PNAS
www.pnas.org

Supplementary Information for ***Nix* alone is sufficient to convert female *Aedes aegypti* into fertile males and *myo-sex* is needed for male flight**

Azadeh Aryan, Michelle Anderson, James K. Biedler, Yumin Qi, Justin M. Overcash, Anastasia N. Naumenko, Maria V. Sharakhova, Chunhong Mao, Zach N. Adelman, and Zhijian Tu*

Send correspondence to Zhijian Tu
Email: jaketu@vt.edu

**This PDF file includes:**

Tables S1 to S10
Figures S1 to S7
SI References

**Other supplementary materials for this manuscript include the following:**

A link to a video showing the flightless phenotype of the sex-converted m/m males:
**https://youtu.be/fwUqN5iKTi0**

RNAseq datasets are available at NCBI with BioProject Accession PRJNA625258.

## Table S1. Plasmid injection and screening of G1 for transformants

| Donor plasmid | Mos1 helper | # embryos injected | # G0 (%survivors) | # pools | # of G1 screened | Pools with GFP+ G1 progeny (#) |
|---|---|---|---|---|---|---|
| pMos-*Nix*-PUbGFP (N1) | pGL3-PUbMos1 | 1053 | 130 (12.3%) | 7 | ~13180 | P4 (5 normal male) |
| pMos-*Nix*-PUbGFP (N1) | pGL3-PUbMos1 | 1107 | 62 (5.6%) | 4 | ~7970 | N3 (12 normal male) |
| pMos-Ntag*Nix*-PUbGFP (N2) | pGL3-PUbMos1 | 1084 | 110 (10.1%) | 5 | ~10600 | P4 (4 flightless♂, 1 normal ♂) |

## Table S2. Screening of the N1 transgenic line

| **Generation** | Transgenic ♂ | Non-transgenic ♂ | Transgenic ♀ | Non-transgenic ♀ | Transgenic flightless ♂ |
|---|---|---|---|---|---|
| G2[a,] | 342[b] | 71 | 0 | 291 | ND |
| G3 | 419[b] | 95 | 0 | 329 | ND |
| G10 | 1166 | 309 | 0 | 808 | 400 |
| G11 | 280 | 91 | 0 | 251 | 83 |
| G13 | 1250 | 429 | 0 | 1127 | 419 |
| Total (G10-11, G13)[c] | 2696 | 829 | 0 | 2186 | 902 |

a. These numbers represent the progeny of just one of the 5 G1 males (N1, P4-1). G3 and subsequent generations were all derived from this cross. No transgenic females were found in G2 progeny of crosses of any G1 males.
b. In early generations we did not screen for flightless individuals, so these numbers include flightless males.
c. The numbers are in general consistent with that the N1 transgene is distantly linked to the M locus. No Chi-square test was performed as the recombination rate is not known.

**Table S3. Screening of the N2 transgenic line**

| Generation | Transgenic ♂ | Non-transgenic ♂ | Transgenic ♀ | Non-transgenic ♀ | Transgenic flightless ♂ |
|---|---|---|---|---|---|
| G2[a] | 242[b] | 152 | 0 | 124 | ND |
| G3 | 142 | 168 | 0 | 151 | 117 |
| G7 | 786 | 749 | 0 | 763 | 348 |
| G8 | 138 | 188 | 0 | 154 | 122 |
| G9[c] | 342 | 531 | 0 | 501 | 301 |
| G14 | 202 | 212 | 0 | 202 | 210 |
| G15[c] | 235 | 231 | 0 | 256 | 215 |
| Total (G3, G7-G9, G14-G15)[d] | 1845 | 2079 | 0 | 2027 | 1313 |

a. These numbers represent the progeny of the only flying G1 males (N2, P4-2).
b. In this generation we did not screen for flightless individuals, so these numbers include flightless transgenic males.
c. 37 and 13 dead positive males from G9 and G15, respectively, were not included in the table as we did not determine whether they were transgenic males (M/m; N/+) or transgenic flightless males (m/m; N/+). It is likely that they are flightless males, some of which have difficulty emerging from the pupa case.
d. G2 numbers are not included in the total. As N2 is inserted in chromosome 2, we expected a 1:1:1:1 ratio among the four possible genotypes, which was observed in G14 (Chi-square test, $p=0.94$) and G15 (Chi-square test, $p=0.29$ when the 13 dead positive males are not included; $p=0.24$ and 0.57 when the 13 dead positive males are arbitrarily assigned to transgenic males or transgenic flightless males, respectively). However, this expectation was not met in earlier generations (Chi-square test, $p<0.05$).

**Table S4. Primers, probes, sgRNAs used in this study**

| Primer/probe/sgRNA name & description | Sequence |
|---|---|
| Primers used to amplify and clone the *Nix* promoter region | |
| PF2 | AGAACAAAAACCGCTCCAG |
| PR2 | CAAATTCTGCATTGATCTCAGC |
| | |
| Inverse PCR to determine transgene insertion sites | |
| 1446F | CAAATTGCCCGAGAGATG |
| 1448R | ATTCCACGCCAGTAGAACTC |
| 1447F | GCGACGGCAAATACTTTG |
| 1449R | TAAGAATCGAAGCGCTGC |
| | |
| Primers for distinguishing M/m and m/m individuals | |
| *myo-sex*-F (amplify *myo-sex*, a gene in the M-locus) | CCTTCAAGCACACCGTTACA |
| *myo-sex*-R (amplify *myo-sex*, a gene in the M-locus) | TCACTATGCAGGAGTTGTTTCG |
| Endo*Nix*-F2 (span exon 1 and intron 1 of *Nix*, thus specific to the endogenous *Nix* gene) | CAAAGAATGATGTGGTTGTCAAA |
| Endo*Nix*-R2 (span exon 1 and intron 1 of *Nix*, thus specific to the endogenous *Nix* gene) | TGTTTGCGCGTTGTGTATTT |
| | |
| Primers for RT-PCR | |
| Endo*Nix*-F1 (1 primer spans the 5'UTR and the ORF, thus cannot amplify the N2 *Nix* transgene) | CGTGCAAATGTGTAAAAAAAGAAATGC |
| Endo*Nix*-R1 (1 primer spans the 5'UTR and the ORF, thus cannot amplify the N2 *Nix* transgene) | AGCTCAGCCATTGATGTCTTACAA |
| N2*Nix*-F (span the N2 *Nix* transcribed from the *Nix* transgene) | ACCACATTGCGACGTTAAAC |
| N2*Nix*-R (span the N2 *Nix* transcribed from the *Nix* transgene) | AAATGTGGTATGGCTGATTATG |
| | |
| rps7 control primers (for RT-PCR and gDNA PCR, RT-PCR product is shorter due to intron splicing) | |
| RPS7-F | ATGGTTTTCGGATCAAAGG |
| RPS7-R | CTTGTGTTCAATGGTGGTCTG |
| | |
| Digital droplet PCR primers and probes (Forward and Reverse primers followed by a Taqman probe) | |
| 829F | GAACTTGTCAAACGATCTCAATG |
| 834R | ATGTTCAGATTGTGCAATCG |
| Dsx F | 5'FAM-TGACGAAGGTCAAGCCGTG -BHQ-3' |
| | |
| 837F | GATACCCCTGGGAGATGATG |
| 839R | TGGAACGCTTCGGAAAGTAG |
| Dsx M | 5'FAM-CGGATTGACGAAGGATACGACATT-BHQ-1 3' |
| | |
| 875F | CCAGCACTTCTCCCTCAATC |
| 878R | AATACTGCTGGTCCATTGTTCC |
| Fru F | 5'FAM-CTGTTCCAGCCGTTAAAGGTCG-BHQ-3' |
| | |
| 870F | CTATCCCCAGATCAACGGTG |
| 874R | ACTGATGGTTGTTCCAGCG |

