Supplementary materials for

Extensive recombination suppression and epistatic selection causes chromosome-wide differentiation of a selfish sex chromosome in *Drosophila pseudoobscura*

Zachary L. Fuller[1, 2*], Spencer A. Koury[3*], Christopher J. Leonard[3], Randee E. Young[3, 4], Kobe Ikegami[3], Jonathan Westlake[1], Stephen Richards[5], Stephen W. Schaeffer[1], & Nitin Phadnis[3**]

[1] Department of Biology, The Pennsylvania State University, University Park, PA 16802.

[2] Department of Biological Sciences, Columbia University, New York, NY 1002.

[3] School of Biological Sciences, University of Utah, Salt Lake City, UT 84112.

[4] Current address: Department of Genetics, University of Wisconsin, Madison, WI 53706.

[5] Human Genome Sequencing Center, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX, 77030.

* These authors contributed equally.

** address correspondence to:
Nitin Phadnis, 257 South 1400 East, School of Biological Sciences, University of Utah, Salt Lake City, UT 84112; tel: (801)585-0493 email: nitin.phadnis@utah.edu

Supplementary text: Sequences at the breakpoints of the basal and medial inversions

1. Basal inversion, proximal breakpoint
XR_group6:1747336-1747636-BREAKPOINT-XR_group6:4381020-4380840

GCACGTCACMCAATACAAACACACAAGCAGTTATTCGACTAACAAATCACTTTCTTTTCACA
TTTCACTTGTGCTTGGCTGTACATTCGATTACATGGGTATTTTGTTGCTCAAAACTTGTATTT
ATTTCTTTATATAGTATACATTACAGTGGTTCAAGTGCATATTTTTTTTGTGTGCGAATATTTT
ATTTGCAAAAAAGTAGAGATGGAGAAACGTCGATTTTACTATCGATAAATACTCAGCCGGCT
GGTAGCGATGTCTATCGATCCTACGCTGCCAGATCGAGAGAAAACGGTTACCTATATCTTA
ACCATCTGTATGGTTAGAAACAAGACAAGACAAGGCAAAACTTTCCCATCTCTAAGAAAGTA
TAATATGATATTTACAAATTCTTTGAAAAAGATAACTTCAGTGCGTACCAAACACTAATTGAC
TTGATTGACGAAATTGATGTAAAACTATAACCGGTACTCCAACACCTTAAAAAAAGCGGGC
GGTCAAGTAATCCCACGCTACTCAATGGTAGCATGGTAACTTAAGCCCTC

2. Basal inversion, distal breakpoint
XR_group6:1747856-1747736-BREAKPOINT-XR_group6: 4381024-4381264

ATAGGGATGTTAAGTATTCATTGATATCAATTATATGTATATGTTTTGCATATCATCATCGAT
ATTTTTGTGCGATAAATATTATTTCGATTGAAATTGTAAATACCTAATGATGTATTTCAATTAT
TTTCGCGGTATATTCTGCTGATTTTTTGTATTTTAAGGCACTAGCACCACTAACCATATCCA
GTGGGATATATACTGTATATATATATATACCATACAGTATACTGTATGGTTAGTGGTGGTAC
AATGGCGCTATCGATACTATCGCATGCATCACCTATGTTACTAGGAACTACATACATTTTTT
CGAGTGTGTGCCTATCAAATGCGTTACCGTCTGTTGATTCTTTAATCTTTAAGTTTAGTTTTA
TTTGGGAAGCACGGTTTGTACGATGTTGAGCGCAAAAGCGAGAGAAAACCTCCATCGAAT
GACTCTCCCGCGTTATGTAATTAAACGCATTGTTCAGTTGCTCCTCACTCAACTGG

Repetitive motif at breakpoint: CCATACAGTATACTGTATGGTTAGTGG

3. Medial inversion, proximal breakpoint
XR_group6: 5055855-5056104-BREAKPOINT-XR_group8: 706130-705579

ATTACCGGAGAATATTTAAATTTAATACGTTCCAGCGATAAATATACCGATTCGAGTAATAG
CGAGTTGGATAAAATAATAACACAATTAGAAATGTGTATATCTATGTGTTTTTTTATTTTAAG
CCTATTTTTTGGCCTTTTTCAATTTTTAGCTTGTTTTAGCTTGTGTTTTTGTGAAATCTAGCTT
ATTTTTACGCTCAAAATCTGGCAACACTGCTGCAGTAGTGACCGCGGATTTATCGATAACG
ACCATCAGGAATATACTGAAACATACCATCTCATTTTAAAAATATACCGTAAATATACTGACG
AATTCAAGTTCTATTTTACTTATTTCTCGTTTTTGATATTTCGTCGAATATTACCAGCTATATA
GAACATTTAGCCATGCCCACTCAATTTTGTACGATTGATGAATCAATTCTATACATGATTGG
CCTATTCGAAATACTTGCTTTTATTGGATTTTGGCTAAAACAAGGCATAAACAAAATAGTTCA
ACAAAAAGAAACAGTGTTAAGATAGCAAGGTGTATTAATATAAGGTAACGTGTACAAGGTGA
ACAGTGTGTGCACTCCAATGGTTATCCTTTTTTAGAGGATTATCGATTTTCGATAAAAACCTT
TACTGGTCTTTTTCGAACTCAAATCAAATTTTAATATATAAATTGCTAGTTAACTTTAAATTGT
TGAACTTTATTTCTGTTAGCTTACAATATACATAATTCAAAATGTTTTAGGCCCACATACGAA
TGAAATGTGGGAAAATTGGTTTTACATGAATAGAATTCACCATTTTATAAATTCTTACAATTT
TCGTGTTTTCTTTTTTGTTTTTTCTTCACTTAACCTTAACTGGAACCGAGAGCGAAAGTAGAA
AAACATTCAATTTCTTAAACAGTAGCGATTTTTTGTTGTAGTTAAATATTCGTGTATGTCCCG
CCTAGGAATTCTTCCGATATTATAACTTTTTGAGAGAGATTTAGATCAAATAGTAGACATACA
GATACATATACATATACATATATTTCATATATGTATGTGTATATA

Repetitive motif at breakpoint:
TCAGAAATATACCGAAATATACCATCTCATTTAAAAAATATACCGTAAATATACTGACGAATT
CAAGTTCTCTTTTACATATTCCTCGTTTTTGATCTTCCGTGGAATATTCTCAGCTATATAGAA
CATTTAGCCATGCCCACTCAATTTTGTACGATTGATGAATCAATTCTATACATGATTGGCCT
ATTCGAAATACTTGCATTTATTGGATTT

