2007) Genetic sexing through the use of Y-linked transgenes. *Insect Biochem Mol Biol* 37(11):1168–1176.
30. Marois E, et al. (2012) High-throughput sorting of mosquito larvae for laboratory studies and for future vector control interventions. *Malar J* 11(1):302.
31. Hamilton WD (1967) Extraordinary sex ratios. A sex-ratio theory for sex linkage and inbreeding has new implications in cytogenetics and entomology. *Science* 156(3774): 477–488.
32. Huang Y, Magori K, Lloyd AL, Gould F (2007) Introducing desirable transgenes into insect populations using Y-linked meiotic drive - a theoretical assessment. *Evolution* 61(4):717–726.
33. Dobin A, et al. (2013) STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1): 15–21.
34. Trapnell C, et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28(5):511–515.
35. Timoshevskiy VA, Sharma A, Sharakhov IV, Sharakhova MV (2012) Fluorescent in situ hybridization on mitotic chromosomes of mosquitoes. *J Vis Exp* (67):e4215.

Downloaded by guest on January 14, 2020