| | |
|---|---|
| Fru M | 5'FAM -CGCACCGAAGGAACAGCAC-BHQ-3' |
| 1393F | TACAAGATGCGCAATGGATA |
| 1394R | TGGCCAGATAGTCGATGTAAT |
| Gene AAEL002401 (internal reference) | 5'HEX-CGTATTGGTTGGAGGCTATGACGA-BHQ 3' |
| *Myo-sex* sgRNA | |
| *Myo-sex* sgRNA1 (PAM site underlined) | GTAACCGTTGCTTTACCAGGTGG |
| *Myo-sex* sgRNA2 (PAM site underlined) | GAAGCCGAAGGATACGTTCAAGG |
| *Myo-sex* sgRNA3 (PAM site underlined) | GGTTACCAAGTCACCCTTGGTGG |
| *Myo-sex* sgRNA primers | |
| *Myo-sex* sgRNA 1F (Blue: T7 promoter; red: target sequence; green: overlap with sgRNA R) | GAAATTAATACGACTCACTATAGGGGTAACCGT TGCTTTACCAGGGTTTTAGAGCTAGAAA |
| *Myo-sex* sgRNA 2F (Blue: T7 promoter; red: target sequence; green: overlap with sgRNA R) | GAAATTAATACGACTCACTATAGGGGAAGCCG AAGGATACGTTCAGTTTTAGAGCTAGAAA |
| *Myo-sex* sgRNA 3F (Blue: T7 promoter; red: target sequence; green: overlap with sgRNA R) | GAAATTAATACGACTCACTATAGGGGGTTACCA AGTCACCCTTGGGTTTTAGAGCTAGAAA |
| sgRNA R (Common sgRNA reverse primer sequence) | AAAAGCACCGACTCGGTGCCACTTTTTCAAGTT GATAACGGACTAGCCTTATTTTAACTTGCTATT TCTAGCTCTAAAAC |
| Primers used to amplify the region targed by *myo-sex* sgRNAs | |
| *Myo-sex* PR | CATAAAGAAATGTGAAGAAAAT |
| *Myo-sex* PF | TTTGGGCTACATCATCCTTCAACGATTC |

**Table S5. Length (millimeter) of the right wing of individuals of the N2 transgenic line**

| Non-transgenic male (M/m; +/+) | Transgenic male (M/m; N2/+) | Transgenic flightless male (m/m; N2/+) | Non-transgenic female (m/m; +/+) |
|---|---|---|---|
| 2.42 | 2.298 | 2.49 | 2.643 |
| 2.394 | 2.35 | 2.426 | 2.987 |
| 2.198 | 2.343 | 2.562 | 3.099 |
| 2.248 | 2.223 | 2.363 | 3.107 |
| 2.28 | 2.151 | 2.408 | 3.133 |
| 2.411 | 2.347 | 2.364 | 2.999 |
| 2.39 | 2.298 | 2.391 | 3.127 |
| 2.363 | 2.302 | 2.392 | 3.112 |
| 2.346 | 2.306 | 2.443 | 3.229 |
| 2.352 | 2.325 | 2.483 | 3.113 |
| 2.367 | 2.267 | 2.369 | 3.115 |
| 2.23 | 2.391 | 2.554 | 3.197 |
| 2.438 | 2.364 | 2.472 | 3.184 |
| 2.496 | 2.252 | 2.465 | 3.11 |
| 2.384 | 2.284 | 2.379 | 3.009 |
| 2.374 | 2.145 | 2.443 | 3.114 |
| 2.316 | 2.303 | 2.533 | 3.176 |
| 2.201 | 2.222 | 2.478 | 3.208 |
| 2.35 | 2.411 | 2.512 | 3.256 |
| 2.293 | 2.32 | 2.551 | 3.061 |
| 2.344 | 2.347 | 2.531 | 3.095 |
| 2.41 | 2.351 | 2.452 | 3.127 |
| 2.438 | 2.428 | 2.498 | 3.037 |
| 2.444 | 2.359 | 2.498 | 3.171 |
| 2.387 | 2.432 | 2.45 | 3.074 |
| 2.452 | 2.36 | 2.462 | 3.149 |
| 2.236 | 2.206 | 2.445 | 3.068 |
| 2.351 | 2.398 | 2.503 | 3.151 |
| 2.362 | 2.241 | 2.279 | 2.934 |
| 2.333 | 2.288 | 2.456 | 3.026 |
| Average 2.354 | Average 2.310 | Average 2.455 | Average 3.094 |

Notes:

1) Thirty individuals were measured from each of the four groups, which were the same $G_{14}$ cohort emerged from the same pool of larvae.
2) One-way ANOVA was performed ($p<0.001$) with the Tukey Simultaneous Test for Difference of Means. All pairwise comparisons were significant ($p<0.0001$), except for non-transgenic males vs. transgenic males ($p=0.1982$). The assumption of equal variance is met using Levene's method ($p=0.68$).

**Table S6. Pearson correlation based on log(FPKM+1) of all genes in each sample.**

| | WTF2 | WTF3 | WTF4 | WTM2 | WTM3 | WTM4 | TRF2 | TRF3 | TRF4 | TRM2 | TRM3 | TRM4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WTF2 | 1.0000 | 0.9864 | 0.9873 | 0.9068 | 0.9067 | 0.9028 | 0.9042 | 0.9028 | 0.9060 | 0.9074 | 0.9061 | 0.9059 |
| WTF3 | 0.9864 | 1.0000 | 0.9886 | 0.9040 | 0.9039 | 0.8987 | 0.9015 | 0.9005 | 0.9032 | 0.9054 | 0.9059 | 0.9042 |
| WTF4 | 0.9873 | 0.9886 | 1.0000 | 0.9057 | 0.9069 | 0.9014 | 0.9029 | 0.9010 | 0.9061 | 0.9075 | 0.9082 | 0.9074 |
| WTM2 | 0.9068 | 0.9040 | 0.9057 | 1.0000 | 0.9892 | 0.9861 | 0.9833 | 0.9831 | 0.9838 | 0.9886 | 0.9883 | 0.9887 |
| WTM3 | 0.9067 | 0.9039 | 0.9069 | 0.9892 | 1.0000 | 0.9875 | 0.9840 | 0.9833 | 0.9851 | 0.9878 | 0.9877 | 0.9900 |
| WTM4 | 0.9028 | 0.8987 | 0.9014 | 0.9861 | 0.9875 | 1.0000 | 0.9825 | 0.9827 | 0.9863 | 0.9835 | 0.9836 | 0.9840 |
| TRF2 | 0.9042 | 0.9015 | 0.9029 | 0.9833 | 0.9840 | 0.9825 | 1.0000 | 0.9886 | 0.9864 | 0.9835 | 0.9835 | 0.9840 |
| TRF3 | 0.9028 | 0.9005 | 0.9010 | 0.9831 | 0.9833 | 0.9827 | 0.9886 | 1.0000 | 0.9882 | 0.9843 | 0.9820 | 0.9837 |
| TRF4 | 0.9060 | 0.9032 | 0.9061 | 0.9838 | 0.9851 | 0.9863 | 0.9864 | 0.9882 | 1.0000 | 0.9839 | 0.9821 | 0.9842 |
| TRM2 | 0.9074 | 0.9054 | 0.9075 | 0.9886 | 0.9878 | 0.9835 | 0.9835 | 0.9843 | 0.9839 | 1.0000 | 0.9891 | 0.9891 |
| TRM3 | 0.9061 | 0.9059 | 0.9082 | 0.9883 | 0.9877 | 0.9836 | 0.9835 | 0.9820 | 0.9821 | 0.9891 | 1.0000 | 0.9887 |
| TRM4 | 0.9059 | 0.9042 | 0.9074 | 0.9887 | 0.9900 | 0.9840 | 0.9840 | 0.9837 | 0.9842 | 0.9891 | 0.9887 | 1.0000 |