4. Medial inversion, distal breakpoint

XR_group6: 5058390-5057728,5057645-5056460-BREAKPOINT-XR_group8: 706177-708097

GTTGTACTATAATATATATTGGCAGGCCTTTTGGGTTCGGCATTCGATTCGCATACCCCATA
ACCATAGCCATAACCATACTATATCTTATGGCTGCCATTAAGTTTCATTATTTCTGAATTTGT
TTATCCGTACAGGACATCAATCAGACACGCCTGCCCCTCATTCCCGAGTATGCCGTGCAGA
CGCTTACGCCGCAGGAAATGCTTCAGGTGGAGCAGCGAATGGATCAGCGGAGGGAGCGA
CTGAAGGACAAGTGCTCGGCCTATGGCCTGGATGTGTTAGGTGGGTCAATGGGCCACAGC
CACAGAGAGAGATAGATAAATAAATAGAGAGAGAAAGAGAGGAAGAGTGAGAGTGTGAAA
ACGAACTTCTCCTTCACATCAATCTTATTTACAGGTCACGACTCGTGGCACACCCCAAACAC
ATGGGAGTTTTTGGTCAACAAAAAGTATCACATTATATGGTGCGTATACAAATTACACATCA
CATCACATTCACAGGAACTCCTCCACTAAAGCCACGAATCTTGCGTAGGTGCAATGTGTTT
AAGGCTGCCTCCTCATCGTGGATGTTCAACTTCAATGTTCTGGCCGGTTACTCACCCAGTT
ATTTGCGCAAAACCAAAAAGATTCTCCTGAATCTGGCCAGAGAACGCTATCCGAGAGTGAC
CCTTGACGAGGTGAGTGCGGAAATGGAACGGAAACTAGCATCATACATTTTTTATACCCGA
TACTCAAAATGAGTATTGGGGTATATAAGATTTGTGGTAAAAGTGGATGTGTGTAACGTCCA
GAAGGAATCGTTTCCGACCCCATAAAGTATATATATTCTTGATCAGCATCAATAGCCGAGTC
GATAGAGCCCTGTCTGTCTGTCCGTCCGTCCCCTTCACCGCCTAGTGCTCAAAGACTATAA
GAGCTAGAGCAACGATGTTTTGGATCCAGACTTCTGTGATATGTCACTGCTACAAAAATATT
TCAAAGAGCAACCAAATTTGGTATCCACACTCCTAATATATCGGACCGAGACGAGTTTGTTT
CAAAATTTCGCCACACCCCCTTCCGCCCCCGCAAAGGATGCAAATCTGGGGATATTCACAA
ATCTCAGGGACTATTAAGGCTAGAGTAACCAAATTTGGTATCCGCACTTCTGCTAGATCTCA
CTATAAAACGTATATCTCAAAATTTCGCCCCACCCCCTTCCGCCCCCACAAAGGACGAAAA
CCTGTTGCATCCACAATATTGCACATTCGAGAAAACTAAAAACGCAGAATCATAGATAATGA
CCATATCCATCAGATTGCTGAATCTGGATCACATCAGACAATTTTTATAGCCAAAAGGAACA
AATCAATTTGCACTGGCTACGCAGCGCCCGACGTCACGCTCAGACTGATTTTCTGTCTCTC
TCGCACGCACTCTTTGTCGTGTCGTTTAATATTAGTGGCGTCTGCCGGAGGAGAGCCATAC
TGACTTAGTATCGGGTATAACTGTAGAGTTGCGGTGTCCGCTCATAACTCATAACGTTCCC
CCTCGTTTTTATACCCGATACTCAAAATGAGTATTGGGGTATATTAGATTTGTGGTAAAAGT
GGATGTGTCTAACGTCCAGAAGGAATCGTTTCCGACCCCATAAAGTATATATATTCTTGATC
AGCATCAATAGCCGAGTCGATTGAGCCCTGTCTGTCTGTCCGTCTGTCCGTCCGTCCGTCT
GTCCGTCTGTCCGTCCCCTTCAGCGCCTAATGCTCAAAGACTATAGAGCTAGAGCAACGAT
GTTTTTGGATCCAGACTTCTGTGATATGTCACTGCTCGAAANATATTTCAAACTTTGCCCGC
CCACTCCGTCCCCCACAAAGGGCGAATCTGTGCATCCACATTTCGACAATACGAGAAAACT
AAAAAACGCAGAATCGTAGAAGATGACTATATCTTACAGAGTGCAAAATCTGAACCAGATC
GTATAATTATTACAGCCAGAATCACGAAAACAATTTCACTCTTTCTCGCTCTGTCTCACTCTA
ACACACAGGTTTCATGGTCGGTTTTGCCAATTTCAAAATATGAGTTCAAGGATCTCAGAACC
TATAAAAGCCAGAGCAACCAAATTTGGTATCCACACTCCTGTGATATCGGACCTTGACCGT
TTCATGTCCACATTTCGCCACATCCCCTTCCGCCCCCGCAAAGGACGAAAATCTGAGGCAA
CCACAAATCTCAGAGACTATTAAGGCTAGAGTAACCAAATTTGGTACACACACTCCTTTAAG
ATGTCACTATAAAACGTATATCTCAGAATTTCGCCCCACCCCCATCCGCCCCCACAAAGGA
CGAAAATCTGTTGCATCCACAATATTGCAGATTTGAGAAAACTAAAAACGCAGAATCATAGA
TAATGACCATATCTATTAGATTGCTGAATCTGGATCAGATCAGATAATTTTTGTAGCCAAAA
GGAACAAATCAATTTGCATTGGCTACGCAGCGCCCGACGTCACGCTCAGACTGATTTTCTG
TCTCTCTCGCACGCACTCTTTGTCGTGTCGTTTAATATTAGCGGCCTCTGCCGGAGGAGAG
CCATACTGACTTAGTATCGGGTATAACCGTAGAGTTGCGGTGTCCGCAGCAACTCACAACG
TTCCCCCTCGTTTTATGTATCTCGTATGACATTTTTGCTTTTCGTTTGCAGCTGCGCGAGGC
GCAGAATGACTCGCTAACCTTTATCATTGCACGCAATCCCTTTGAGAGGCTGTTGAGCGCT
TATCGCGATAAGATGGTGTTCGCCCTGCCCTATTCCTTCCATGACAAGCTGGGCCGCAGC
ATTGTGCGCAATTATCGCAAGAAGGTAAGAAGGCAATCAAGACTGACGGAAACCTCCGAAC