Pearson correlation analysis was performed in R
https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cor

**Table S7. Spearman correlation based on log(FPKM+1) of all genes in each sample.**

| | WTF2 | WTF3 | WTF4 | WTM2 | WTM3 | WTM4 | TRF2 | TRF3 | TRF4 | TRM2 | TRM3 | TRM4 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WTF2 | 1.0000 | 0.9687 | 0.9696 | 0.8890 | 0.8893 | 0.8866 | 0.8872 | 0.8865 | 0.8893 | 0.8905 | 0.8874 | 0.8877 |
| WTF3 | 0.9687 | 1.0000 | 0.9691 | 0.8861 | 0.8865 | 0.8827 | 0.8843 | 0.8832 | 0.8864 | 0.8881 | 0.8861 | 0.8859 |
| WTF4 | 0.9696 | 0.9691 | 1.0000 | 0.8877 | 0.8890 | 0.8850 | 0.8859 | 0.8836 | 0.8884 | 0.8889 | 0.8891 | 0.8890 |
| WTM2 | 0.8890 | 0.8861 | 0.8877 | 1.0000 | 0.9722 | 0.9699 | 0.9663 | 0.9654 | 0.9668 | 0.9715 | 0.9702 | 0.9719 |
| WTM3 | 0.8893 | 0.8865 | 0.8890 | 0.9722 | 1.0000 | 0.9706 | 0.9661 | 0.9655 | 0.9687 | 0.9706 | 0.9701 | 0.9737 |
| WTM4 | 0.8866 | 0.8827 | 0.8850 | 0.9699 | 0.9706 | 1.0000 | 0.9656 | 0.9657 | 0.9693 | 0.9684 | 0.9669 | 0.9683 |
| TRF2 | 0.8872 | 0.8843 | 0.8859 | 0.9663 | 0.9661 | 0.9656 | 1.0000 | 0.9705 | 0.9686 | 0.9664 | 0.9654 | 0.9666 |
| TRF3 | 0.8865 | 0.8832 | 0.8836 | 0.9654 | 0.9655 | 0.9657 | 0.9705 | 1.0000 | 0.9701 | 0.9677 | 0.9643 | 0.9674 |
| TRF4 | 0.8893 | 0.8864 | 0.8884 | 0.9668 | 0.9687 | 0.9693 | 0.9686 | 0.9701 | 1.0000 | 0.9679 | 0.9654 | 0.9690 |
| TRM2 | 0.8905 | 0.8881 | 0.8889 | 0.9715 | 0.9706 | 0.9684 | 0.9664 | 0.9677 | 0.9679 | 1.0000 | 0.9716 | 0.9728 |
| TRM3 | 0.8874 | 0.8861 | 0.8891 | 0.9702 | 0.9701 | 0.9669 | 0.9654 | 0.9643 | 0.9654 | 0.9716 | 1.0000 | 0.9710 |
| TRM4 | 0.8877 | 0.8859 | 0.8890 | 0.9719 | 0.9737 | 0.9683 | 0.9666 | 0.9674 | 0.9690 | 0.9728 | 0.9710 | 1.0000 |

Spearman correlation analysis was performed in R
https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/cor

## Table S8. PCA based on log(FPKM+1) of all genes in each sample.

| | PC1 | PC2 | PC3 | PC4 | PC5 | PC6 | PC7 | PC8 | PC9 | PC10 | PC11 | PC12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| WTF2 | 0.284 | -0.489 | -0.084 | -0.221 | 0.708 | -0.104 | 0.007 | -0.134 | 0.159 | -0.26 | 0.017 | -0.011 |
| WTF3 | 0.29 | -0.512 | 6E-04 | 0.191 | -0.325 | 0.536 | 0.186 | 0.368 | 0.139 | -0.094 | -0.16 | 0.044 |
| WTF4 | 0.289 | -0.501 | 0.066 | 0.011 | -0.362 | -0.437 | -0.2 | -0.235 | -0.295 | 0.363 | 0.157 | -0.028 |
| WTM2 | 0.291 | 0.167 | 0.275 | -0.1 | 0.254 | 0.257 | 0.137 | 0.164 | -0.787 | 0.029 | 0.068 | 0.04 |
| WTM3 | 0.292 | 0.167 | 0.22 | -0.226 | -0.023 | -0.192 | -0.173 | 0.456 | 0.207 | 0.049 | -0.057 | -0.681 |
| WTM4 | 0.286 | 0.175 | -0.053 | -0.707 | -0.162 | 0.302 | -0.124 | -0.158 | 0.239 | 0.255 | 0.09 | 0.312 |
| TRF2 | 0.288 | 0.169 | -0.434 | 0.315 | 0.161 | 0.096 | -0.629 | 0.065 | -0.082 | 0.141 | -0.358 | 0.106 |
| TRF3 | 0.284 | 0.172 | -0.494 | 0.246 | 0.07 | 0.026 | 0.315 | 0.064 | 0.094 | 0.2 | 0.64 | -0.135 |
| TRF4 | 0.283 | 0.159 | -0.406 | -0.193 | -0.318 | -0.253 | 0.308 | -0.159 | -0.195 | -0.504 | -0.33 | -0.078 |
| TRM2 | 0.287 | 0.158 | 0.243 | 0.238 | 0.152 | -0.057 | 0.453 | -0.304 | 0.217 | 0.446 | -0.455 | -0.011 |
| TRM3 | 0.296 | 0.162 | 0.361 | 0.278 | -0.123 | 0.245 | -0.258 | -0.511 | 0.122 | -0.411 | 0.254 | -0.164 |
| TRM4 | 0.292 | 0.166 | 0.273 | 0.153 | -0.024 | -0.424 | -0.004 | 0.379 | 0.186 | -0.206 | 0.127 | 0.61 |

PCA was performed in R:
https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/prcomp