TATGCCTATCTATCATATATATCCTGACAGCCTTCCCTGGTGGCCCGAGCGCCCAACACCA
AGTATCCTTCGTTTCCGGAGTTTGTCAACTGGCTGCTCGACCAGGTGAAGCGGGGCAGCT
TCATTGACATGCACTTTGTGGCGGCCACGTCGTTCTGCACACCCTGTCTGATTCGTTTCGA
CATGATACTGAAGTTCGAGTCGCTCACAGAGGATCAATTATATCTAATCGAAAAGACTGGC
CTGAAAAGGGTGATAGCGCCCGTGTGGCGCAACATGGGCAAAGGTGGCCGCAAGACGCA
CGAACTCCAGCAGCAGTTCTATGCCCAGCTCACAAGGCACGAAATGCTGGAGCTCTACGA
GTATTACAAGTGAGTGGGACAATCTAATACCCCTTTTGATAGTGCTCTAACAGTCTTTGTTC
TATTCCAAGATATGATTTCGAGCTGTTTGACTACGATATCCAGGAGTATCTACAGGTAGCCA
GACCAGACGAATCATCCGGCAGTCCAACGGCCACAAAGAATTAACCCAGCTCTAAATTCGT
AGAATTTGCTAACGAAATGAATTAGTTCAACGAAAATACATATTAGCCGTCACAGTACGTAA
TTGCAAATGTTGATAATACTGAATGAATACAATAATAAACTTAAATATATTTAAACGAAACGT
TTTTACCGTCTATCGTTTTTTGGGCTGCTATCGATAAGCGGGATAACATTATATCGAACAAA
TATCGGAACAAATATCGTTTGGGCCCATTTGGCATATTTGGAATTAGTGTGCCACAACAACT
TTCTTCTGTTGCTAATTCCATGCACACATACCCTGGGGGGAAATGAATGTTAAAAATTGAGAA
AATGTTTGTCTTCCAAGGCTCAGTAGGTATCAGGATCGCTATAAATATATACAAGATACTAA
TAATGTACTTGAATATGTCTAAAGAATATCAATTTGAGAAAAGGTCTTTATAAATCACGTTTC
ATCTTCATATAAAATTAACAAAATAGGGTATACAAAGGCCAACACGCAACATAGAGACTCTT
TTTTGTTGGACTTTTTCGTTTAAACCTTTTTTTACAATAAATACAATAAAAGCAAGGATTAGG
TTAGGATAGGTTCAGGCGGTAGCCTGGGGGAGTGATGAAACCCTCCCAAGCTCACTTGGA
CCTGTAGAAGGTCCGTTGTGGTGCCGCATGAGCGTGTGCCCCTTCCCAGTGCCTAAAAAC
CGTGGAGAAACAGGCTGTTGCAAATTAAGGGCAGTCCCATCCGGAGGCTTGGATGAATCT
GGACAAGGTCCCCGGTTTGATTTCGGACAGGTCCGAAATGTCTGTGAGGAAAGCAGCCCC
GAGAAAACAAAGACGACGATTTCCAAGGGCCGAGCAGTGTCAGAAGAAGTGGGGGACCG
ATTCCTCCTCTTCCTCGTCGCGACAACTCCTGCAGAAGTCGTTATGAGGGATTGACAGTCT
AGAGGCATGAGTGCCAATCTGCCAGTGTCCCGTAATGGCACGAGTAACTGCAGAGCACTG
TGCTCTGCTGAGCCTATACAGCTCGGCTGAGCGCTTCTTCGATCGATATGGCCATATCAAT
CTTGAGATTTTACAGGAAATCTCCTATTGGCATTTTGCTCGAACAATTCGTGCGTGAAGAGC
CTGCACGTAGCCAGAGGCATCGCTACCTGCTCCCTCTCCATAAGGAGAGGAATTGTAGTA
CCCTGCCTTGCTAGCTCGTCGGCGGCGTCGTTCCCCTCGATGTCCCTGTGGCCTGGGACC
CATATAAGGAAGAGGTCTAACTGATCTGCGATCTCGTGCAGAGACCTGCGGCAGTCATTTA
CAGTCGCTGAGTTCGACGATATTGAGCCTAGAGCTTTAATAGCTACTTGGCTGTCCGAGAA
AACGCACACTAAGTGGCGGTNGTCTNAGAAGTAGTTGCAGACGATGGGAATGGGCCTCAC
TCCNACAGCGTCTGCGGGAATTAACCTGAATGCCTATAGCAACATCGTCCGCATACGCCAC
CACACGGCAGCCACCCCCCTCTATCTCCCGCAGCAGTTCGTTCACTGCCACGTTCCATAG
GAGGGGCGAGAGGACTCCGCCCTGCGGGGTGCCTCTACTGACAAATCTGGTGGACGTTG
ACGTCCCCAGTGATGCCTCAACCGTCCTGCATTGTAGCATCTGATCGATCAGTTTCACCGT
CCGGGAGTCAACCCCCAGGTCCGTCAGTGCACCCGTGATGGCGGTCGGGAGGATGTTAT
TAAAGGCGCCTTCTATGTCGAGGAAGGCTACCAGGGTGTACTCCTTACAGTTGAGGGACG
CTTCTATGATCGATGTGATCCTGTGTAGGGCCGTTTCGGTGGATCTTCCTTTCCGATAGGC
ATGCTGGGAGTCTGAGAATAGACTAGGCGGAATATTTACCGTGAGATGCATCCCCAAAAGC
CGTTCCATCGTCTTCAGAAGGAAGGACGATAGACTTATGGGTCTGAAATCTTTGGGGGCAG
TGTGCGAGGGCTTGCCTGCCTTGGGGATGAAGACGACTTTGGGATGAAGACAGGTTTCCG
GGATATATCCATGGGAGAAGATTCCCTCGTATATCCTTTTGAGCCAATTAGTGGCTTNTTGT
CCAGCGTGAATCAGTTAGGGCAGGGATGATGCCATCTGGTCCCGCAGACTTGTATGGTTT
GAAGCTTTTTATTGCCCGGGCTCCTGGCTAAGCAGGTTCATGATAACACTTGAGGCAACGG
AGGGAGCCGCGATGTAGTGTGGTCTGTTTTCATCGCAGCCGGGAAAGTGGGTGTTAAGGA
GAAGATTTAGCGACTCATTGCTGGACGTTGTCCATGATTAGTCGGCGTTCTTCAGGTAGCC
CAGAGTGGGTGTAGTCTTCGAGAGAACTTTGCGCAAGCGCGAGGCTTCCGAGTTATTCTC
AATGTCGCTACAGAACTTACGTCATGAGGCGCGTTTGGCTTTCCTCAGTTCTTTATTGTAAA
CTGACAGACTGGCCATGTAATCGGCCCAGTTCTGCTCAGCGTTTTCAGCCTTATCTTTATT
GAAGAGTCTTCGGGAGGCTTTCCGGAGTTCCCTCAGTTTCGGATTCCACCATTCAGGTTTC

TTCCGGCCTCTGCTCTTGCTAGAGGGGCACGATCTATCGAGCGCGTCATTGCAGGCGTCG
GTACATAGCTTAACAAGATGAGTCGTGGCTTCCCTGGAGATCTCCTCCTCAGGGAGAACAC
GACAGAGGTGGTCTGAGTACCGAGCCCAGTTTGTCCGCCTAAGGTTCCGAAATGAGCTGG
CCTCGGTAGTGAAAAAGAGAATACGGTTTCCACGTACCTGTGGTCCGAGAAGGAGTGCTC
TTCCAGGACCCCCCAGGATACGATGTTCCCGTGTAACTCATGAGAGGCAAGTGTGAGGTC
CAGGACCTCCTTGCGGTTTTTAATAATGAAGGTGGGATCACTCCCCCTATTTAAAAGGACC
ATCTGAGTGGTCAGTAAATAGTTGAAAAGGTACTCACCCCTTTCGTTGGTGTCAGTGCTTC
CCCACTGGCAGTGGTGTGCATTGGCGTCGCACCCTACGATTAGGCCGGTGTCCTTGGCCT
CGCAGTCTATGATTAGCTACGTATAGCTTATAGTCAGGAGTCCTCAGACCAGAGACCCTGT
TTCCGAGGATCCAGGGCTCCTGAATGAGGACTATGTCGGCTCCACCCTTGGNCTAGNGTG
GNGCAGGAGAGCAGCGCATGCTGCCTTACAAAGGGTGGNGGTTTATCTGCAGGAGTATCA
GTGACATTCCTGAGGTTCACCTCCACCATTGTCCTGTCGTGGTCGCTTTCCGAAGAGTCAA
ACTCCTCCGCGACCCCGATGTGCTTGAGATCCCTAGTAAGATCGCTCACGTCTGACACGTA
ACCATTAAGGATGTCATCCGACACCACTGATGCCACGTCTGACACCGCAGGCTTGATCTCC
ATACTAGGGATCTCCGGGAGCTCCAGGTCTGGCTCGACCGTCGGTTGCATGCCTTTGGAG
TCGGACTTGTACGGTACTACGACTACCTTCTTGTAGCCGTAGTCCCACAGCGCCCTCCGTG
TCGTGGAGAAGTCTGACCGACTCCTCGTTCAGGACGAGGATGATTTGGCGCCTCTGCCCG
TTGCTCTCCTCTACCTTGGACACCTNCCAATCGGCTGTGGGAAGGTGGGGGTTGCAGACC
TGCAGAATCCGGAGAATCTTATCCGACTCTGCANNNCTTCGCCGAGACCCACGCTCTGGC
CCTCGGCTTGCTTCACTTCACGACCTCCAGCTGGGCTCCTGGGTAGACCTCCTTGAGCGA
CGCTACCGCCTTCGCGTAGAGGGCCGCCGAGCGAGCATGTTGGCAGGCGATGGCTTTCA
CCCTGCCTTGGTGCCACCCGGCGTCCTCTCACTTGGGCGGTGGTCCTCCATTCTCCTCCA
ACTCCTGGAGGAAACAGTCTTGGAGCTCGTTTTCCACGAGGTGCCACTTGTCCCTAGGGA
TCTGCCCGTCTTTAGAGCCCCTGTCTAGGACTCCAATCAGGGTTTTTCCCCTCGCCACCTC
GGCAAACGAGCGGTTGCTCAGTGTCTTCGTCCTCTTGGCCTGTTGGACTTGTGTGGCGGT
GTCCTCCGCCAGCTTTGGCTGGCTGGTAGTCGGGCATGACTGTCTTGGCCCATCGCGGC
GATTCGATTCCCTGGGCAGACGCCCGGAACACCGGGTCGTCGTTGCCCAGGATGAAGGC
CGCACGTCTTTTGTCCTGGAACTTGGGCCTATCTTTCGAGGCAGAGGCGCCGCCTGCCGG
GGTTGTCAAGGGGTTCCCGGTCCCTAGGGGTCCGCCAGCCGCGATATCGCTCTGCATGAT
CGCCCCCCCCGGTAACCGGGTCGCCCTTGGGGCCCCCGATGTGGATTGGCCCGATTGGCT
TTTCGGTCCTCTCCTCAAAGCGCCTGCTGCCTCGAGACGGTGGACCGACTTAGTCTCCTTC
CCCGTCTTTTTGGGCACTGGTTTCCAGATGCGTTGTAGAGGGATGGTCTTTCCTCCGGCAC
TTAGGAGGAGTGCCGTGCTCCTTGCGGTGTTTTTAGTTTTAATATTAAATTTGGCATATTTG
ATCCCACGAGTAGGCGGGAAGGGGTTAGTCGCCCGAGCATAGCCCGCGTGCCCCGGGAA
GCCTTATTAAAACTGGAGGTCGGCAGGTATCCAGAGTCCGCATACATACATATGTATGTAT
GTATGTATGTATAAGCGACCGAGCACCCCCTCGGCCATGCATCCCTCGGCATATGAGAGT
GTCACCTTGGATTGGGGTTATTTCACTGAGAGTTTTCTTCTCTCCGCCCGACTACAATTAGC
CAGTATCCTGGCAGAGACATGACCGTACCAGGACCTGTTTCCAGCCGCCGTCACGGAGAG
AGTATAGTCGCACTTCCGCCGGTGTGTACAGTTACGGCGGGTGCGGGTTCGCTCGTGGCC
ACCTGTATCCTATGACGCGGAGCACAGTCGCCTTGGATCCAGGGCAAGCCATGAACCAGA
GGCGCCTGTGCCCCGTTCCGAGTCTACAGTTGTGACGAGCCGACCCGACATCCGTTTATC
CCCCTGTAGCACCCGACGGCTGACGGCCAAAAGCAAGGATTAAAAACTAACTAATCTTGTA
GAAAATTGACTCATCAATGGAATAAAACTATGTGGGCATTGCTGAGAGATTTAGTTAGTCAG
CATTATTCAATGGAATATCACAATTGAGCAATATGAACGAATTTCATTTTTCGTTTGTCTTGA
TTGACCTATTTCGCTACTTGTTTTTTTGTTTGACTTATTTCGCTAAAACCCTTTTTCACGCAA
AACGCAATAAAAGCAAGTATTAAGAATACACCAGTTAAATAGAAAATTGATTCATCAATCGT
ATAAAATTATGTGGGCATGGTCAAATTTTCTATCTAGCTAATAATATTCCTCGGAATATCAAA
AACGAGGAATATGTAAAATAGAACTGGAATTCGTCAGTATATTTAAGGTATATTTTTAAAATG
AGACGGTATATTTCTAAGGGTCGGACGGTATATTTTAACGCGGTCAGACTGCAGCACTGCA
ACATCAACAAACGCCACGACCTATAAAAATTGTTAAATTAACTGACACCGAAAATGAATGAG
TAATTCAGTGAGAATCAATCGAAATTTATTAGCTAGAAGCATAGGCAAGGCATAAGCAAAAA