## Table S9. Data from competition assays.

### N1 transgenic line

| Replicate | Brood # | + | - |
|---|---|---|---|
| 1 | 1 | 11 | 34 |
| 1 | 2 | 21 | 21 |
| 1 | 3 | 0 | 48 |
| 1 | 4 | 0 | 36 |
| 1 | 5 | 30 | 52 |
| 1 | 6 | 10 | 61 |
| 1 | 7 | 0 | 69 |
| 1 | 8 | 0 | 28 |
| 1 | 9 | 0 | 23 |
| 2 | 1 | 0 | 94 |
| 2 | 2 | 0 | 36 |
| 2 | 3 | 0 | 75 |
| 2 | 4 | 32 | 31 |
| 2 | 5 | 0 | 3 |
| 2 | 6 | 0 | 74 |
| 2 | 7 | 34 | 30 |
| 2 | 8 | 13 | 20 |
| 2 | 9 | 25 | 31 |
| 3 | 1 | 0 | 68 |
| 3 | 2 | 16 | 19 |
| 3 | 3 | 0 | 53 |
| 3 | 4 | 0 | 74 |
| 3 | 5 | 15 | 53 |
| 3 | 6 | 29 | 33 |

| | | | |
|---|---|---|---|
| 3 | 7 | 24 | 22 |
| 3 | 8 | 0 | 75 |
| 3 | 9 | 0 | 42 |
| 4 | 1 | 0 | 100 |
| 4 | 2 | 0 | 106 |
| 4 | 3 | 0 | 84 |
| 4 | 4 | 19 | 28 |
| 4 | 5 | 6 | 7 |
| 4 | 6 | 13 | 17 |
| 4 | 7 | 28 | 20 |
| 5 | 1 | 4 | 3 |
| 5 | 2 | 0 | 69 |
| 5 | 3 | 34 | 28 |
| 5 | 4 | 29 | 17 |
| 5 | 5 | 58 | 31 |
| 5 | 6 | 16 | 15 |
| 5 | 7 | 15 | 19 |
| 5 | 8 | 0 | 52 |
| 5 | 9 | 14 | 13 |
| 5 | 10 | 0 | 21 |
| 6 | 1 | 0 | 41 |
| 6 | 2 | 12 | 18 |
| 6 | 3 | 36 | 38 |
| 6 | 4 | 0 | 19 |
| 6 | 5 | 13 | 10 |
| 6 | 6 | 30 | 28 |
| 6 | 7 | 0 | 22 |
| 6 | 8 | 0 | 75 |
| 6 | 9 | 0 | 70 |
| 6 | 10 | 16 | 43 |

**N2 transgenic line**

| Replicate | Brood # | + | - |
|---|---|---|---|
| 1 | 1 | 23 | 58 |
| 1 | 2 | 0 | 71 |
| 1 | 3 | 0 | 46 |
| 1 | 4 | 0 | 13 |
| 1 | 5 | 0 | 28 |
| 1 | 6 | 0 | 23 |
| 1 | 7 | 34 | 40 |
| 2 | 1 | 0 | 94 |
| 2 | 2 | 0 | 58 |
| 2 | 3 | 0 | 44 |
| 2 | 4 | 37 | 41 |
| 2 | 5 | 43 | 42 |
| 2 | 6 | 49 | 42 |

| | | | |
|---|---|---|---|
| 2 | 7 | 0 | 57 |
| 3 | 1 | 0 | 79 |
| 3 | 2 | 35 | 29 |
| 3 | 3 | 1 | 4 |
| 3 | 4 | 27 | 42 |
| 3 | 5 | 51 | 55 |
| 3 | 6 | 0 | 11 |
| 3 | 7 | 40 | 37 |
| 3 | 8 | 0 | 72 |
| 4 | 1 | 60 | 67 |
| 4 | 2 | 0 | 19 |
| 4 | 3 | 30 | 37 |
| 4 | 4 | 0 | 30 |
| 4 | 5 | 34 | 21 |
| 4 | 6 | 20 | 30 |
| 4 | 7 | 4 | 10 |
| 5 | 1 | 8 | 10 |
| 5 | 2 | 3 | 0 |
| 5 | 3 | 0 | 24 |
| 5 | 4 | 49 | 45 |
| 5 | 5 | 12 | 9 |
| 5 | 6 | 4 | 10 |
| 5 | 7 | 32 | 20 |
| 6 | 1 | 0 | 62 |
| 6 | 2 | 38 | 36 |
| 6 | 3 | 0 | 85 |
| 6 | 4 | 43 | 39 |
| 6 | 5 | 0 | 1 |
| 6 | 6 | 44 | 28 |
| 6 | 7 | 0 | 13 |

Note that females that did not lay eggs were not included in the analysis.
+: number of positive or transgenic progeny
-: number of negative or non-transgenic progeny

**Table S10. Proportion of females that produced broods that contained transgenic progeny** [a, b, c].

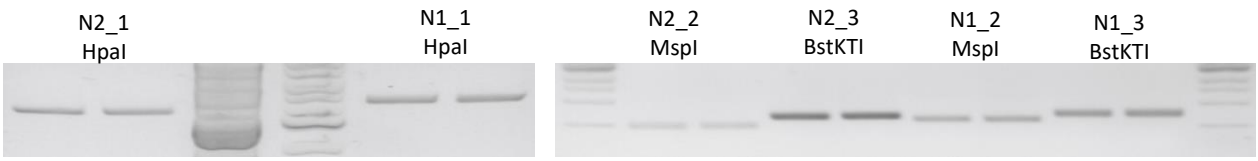
| Line | Brood with transgenic progeny [b] | Brood with no transgenic progeny | Observed proportion [c] |
|---|---|---|---|
| N1 | 28 | 26 | 0.519 |
| N2 | 24 | 19 | 0.558 |

a. Females that did not lay eggs were not included.

b. Females that produced broods that contained transgenic progeny must have mated with a transgenic male. However, these females may or may not have mated exclusively with a transgenic male. Although there is evidence to support the assumption that females mate only once in laboratory cage studies when mating is not interrupted (1), there are other reports that show evidence of polyandry (2). We should note that the analysis of the proportion of transgenic progeny (Table 2) does not assume monogamous mating or monogamous sperm contribution.

c. Overall, the percent of females that produced broods that contained transgenic progeny was 51.9% and 55.8% for the N1 and N2 lines, respectively. For reasons described above, these percentages do not necessarily mean that the transgenic males are more competitive in mating than their non-transgenic siblings. Single mating is an underlying assumption for the null hypothesis that the proportion of females mated with a transgenic male is 50%, which needs to be tested.

**A**




**B**

>N1_inverse_PCR_sequence (raw Sanger reads):