TACTTACTTGTTTGTGATTTAGAGAGCTTCGGCATTGCACTGACGTGGACCTGAAAGATTAG
AAACAGGTGAGGCGACATGTGCATGCATTATCTAGATTGAAATCAGTCTGGTACACCTTTT
GGCAGACTTCCAATATGTCCACAGCACGGCGGCGGCCCCAGGGTGGCAGTGCGGGCGAT
GTACAGAGCTACCACCCGAAGGAAAGGCTGGTGGAGAAGGACGAGGGAACGCCGGACAT
ACGCGGCGACAGACGGAACATAGCCATCCTGCTGTTCCTGTATATACTGCAGGGCATACC
CATCGGCCTGATAGCCGCCATTCCCATGCTGCTGCAGAACCGGGGAGCCAGCTACAAGCA
GCAGGCGGAGTTCTCCTTCGCCTACTGGCCGTTCAGCTTGAAGCTACTGTGGGCCCCCAT
CGTGGACTCGCTTTACGTCCGGAGATTCGGGCGCCGCAAATCGTGGCTGGTGCCGGTGC
AGTATCTGCTTGGCGGATTCATGATGTTCCTTTCCTACTACGTGGATCGTTGGCTGGGCGG
CGATGGTGTGGAGCCGAATGTGGCGCTGCTGTCGCTGCTCTTCTTCCTGCTCAATTTCCTG
GCTGCCACCCAGGACATCGCTGTGGACGCTGGCTCTGACTATGCTGAAGCGCTGCAATGT
AGGCTACGCCTCCACCTGCAACAGTGTCGGCCAGACAGCCGGCTACTTCCTTGGATATGT
GTGTTCATTGCCCTCGAATCGAAGGACTTTTGCAACAAGTACATGAGGGATGTGCCCCTGA
ACGAGGGCATGATCACGTTACCACGCTTCCTCTGGTTCTGGGGTATTGTCTTTGTGGTGGC
CACCACTCTGGTGGCAATCTTCAAGAAGGAGAATGACATTGAAGATGCCCATACGGAATCT
CGCTACACGGAGGAGCATGAGCTGAACATTCGTCAGAGCTACAAGATCCTCTGGGACATG
GTGAAGATGCGACCAGTGCAGATCCTAGCCGCCATTCTGCTCACCGTTAAGGTAACCTTCT
CCGCATCGGATGCGGTCACCAGTCTGAAGCTCATCGATGCCGGGGTGCCCAAGGATCAG
CTGGCTCTGCTGGCCATCCCCCTCATTCCCCTTCAGCTCGTCCTGCCCCTGGTGATGGGT
CGCTACACCAACGGCCCGCGTCCCATGGATGTGTATCTGAAGGCCATTCCATATCGAATTC
TTCTGGCCGCCGTGGCGACAATTTTCGCCTACGCCACACCTTTTATGGTGCAGAAAGGGC
ATGTCCCAGTGTACTATTATGTCCTTCTGATCGCCCTGTATGCCTGCTATCAGGTATTCCTG
TACTCAATGTTCGTGGCGGCCATGGCATTCTTTGCAAAGATCTCGGACCCCGCTGTGGGT
GGCACATACATGACATTCCTCAACACGCTGTGTAATCTGGGCGGCAATTGGCCCAACACG
GTGGTGCTCTGGCTGGTGGACGTGCTCACATGGAAACAGTGCACCACCAATACGGATAAT
ACGTGCCTCAATAAGGATGAGCAACAGGTACGAGAAGACTCTACCCACCTTCTGTGGTGTT
TTTCAACTAATCTAGTGCCCTCCCTCACTCCTTATAGAGCTGTGAATCGTCACACGGCAATT
GCGAGAT

Details and caveats for the crossing scheme used to generate *SR* and *ST* lines:
From the analysis of differential gene expression, a total of 868 significant genes were detected off the *X*-chromosome. Although these differences may be the result of *trans*-acting factors associated with genetic variation located on *XR*, our crossing scheme maintained lines through different sets of marker strains, possibly creating structured variation on the autosomes. Additionally, the *SR* and *ST* strains carry different third chromosome arrangements (Arrowhead for *SR* and Standard for *ST*) which are known to harbor an abundance of *cis*-acting expression differences. Indeed, a principal component analysis (PCA) of SNPs called in autosomal genes from the RNA-seq reads confirmed the presence of variation distinguishing *ST* and *SR* individuals (Supplementary Figure S5). It is therefore possible that these autosomal transcriptional differences may arise from *cis*-acting factors associated with structured genetic variation in the stock marker stains. As a result, autosomal comparisons to the *X*-chromosome were not considered for this study.

Additionally, to establish the *SR* stocks, lines were repeatedly backcrossed to *ct sd y se* marker strains which carry a *ST X*-chromosome. This backcrossing procedure was not used for the *ST* lines. As a result, recombination may have had a homogenizing effect on *XL* variation, potentially biasing patterns of F_{ST} and d_{XY} downward. While the exact extent of differentiation for *XL* between the *SR* stocks used here and natural populations is unclear, because of these potential biases we limit our comparisons between patterns of genetic variation on *XL* and *XR*. Moreover, our main conclusions regarding the age of the *SR* chromosome and the levels of differentiation and recombination suppression across the *SR* inversions do not rely on these comparisons.

Supplementary Methods and Results

Standard / Sex Ratio Females Fail to Produce Inversion Recombinants in cytogenetic mapping experiments. Two isofemale lines one each collected from two localities (KBPN2, Kaibab National Forest, AZ, September 2017 latitude: 36° 24' 42" N; longitude: 112° 18' 48" W; AO4, Allred Orchard Farm Market, Provo, UT September 2017, latitude: 40° 15' 43" N; longitude: 111° 39' 32" W) produced all daughters suggesting each female had mated to an SR male. We crossed the putative ST/SR daughters to a hemizygous male with a multiply marked *X*-chromosome. We individually crossed male offspring to ST/ST females (107 males from KBPN2 and 96 males from AO4). A single female larva (Figure S1) from each of the 107 KBPN2 and 96 AO4 male test crosses were karyotyped using standard cytogenetic methods (Painter, 1934). Specifically, we scored the XR chromosomal karyotype where all female offspring carried a ST chromosome derived from the F1 female parent plus the parental or recombinant chromosome from the tested male (Figure S2).

In the sample of 107 KBPN2 and 96 AO4 males from a ST/SR female, we found no recombinants among the three non-overlapping inversions. Thus, non-overlapping inversions are inherited as a single unit. The frequency of the ST gamete is not significantly different from the expected frequency of 50 % (see Table S9, KBPN2, binomial sign test P=0.281; AO4, binomial sign test P=0.063). The sex ratio for the ST arrangement was near 50% for both strains while SR males sired > 96 % daughters (Table S10). A t-test with unequal samples shows that the mean sex ratio is significantly different between ST and SR males (KBPN2, t=51.64, df=56, two tailed P =6.3 x 10^{-49}; AO4, t=33.73, df=91, 3.18 x 10^{-53}). Interestingly, the mean number of offspring produced did not differ between ST and SR males (KBPN2, t=0.34, df=100, two tailed P = 0.74; AO4, t=-0.00, df=92, two tailed P = 0.99). Despite a reduction in Y-bearing sperm, SR males produce a similar number of offspring as a ST male.

Random Union of Gametes Model of Linkage Disequilibrium Decay

The comparative genomics analysis of *SR* chromosomes revealed an ancient origin of the inverted arrangements with extensive genetic differentiation stemming from greater than two million generations without gene flow. Consistent with this work, natural population sampling indicates all three inversions are found in near perfect association ($r^2 = 0.998$, Table 1) (Nielsen and Slatkin 2013). However, direct experiments demonstrate the array of three inversions can be readily broken apart, which qualitatively suggests recombinant chromosomes should be more common in nature and genetic differentiation should erode in the ~5 Mb collinear region between the medial and terminal inversions. To quantitatively understand the interaction of low levels of recombination (on the order of 10^{-3} per meiosis) over long periods of evolution (on the order of 10^6 generations) we introduce a two locus, two allele population genetic model for the decay of gametic phase disequilibrium (Crow and Kimura 1970).