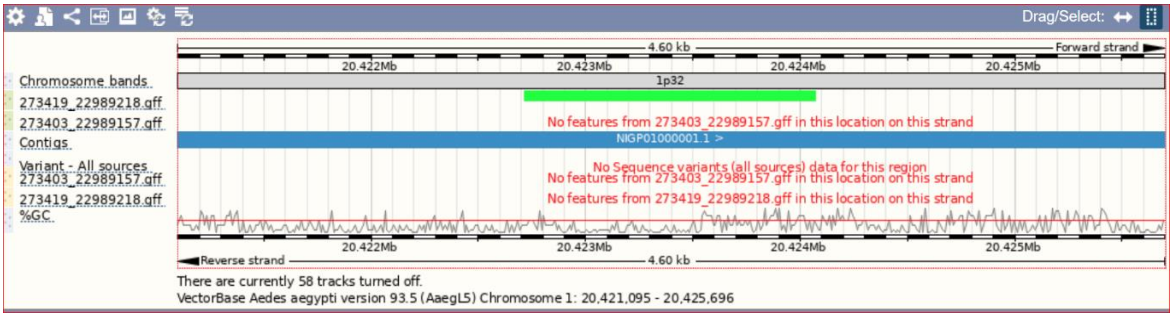
TTATCGTCAGGAAAGTTTGTGTTCGCATCTAGAACGCGTAAGATCCCCCAGATCTAAGCATCTTTTACCGCGACAGCGACAAAAAATCGACGCGATTTCGCTGATTTACGCTACATGCA
CTGTTTTATGACCACCGGATAAAACGTATCATTCAGTGATGATAACATTATGTACATATTACGTATTGGATACATACAGCGTATGTAAGTGAATGTTCGAAGCAATATTATTCTTGTTT
GGATAACAATTCAAAACATAATAATATATTGTAACAGAACATTGTATTGTTTTTAATTGGTTTATACCTTACAATTTTGGTCAAATTCCTATTTTCGAGTAAAACAACTGTTGCTTTTT
TCTATGTAGCTTTGCTGAAAAAAATAAACAAAACAAGTTCAGAAAGCAAGAAATGTTGGGTGAAAAATATTCAAAAAGTAAAAATAAGTAGTTTTTTAGAAGTCACTTGACATTAGCTG
TAACGACGAATGCAATCATGATACAGTTCGTTGAATTGTTTTTGTATTGTACAACAATACACGAATTGCTAATCAAGACAATACATGAACAATATATCGTATTGTATTTTCAAAATGTT
TGTATGAGAAAATATCTATATTTGTATTGTTACGATACTTAACCACCCATACATTATATTGCCAAAATCTCATATACAATACAGTGAACTGTTCAAAACAATATATTGTACTGTAATTG
TATTGTCATTGTACAATATACTGTATTGTATTTCCAATATATTGAATGGTACTGTTGCAATACGCTATATTGTTTTTGTATTGTATTTTTTATCCGGGCATCCAAGCCACAGCGACGCG
ATTTCGCAGCGATTCGCGTCGTGTCAGTCAAGATGGTTTCACAGAAGAGCGCTTGCAGCGACACAACAATAAAATCGCGACGATTTCTGGATATGATGATCGAGCGTGAGGTAATTGAA
AGGTGCATACAGCACCACGATGAACACTTGAATGGCCTTGAGGTAGATAACCAAGATGGCAGAATGCACAGCTTTATGTCGCTATGGTGATAACACTTTAAAAGTCATGGTTATTCAAG
TGTGTAGTGTCGTAATTTGGAGGAATAGATCCACCGGCCGAACTCTCATTTTCTGCTGCGCGTCCTTGAGCCGAAAGGACGTACGCGCCAGCCGAACTATCGCCGGCTGAATTTTCGGA
AAACTGTTCCATTTTAATGACCGTCCGTGTTCGCTATATTCGTAAGTCTTCTTCTTCTTTCTTGGCGTTACGTCCCCACTGGGACAGAGCCTGCTTCTCAGCTTAGTGTTCATATGA
GCACTTCCACAGTTATTAACTGAGAGCT**TA**TCAGGTGTACAAGTATGAAATGTCGTTTTTTTAAATCAAAAAACACGTTAAATTTTGAAAGTTTAGGTCC

>N2_inverse_PCR_sequence (raw Sanger reads):

GACTCGGTATTCTTAAGAACGAGCCTTGAAAAGGTTTTTCTGTGAACCTGTCACATTTTTTTCATGTAATTGACCTTACTGTTCCTAGAAAACTTCTAGCTACCCACGCATCAAATAAA
TTAAATAACGAAAAAGGATAATATTTAGGTTCTTTTATCAAAGGATACAGCTTTTTTCTTAGTGATTCGGATTATCCGGAGGTGAAAAAAATCTATAACTCGGATAATCGAGTCGGACC
TGTAAACATATACTCATCATAATCTTCAGCGTATCACAGAGATCATCATTATGAGTACATCGCACAGTGAGTCTTCCTCGTTACGAGTCGTACAAAAGTAAAAAAATGAGTCGTGAAAT
AGATTTATCAGTGTGGAGTGGATACGAGTAATTTGGAATAATGTATGAATAAATCTTCTTCTTCTTTCTGGCGTTACGTCCCAACTGAAATAATGAACTGAAATATGAATAAATAT
GTTCTTCTATTGATACTCTAGCGTTTGTCACTAGTAGATTTGGGGTTGTTGCATCTGATACCGTTCTTAGAAATGATCCAGCACGTCACAATTTTGAGCTACATGTTGCTAAGGTTCTA
TTATTAAGTCTTTATTCGATCAACTACCCTAAATATGTGTCTAAATTGTAATAATCTAAGTGCACCAAATGATTATTACGGAGAGCATGGGCACCATTCATTCCAAGATCTTAATAATG
ATGTTGATTTCTTTCTAAATCATCCTCATTGTCATTCTTTGATACCTTCAATTCAGTCGACGATAATGCCAGCTAAAGAGGGAGTCCGATCGCTTGATCGCGTATTTTGCTTCACGCAT
GTGTGACTCTTGCGTAAACATAATCTTGCATGATTTGAAGAATGTATTTGACAAGTATATATGCCCACGATAGTCCAATGCAACTGCCACCGTTGGAGTATGGGACACAATACTATTTG
GAAAATAGCCATTCAACAACTTTCCCACATAAATATCAGATGATGTGTTTATGGGCTTAATCATTTTATCATTTATAATTTTGCCATGTTGAGCGATTCTAAAATAGCAACCATATTGC
ATAATTTACCCATTTGTATAATTTTCTCATTTGGCTACTTCTATTTTTCGGATGGAACCAAATTGTCTACCACCCGTAAACAGTCATCTGGTCTATATTTGCCATCCAATGTGGATCGA
TTTGTTTGAATTGCGTTTCTTATTGCAGCCTGCCAGTGTATTGAAGGGCGTCATAATATATTCTAAAGAAAACAATTATCAAGGTCGTCTTCGTTTGAGATATTTTCCGGTTTATTAAA
TTGGAGCTTCTCATTAGCATGCGTGTGTTGAACGATATCAATCGATGCGCTCGAATAAAATTA**TA**TCAGGTGTACAAGTATGAAATGTCGTTTTTTTTAAATCAAAAAACACGTTAAAT
TGAAAGTTGGGTCC

**C**

**Map of N1 insertion site sequence to ~20.4 Mb on Chromosome 1**




**Map of N2 insertion site sequence to ~67.56 Mb on Chromosome 2**

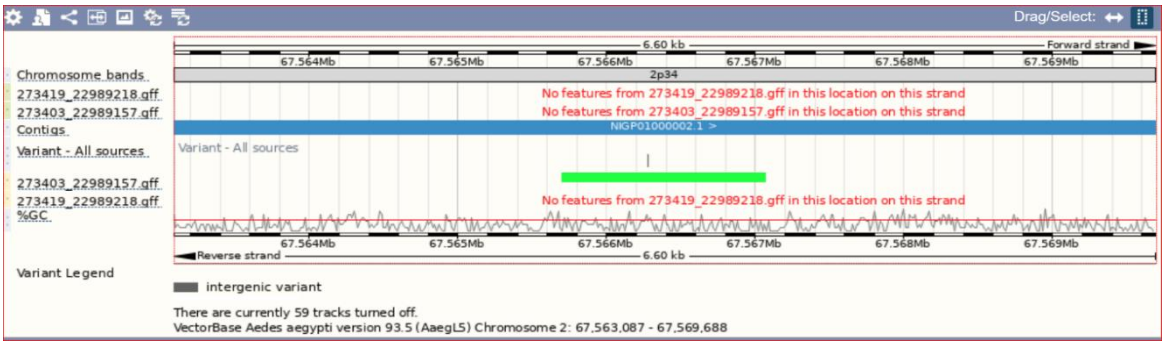


**Figure S1.** Inverse PCR to identify transgenic cassette genomic insertion sites. **A)** 1% agarose gel image showing PCR products resulting from PCR on ligated genomic DNA isolated from transgenic individuals. Genomic DNA of three separate individuals (_1, _2, and _3) from each transgenic line (N1 and N2) were digested with 3 restriction enzymes HpaI, MspI, and BstKTI. **B)** Genomic insertion site sequence identified by cloned and sequenced PCR products. For each transgenic line, all 3 reactions with different restriction enzymes resulted in identical sequences for the same genomic loci. "TA" target site (red highlight), and 5' end of transgenic cassette, including the Mos1 Right Arm (Yellow highlight), are indicated. **C)** Mapping of transgenic insertion sites to the AaegL5 Genome (Vectorbase.org) by BLAST of sequences of cloned inverse PCR products.