Consider a locus with two alleles (*A* and *a*, with frequencies p_A and p_a), a second locus with alleles (*B* and *b*, with frequencies q_B and q_b), and a recombination frequency between these loci denoted *c*. Let $P_{t(AB)}$ be the frequency of a haplotype carrying alleles *A* and *B* among the gametes produced in generation *t*. In any given generation, a particular copy of the *AB* haplotype can either be inherited intact from the previous generation with probability *1 - c*, or alternatively be the product of a recombination event with probability *c*. In the former case, the

frequency of the haplotype is equal to its frequency in the preceding generation denoted $P_{t\text{-}1(AB)}$. In the latter case, under the assumption of random union of gametes (*i.e.,* independence of egg and sperm allelic states in Hardy-Weinberg equilibrium), the probability of a haplotype stemming from recombination having alleles *A* and *B* is simply the product of their respective frequencies in the population ($p_A q_B$). Therefore, we can describe the frequency of the haplotype of interest $P_{t(AB)}$ in generation *t* as a function of its frequency in the previous generation $P_{t\text{-}1(AB)}$:

$$P_{t(AB)} = (1 - c)P_{t-1(AB)} + cp_A q_B \qquad \text{Equation 1}$$

After subtracting $p_A q_B$ from both sides of equation 1, rearranging, and iterating following Crow and Kimura (1970), a recursion equation for this model of gametic phase disequilibrium can be written as:

$$P_{t(AB)} - p_A q_B = (1 - c)^t \left(P_{0(AB)} - p_A q_B\right) \qquad \text{Equation 2}$$

where $P_{0(AB)}$ represents the initial *AB* haplotype frequency, and the left-hand side of equation 2 (the difference of the observed haploytpe frequency $P_{(AB)}$ from the expected frequency $p_A q_B$), is recognizable as the classical coefficient of linkage disequilibrium $D_{(AB)}$ (Ewens 2004). Finally, because recombination is restricted to females in *D. pseudoobscura*, the crossover rate in this population model must be corrected to one half the experimentally determined recombination frequency. In this way the frequencies of all four haplotypes (*AB, Ab, aB, ab*) can be predicted within a population at any future generation *t* given initial conditions P_0 with the series of equations:

$$P_{t(AB)} = \left(1 - \frac{c}{2}\right)^t \left(P_{0(AB)} - p_A q_B\right) + p_A q_B \qquad \text{Equation 3a}$$

$$P_{t(aB)} = \left(1 - \frac{c}{2}\right)^t \left(P_{0(aB)} - p_a q_B\right) + p_a q_B \qquad \text{Equation 3b}$$

$$P_{t(Ab)} = \left(1 - \frac{c}{2}\right)^t \left(P_{0(Ab)} - p_A q_b\right) + p_A q_b \qquad \text{Equation 3c}$$

$$P_{t(ab)} = \left(1 - \frac{c}{2}\right)^t \left(P_{0(ab)} - p_a q_b\right) + p_a q_b \qquad \text{Equation 3d}$$

Here we define locus *A* as the basal and medial inversions of *SR* chromosomes and locus *B* as the terminal inversion, we obtain frequency data for p_A and p_B from published data on natural population sampling (Table 1), and estimate recombination frequency from our direct laboratory experiments (Figure 5B; Supp. Table S10). This parameterization allows for analysis of the asymptotic approach to linkage equilibrium and the evolutionary trajectory of all four haplotypes: Standard arrangement ($P_{(ab)}$), recombinants ($P_{(Ab)}$) and ($P_{(aB)}$), as well as the fully intact *SR* chromosome ($P_{(AB)}$). Under this neutral model the half-life of LD is only 577 generations, with effective linkage equilibrium achieved within 10,000 generations. Furthermore, the near perfect association of SR chromosome inversions should have completely broken up long ago, F_{ST} in (as well as LD across) the collinear regions should be substantially lower, and recombinant *SR* chromosomes should be found at an average frequency of approximately 0.12 in present day natural populations.

By using the observed present-day frequencies of *SR* chromosomes and their rare recombinants observed in nature (Table 1) we can estimate the equilibrium haploid selection coefficients (*s*) associated with recombinants using a modified version of equation 1:

$$\widehat{P}_{(aB)} = (1 - s_{aB})\left(\left(1 - \left(\frac{c}{2}\right)\right)\widehat{P}_{(aB)} + \left(\frac{c}{2}\right)p_a q_B\right) \quad \text{Equation 4a}$$

$$\widehat{P}_{(Ab)} = (1 - s_{Ab})\left(\left(1 - \left(\frac{c}{2}\right)\right)\widehat{P}_{(Ab)} + \left(\frac{c}{2}\right)p_A q_b\right) \quad \text{Equation 4b}$$

Solving for s under these conditions yields s_{aB} = 0.316 for the terminal inversion only recombinant and s_{Ab} = 0.649 for the basal and medial inversions recombinant, respectively. The haploid estimates represent the extremely intense epistatic selection acting against recombinant genotypes and are consistent with the conditions determined by numerical analysis of the drive-selection balance model (viability selection assuming complete recessivity $s_{aB} > 0.01$ and $s_{Ab} > 0.29$). These theoretical considerations demonstrate the only way a two million generation old chromosome-wide coadapted gene complex can show no signs of gene flow given our laboratory estimates of recombination is to have strong epistatic selection against recombinants, not just in present day populations, but throughout the long evolutionary history of *SR* chromosomes.

Drive-Selection Balance Model of Linkage Disequilibrium Decay. In the previous supplemental section a general model for the decay of gametic phase disequilibrium is presented. The general model only treats gametic frequencies using the random union of gametes assumption under strict neutrality. Here we relax these assumptions by incorporating specific properties of *SR* chromosomes: *X*-linkage, male-specific segregation distortion, female-specific recombination, sex differences in allele frequencies, and enforcing equilibrium drive-selection conditions. The resulting model described below is substantially more complex, requiring treatment of additional parameters as well as explicit modeling of both genotypic and gametic frequencies for each of the sexes separately. However, the qualitative result of the general simple model in the main text and the specific complex model in this supplemental remain the same: the three-inversion state of *D. pseudoobscura* should have reached linkage equilibrium hundreds of thousands of years ago and can only be maintained by strong epistatic selection.

Consider a *X*-linked locus with two alleles (*A* and *a*, with frequencies p_A and p_a), a second locus with alleles (*B* and *b*, with frequencies q_B and q_b), and a recombination frequency between these loci denoted *c*. Let $P_{t(AB)m}$ be the frequency of a haplotype in the male gametic pool carrying alleles *A* and *B* produced in generation *t*. Additionally, let $G_{t(AB)m}$ be the frequencies of male hemizygous genotype carrying alleles *A* and *B* in generation *t*. Following this form, $P_{t(AB)f}$ is the haplotype frequency in the female gamete pool and $G_{t(AB/ab)f}$ would be the frequency of the female coupling phase double heterozygote at generation *t*. Here, we divide a generation into three stages: genotypic frequencies pre-selection ($G_{t(AB)m}$), genotypic frequencies post-selection ($G'_{t(AB)m}$), and then gametic frequencies ($P_{t(AB)m}$). These stages are connected by three processes: viability selection (s_{AB}) to transform pre-selection genotypes to post-selection genotypes, female-specific recombination (*c*) transforming female post-selection genotypes to female gametic frequencies, and male specific meiotic drive (k_{AB}) transforming male post-selection genotypes to male gametic frequencies. Finally, sex-specific gametic frequencies are used to calculate the next generations pre-selection genotypic frequencies. A diagram of this model is illustrated in Supplemental Figure S6.