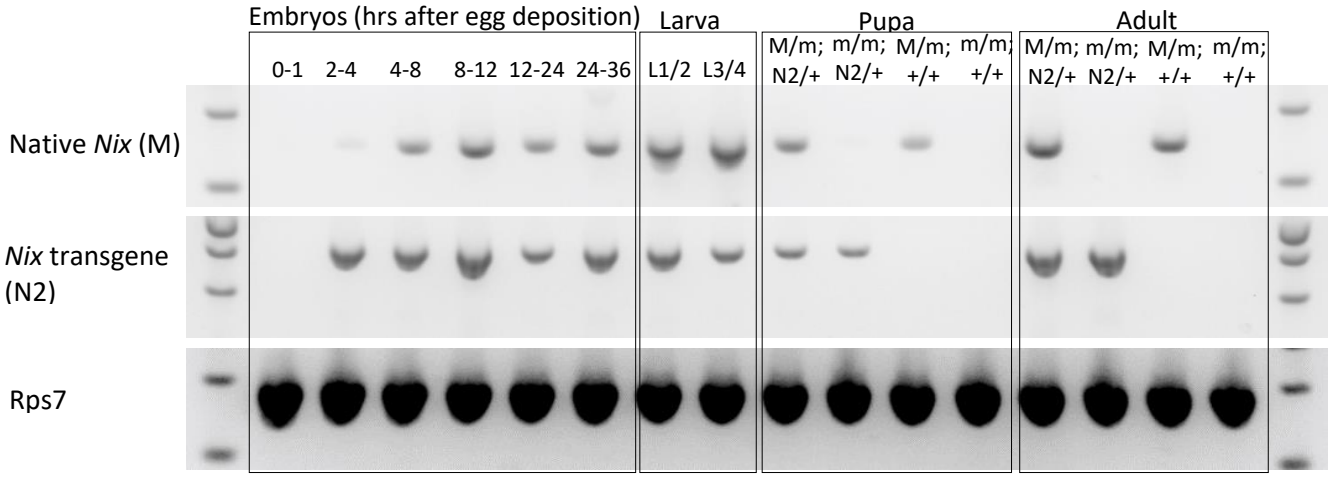



**Figure S2. RT-PCR showing the transcription profile of the N2 transgene and the endogenous *Nix* gene.** L1/2 indicates mixed sex first and second instar larva. L3/4 indicates mixed sex third and fourth instar larvae. Genotypes of the pupa and adults was determined by the presence and absence of the EGFP transformation marker and by the presence and absence of endogenous *Nix* and *myo-sex* (Figure S3D). Note that RT-PCR does not specify spatial information of gene expression.

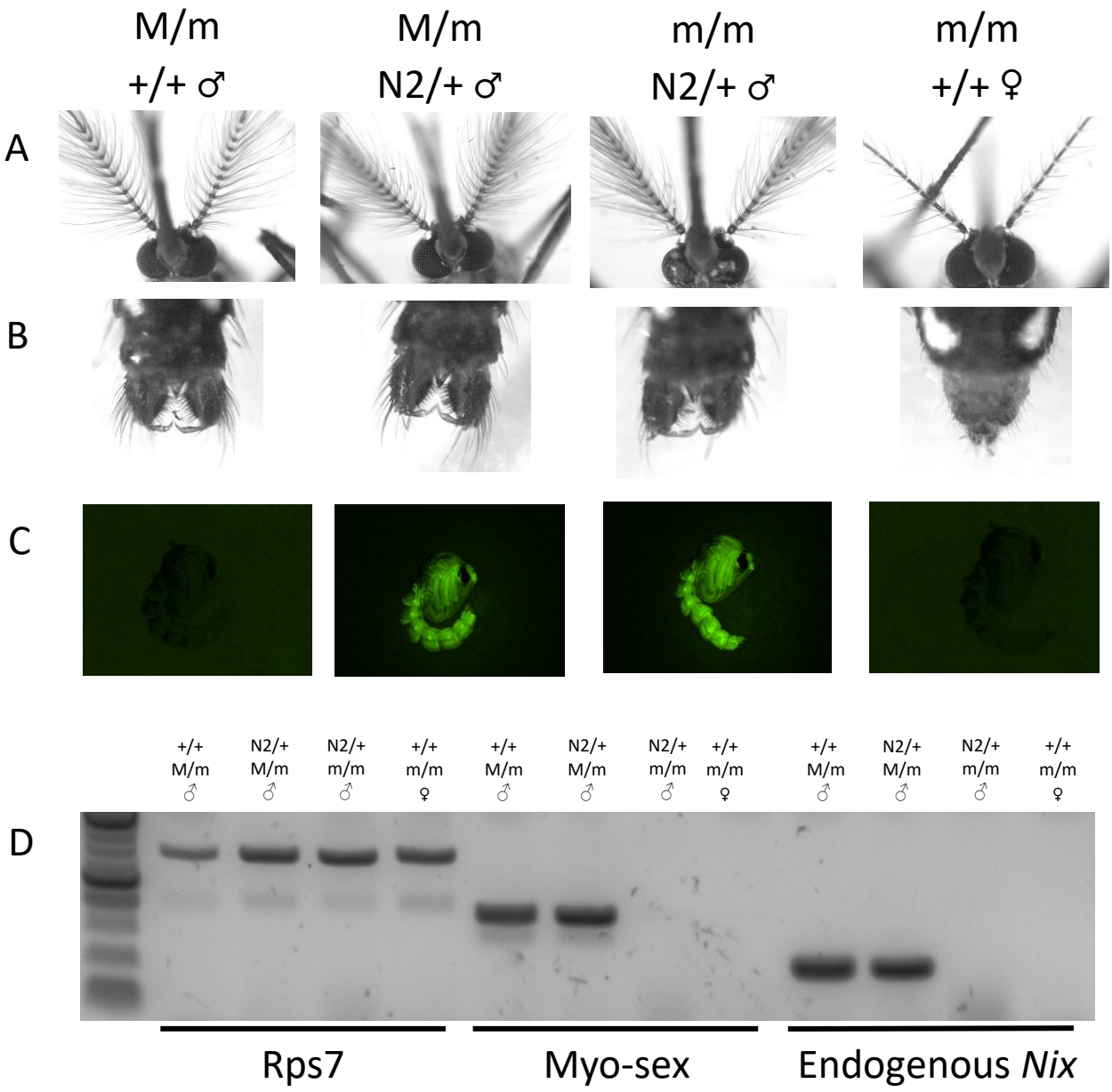



**Figure S3. Images of sexually dimorphic features of all four genotypes and methods used to determine the genotypes.** White light images of heads (A) and genitalia (B) of the four phenotypes observed in N2 transgenic *Nix* lines. ♂/♀reflects the observed phenotype. Images taken using the LAS V4.5 software suite with zoom set to 4 and the following settings: Gain 1, Gamma 1, greyscale (white light) or pseudocolor (509nm). C). The presence of a *Nix* transgene (N2) is determined by the presence EGFP marker. +/+ indicate wild-type or the lack of a transgene. D). The absence of the M locus, as indicated by the absence of two M locus genes *myo-sex* and the endogenous *Nix*, is used to indicate the m/m genotype.

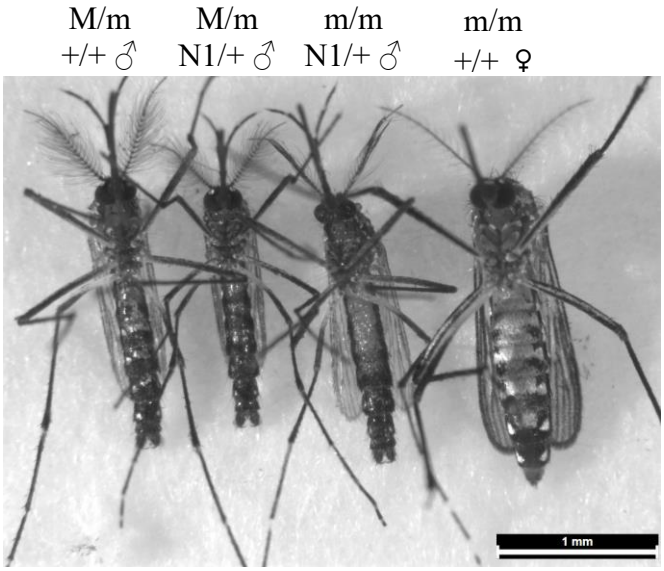



**Figure S4. Full body images of the four phenotypes observed in N1 *Nix* transgenic lines.** Representative individuals of each phenotype for the N1 line. Images were taken using the LAS V4.5 software suite with zoom set to 1, and the following settings: Gain 1, Gamma 1, greyscale

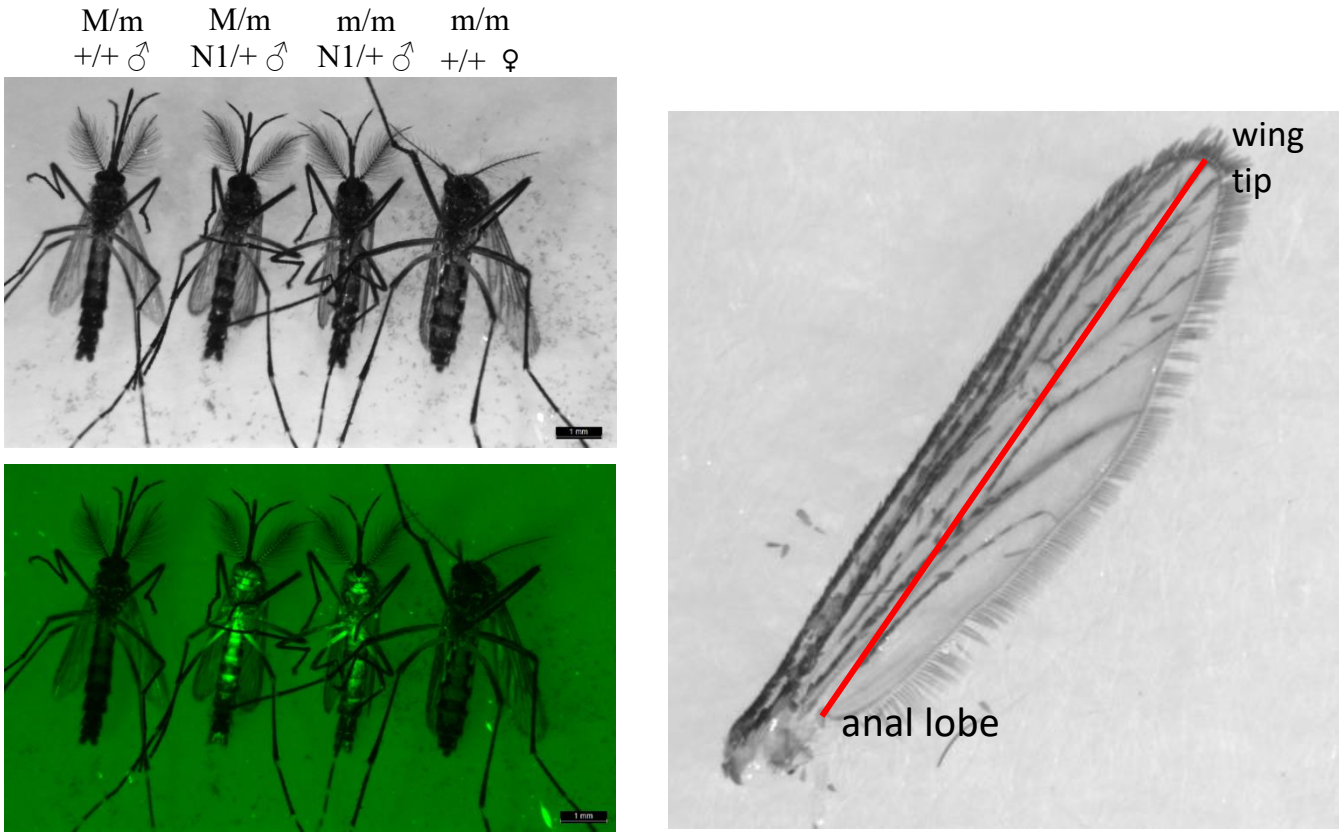
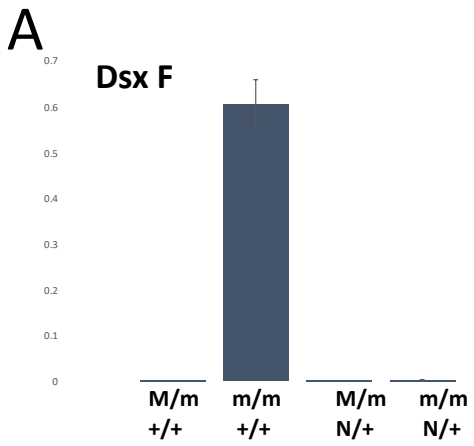
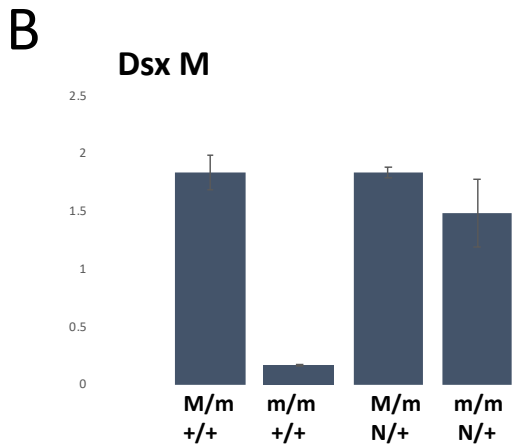
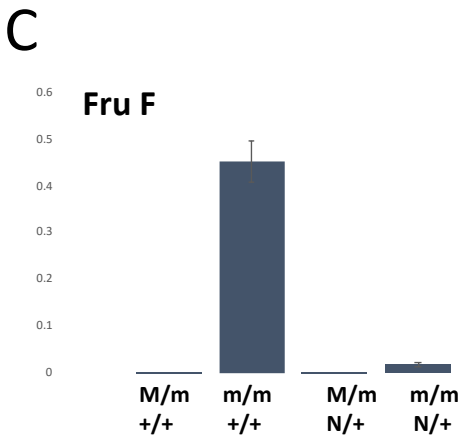
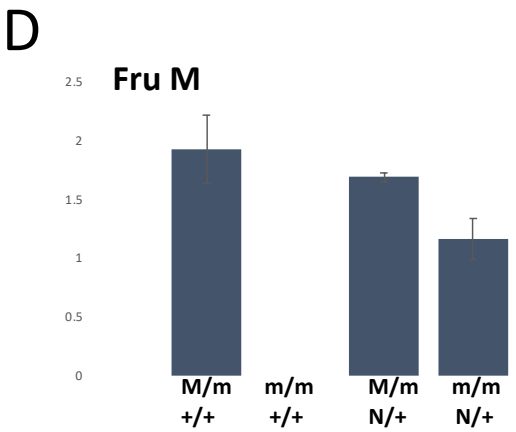