For an *X*-linked locus, male hemizygous pre-selection genotypic frequencies are solely a function of female gametic frequencies in the previous generation:

$$G_{t(AB)m} = P_{t-1(AB)f}$$
$$G_{t(aB)m} = P_{t-1(aB)f}$$
$$G_{t(Ab)m} = P_{t-1(Ab)f}$$
$$G_{t(ab)m} = P_{t-1(ab)f}$$

In contrast, female genotypic frequencies are determined by both female and male gametic frequencies in the previous generation:

$$G_{t(AB/AB)f} = \left(P_{t-1(AB)m}\right)\left(P_{t-1(AB)f}\right) * 2$$
$$G_{t(aB/AB)f} = \left(P_{t-1(aB)m}\right)\left(P_{t-1(AB)f}\right) + \left(P_{t-1(aB)f}\right)\left(P_{t-1(AB)m}\right)$$
$$G_{t(Ab/AB)f} = \left(P_{t-1(Ab)m}\right)\left(P_{t-1(AB)f}\right) + \left(P_{t-1(Ab)f}\right)\left(P_{t-1(AB)m}\right)$$
$$G_{t(ab/AB)f} = \left(P_{t-1(ab)m}\right)\left(P_{t-1(AB)f}\right) + \left(P_{t-1(ab)f}\right)\left(P_{t-1(AB)m}\right)$$
$$G_{t(aB/aB)f} = \left(P_{t-1(aB)m}\right)\left(P_{t-1(aB)f}\right) * 2$$
$$G_{t(Ab/aB)f} = \left(P_{t-1(Ab)m}\right)\left(P_{t-1(aB)f}\right) + \left(P_{t-1(Ab)f}\right)\left(P_{t-1(aB)m}\right)$$
$$G_{t(ab/aB)f} = \left(P_{t-1(ab)m}\right)\left(P_{t-1(aB)f}\right) + \left(P_{t-1(ab)f}\right)\left(P_{t-1(aB)m}\right)$$
$$G_{t(Ab/Ab)f} = \left(P_{t-1(Ab)m}\right)\left(P_{t-1(Ab)f}\right) * 2$$
$$G_{t(ab/Ab)f} = \left(P_{t-1(ab)m}\right)\left(P_{t-1(Ab)f}\right) + \left(P_{t-1(ab)f}\right)\left(P_{t-1(Ab)m}\right)$$
$$G_{t(ab/ab)f} = \left(P_{t-1(ab)m}\right)\left(P_{t-1(ab)f}\right) * 2$$

The pre-selection male genotypic frequencies are transformed to post-selection genotypic frequencies in this model by selection acting on deleterious effects associated with *SR* and recombinant chromosomes with relative fitness of standard chromosomes set to unity and then adjusted by mean population fitness of males $\underline{w}_m$:

$$G'_{t(AB)m} = \left(G_{t(AB)m}(1 - s_{AB})\right)/\underline{w}_m$$
$$G'_{t(aB)m} = \left(G_{t(aB)m}(1 - s_{aB})\right)/\underline{w}_m$$
$$G'_{t(Ab)m} = \left(G_{t(Ab)m}(1 - s_{Ab})\right)/\underline{w}_m$$
$$G'_{t(ab)m} = \left(G_{t(ab)m}\right)/\underline{w}_m$$

The pre-selection female genotypic frequencies are similarly transformed in this model assuming deleterious effects of both *SR* chromosomes and recombinants are fully recessive:

$$G'_{t(AB/AB)f} = \left(G_{t(AB/AB)f}(1 - s_{AB})\right)/\underline{w}_f$$
$$G'_{t(aB/AB)f} = \left(G_{t(aB/AB)f}(1 - s_{aB})\right)/\underline{w}_f$$
$$G'_{t(Ab/AB)f} = \left(G_{t(Ab/AB)f}(1 - s_{Ab})\right)/\underline{w}_f$$

$$G'_{t(ab/AB)f} = \left(G_{t(ab/AB)f}\right)/\underline{w_f}$$
$$G'_{t(aB/aB)f} = \left(G_{t(aB/aB)f}(1 - s_{aB})\right)/\underline{w_f}$$
$$G'_{t(Ab/aB)f} = \left(G_{t(Ab/aB)f}\right)/\underline{w_f}$$
$$G'_{t(ab/aB)f} = \left(G_{t(ab/aB)f}\right)/\underline{w_f}$$
$$G'_{t(Ab/Ab)f} = \left(G_{t(Ab/Ab)f}(1 - s_{Ab})\right)/\underline{w_f}$$
$$G'_{t(ab/Ab)f} = \left(G_{t(ab/Ab)f}\right)/\underline{w_f}$$
$$G'_{t(ab/ab)f} = \left(G_{t(ab/ab)f}\right)/\underline{w_f}$$

In this model, female gametic frequencies are a function of recombination (*c*) and mendelian segregation acting on female genotypic frequencies; in contrast, no recombination occurs in males and the male gametic frequencies are dictated by the action of meiotic drive (*k*) as determined by male genotypic frequencies. Therefore, male gametic frequencies, including Y bearing sperm $P_{t(Y)m}$ are:

$$P_{t(AB)m} = (k_{AB})\, G'_{t(AB)m}$$
$$P_{t(aB)m} = (k_{aB})\, G'_{t(aB)m}$$
$$P_{t(Ab)m} = (k_{Ab})\, G'_{t(Ab)m}$$
$$P_{t(ab)m} = (k_{ab})\, G'_{t(ab)m}$$
$$P_{t(Y)m} = (1 - k_{AB})\, G'_{t(AB)m} + (1 - k_{aB})\, G'_{t(aB)m} + (1 - k_{Ab})\, G'_{t(Ab)m} + (1 - k_{ab})\, G'_{t(ab)m}$$

Female gametic frequencies based on mendelian segregation and recombination in both coupling and repulsion phase double heterozygotes are:

$$P_{t(AB)f} = G'_{t(AB/AB)f} + \tfrac{1}{2}\left(G'_{t(aB/AB)f}\right) + \tfrac{1}{2}\left(G'_{t(Ab/AB)f}\right) + \tfrac{(1-c)}{2}\left(G'_{t(ab/AB)f}\right) + \tfrac{c}{2}\left(G'_{t(Ab/aB)f}\right)$$
$$P_{t(aB)f} = G'_{t(aB/aB)f} + \tfrac{1}{2}\left(G'_{t(aB/AB)f}\right) + \tfrac{1}{2}\left(G'_{t(ab/aB)f}\right) + \tfrac{(1-c)}{2}\left(G'_{t(Ab/aB)f}\right) + \tfrac{c}{2}\left(G'_{t(ab/AB)f}\right)$$
$$P_{t(Ab)f} = G'_{t(Ab/Ab)f} + \tfrac{1}{2}\left(G'_{t(ab/Ab)f}\right) + \tfrac{1}{2}\left(G'_{t(Ab/AB)f}\right) + \tfrac{(1-c)}{2}\left(G'_{t(Ab/aB)f}\right) + \tfrac{c}{2}\left(G'_{t(ab/AB)f}\right)$$
$$P_{t(ab)f} = G'_{t(ab/ab)f} + \tfrac{1}{2}\left(G'_{t(ab/Ab)f}\right) + \tfrac{1}{2}\left(G'_{t(ab/aB)f}\right) + \tfrac{(1-c)}{2}\left(G'_{t(ab/AB)f}\right) + \tfrac{c}{2}\left(G'_{t(Ab/aB)f}\right)$$