**Figure S5. Method to measure the wing length of adult mosquitoes**. Top left panel is the black and white image of the four genotypes to be measured. Bottom left panel is the florescent image of the same individuals. The right panel illustrates that the measurement is taken from the anal lobe to the wing tip.




E

| | Dsx F | Dsx M | Fru F | Fru M |
|---|---|---|---|---|
| (m/m; N/+) vs (M/m; +/+) | ns | ns | ns | ns |
| (m/m; N/+) vs (m/m; +/+) | **** | ** | **** | * |
| (m/m; N/+) vs (M/m; N/+) | ns | ns | ns | ns |
| (M/m; N/+) vs (M/m; +/+) | ns | ns | ns | ns |
| (m/m; +/+) vs (M/m; +/+) | **** | *** | **** | *** |
| (M/m; N/+) vs (m/m; +/+) | **** | *** | **** | ** |

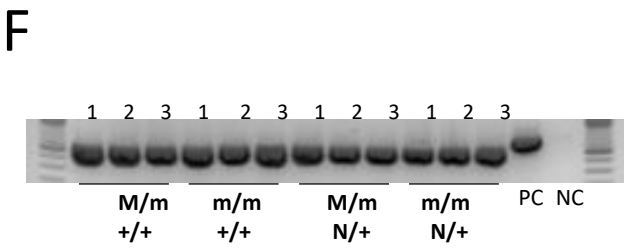



**Figure S6. Digital droplet PCR assays for the N1 transgenic line show *doublesex* and *fruitless* isoform levels in sex-converted flightless males (m/m; N/+) are similar to wild-type males (M/m; +/+). A)** *Doublesex* female isoform **B)** *Doublesex* male isoform. **C)** *Fruitless* female isoform **D)** *Fruitless* male isoform. Values shown are the mean +/- SEM. The gene AAEL002401 was used as an internal reference for gene expression (3, 4). **E)** Summary of statistical significance from all pairwise comparisons using a one-way ANOVA followed by Tukey Simultaneous Tests of Differences of Means ($p > 0.05$, ns; $p </= 0.05$, *; $p </= 0.01$, **; $p </= 0.001$, ***; $p </= 0.0001$, ****). Assumptions of equal variance are met. **F)** RT-PCR to check the quality of cDNA used for *doublesex* and *fruitless* assays, using primers for ribosomal protein S7. Three individual adult mosquitoes (3 biological replicates) were used for each genotype. Positive control (PC) and negative control (NC) templates were *Ae. aegypti* genomic DNA and $H_2O$, respectively.

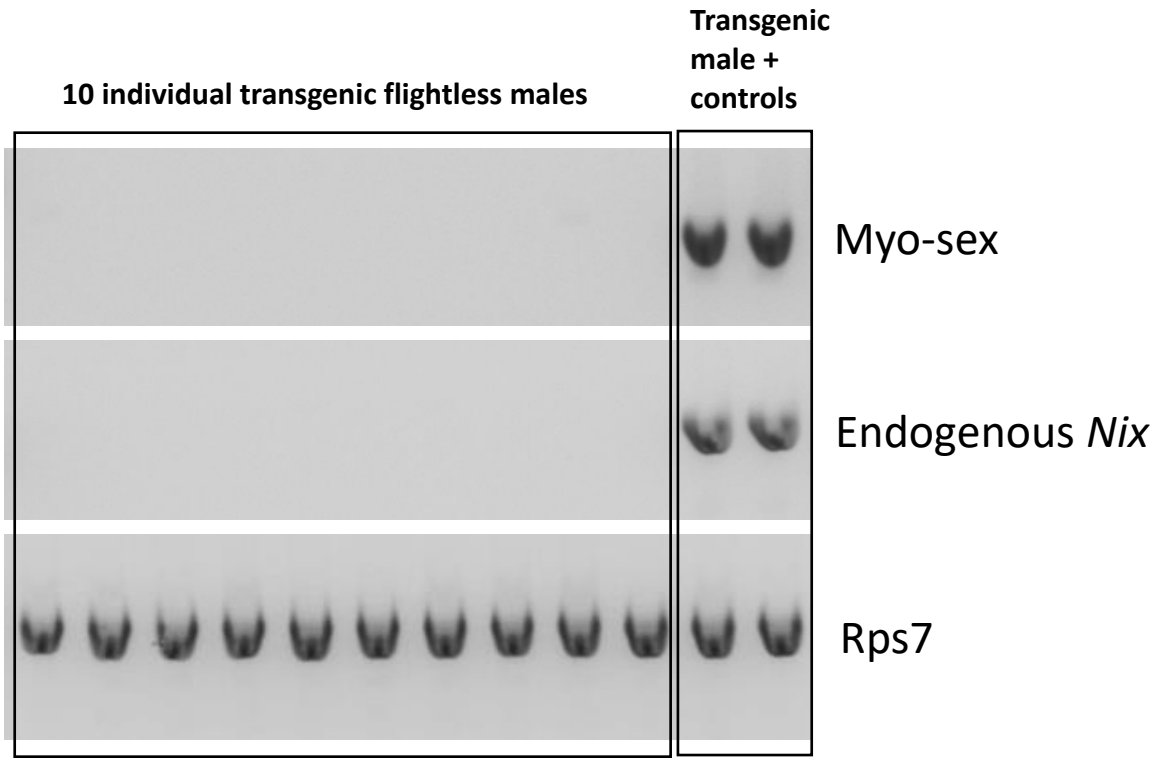



**Fig. S7. PCR assays used to determine that the flightless males in the N2 line are genetic females (m/m) that lack the M locus.** These flightless males are progeny derived from forced mating. The absence of the M locus, as indicated by the absence of two M locus genes *myo-sex* and the endogenous *Nix*, is used to indicate the m/m genotype. Rps7 is used as the positive control. Size markers are not shown.

## SI References


1. D. O. Carvalho *et al.*, Mosquito pornoscopy: Observation and interruption of Aedes aegypti copulation to determine female polyandric event and mixed progeny. *PloS one* **13**, e0193164 (2018).
2. M. E. Helinski *et al.*, Evidence of polyandry for Aedes aegypti in semifield enclosures. *The American journal of tropical medicine and hygiene* **86**, 635-641 (2012).
3. A. B. Hall *et al.*, A male-determining factor in the mosquito Aedes aegypti. *Science* 10.1126/science.aaa2850 (2015).
4. W. Hu, Z. J. Tu, Functional analysis of the promoter of an early zygotic gene KLC2 in Aedes aegypti. *Parasit Vectors* **11**, 655 (2018).