Finally, although the action of the *SR* chromosome alters the sex ratio in progeny, we do not model non-Fisherian sex ratios at the population level. Therefore, at the beginning of each generation the male gametic frequencies of *Y*- and *X*-chromosome bearing sperm are adjusted to a 50:50 ratio by dividing by $P_{t(Y)m}$ and $\left(1 - P_{t(Y)m}\right)$, respectively, before being used to calculate genotypic frequencies in the next generation

To explore the decay of linkage disequilibrium on the *SR* chromosomes of *D. pseudoobscura* we use published data, our own experimental results, and preliminary observations on the behavior of recombinant *SR* chromosomes. The published record of species-wide *S*R frequencies weighted by intensity of sampling is 0.135 (Table 1 main text), and our model is initiated with complete absence of recombinants. In the

absence of recombination, the strong drive (k_{AB} = 0.99) of *SR* chromosomes requires strong counterbalancing selection to prevent rapid fixation. Fitness defects must be present in both sexes, and upon the assumption of complete recessivity, a selection coefficient of s_{AB} = 0.431 yields a stable equilibrium at the natural population frequencies.

In the presence of recombination, the stable equilibrium established by conditions k_{AB} = 0.99, s_{AB} = 0.431 rapidly breaks down. This is because the basal and medial inversion carrying recombinant goes to fixation in under 100 generations, with a LD half-life of only 47 generations. In agreement with published accounts, we observe the recombinant *SR* chromosomes with only the basal and medial inversions drive, while the recombinant *SR* chromosomes with only the terminal inversion segregates according to mendelian ratios. However, in preliminary analyses the drive of the recombinant chromosome is weaker, $k_{Ab} \approx 0.75$ (S. Koury personal observation), and this reduced drive is used in our model. To discover the conditions under which recombinant *SR* chromosomes will not accumulate in nature despite our laboratory estimated recombination rate of c = 0.0012, we performed numerical analysis to identify to intensity of selection on recombination. In this model, where selection against the fully intact SR chromosome is s_{Ab} = 0.431, additional selection with $s_{Ab} > 0.29$ and $s_{aB} > 0.01$ is required to prevent establishment of recombinant *SR* chromosomes in natural populations.

Table S1: *D. pseudoobscura* reference alignment statistics.

Statistic	***D. pse ST***	***D. pse SR***	***D. mir***
Total Reads	139337556	145236658	49649299
Mapped Reads	133512488	139245407	45556023
% Mapped	95.82	95.87	91.53

Table S2: *D. pseudoosbcura* reference alignment statistics for genomic scaffolds.

Scaffold	Mean Coverage	Length	% Covered	Covered (bp)	+ Reads	- Reads	Read GC	Median Coverage	St. Dev. Coverage
D. pseudoobscura ST									
2	83.9129	30819483	97.5788	30073269	13841040	13841340	0.4403	85	40.51
3	85.9616	19787792	97.5288	19298803	9058245	9057166	0.4549	86	62.53
XL_group3a	44.7765	2692213	96.4761	2597342	642496	643846	0.4455	43	67.92
XL_group3b	54.3022	388551	98.341	382105	112865	113304	0.4266	45	122.17
XL_group1e	42.4095	12541198	97.8198	12267770	2878802	2874461	0.4426	42	26.38
XL_group1a	45.7357	9148293	96.8705	8861993	2257478	2255726	0.4523	42	81.27
XR_group3a	40.5871	1469181	97.7692	1436406	324078	324013	0.4676	40	16.46
XR_group8	41.4123	9197557	97.3256	8951575	2056823	2057328	0.4571	41	20.03
XR_group6	42.6727	13333775	97.5782	13010864	3061992	3063993	0.4477	42	34.97
XR_group5	51.3656	740970	98.7634	731807	203154	203069	0.4408	44	49.9
4_group1	84.4553	5287126	98.1752	5190646	2402053	2400617	0.418	84	201.76
4_group2	81.5053	1235759	94.4519	1167198	534344	533721	0.4381	85	62.7
4_group5	80.7469	2439919	93.5238	2281905	1054607	1055036	0.4205	86	36.3
4_group3	83.7848	11685562	97.9635	11447582	5288757	5290633	0.438	84	38.18
4_group4	85.1283	6594820	96.8235	6385336	3016316	3013210	0.4307	86	58.84
D. pseudoobscura SR									
2	85.6435	30819483	97.5044	30050367	14142848	14139655	0.4399	87	39.7
3	87.4484	19787792	97.5056	19294208	9232111	9227540	0.4545	87	61.44
XL_group3a	46.4184	2692213	96.4888	2597685	666643	668042	0.445	45	70.15
XL_group3b	55.1312	388551	98.2872	381896	114618	114974	0.4253	46	102.94
XL_group1e	43.8398	12541198	97.8332	12269454	2978752	2973251	0.4421	43	27
XL_group1a	46.3822	9148293	96.8344	8858691	2295757	2290161	0.4523	43	92.78
XR_group3a	44.1206	1469181	96.2329	1413835	356858	355122	0.4634	42	60.14
XR_group8	42.891	9197557	96.4071	8867101	2138713	2136598	0.4547	42	33.05
XR_group6	44.1551	13333775	96.608	12881487	3185838	3187677	0.4467	43	50.19
XR_group5	47.4404	740970	96.7479	716873	188068	188616	0.4374	45	39.7
4_group1	84.1611	5287126	97.9979	5181270	2393844	2394017	0.4193	87	60.4
4_group2	83.6336	1235759	94.4369	1167013	548437	547715	0.4372	87	58.68
4_group5	83.2656	2439919	93.3848	2278513	1087501	1088924	0.4196	88	44.79
4_group3	85.4123	11685562	97.8549	11434895	5396249	5400912	0.4376	87	40.35
4_group4	86.743	6594820	96.7652	6381492	3076471	3073661	0.4303	88	52.09
D. miranda									
2	18.7027	30819483	93.6237	28854340	4027647	4027315	0.4583	18	18.38
3	20.1168	19787792	94.4606	18691673	2779677	2785477	0.4691	19	26.53
XL_group3a	19.4302	2692213	92.6627	2494677	368119	368015	0.4623	18	32.81
XL_group3b	29.0281	388551	91.2452	354534	79627	79773	0.4578	19	47.92
XL_group1e	17.8448	12541198	92.6182	11615438	1585445	1585920	0.4625	17	21.8
XL_group1a	19.6939	9148293	91.9186	8408982	1275755	1276178	0.4701	18	44.68
XR_group3a	19.3349	1469181	93.2392	1369852	201744	200722	0.481	17	69.58
XR_group8	17.3782	9197557	94.1218	8656902	1121618	1119342	0.4722	17	11.37
XR_group6	18.1368	13333775	94.4924	12599410	1690148	1689856	0.4622	18	24.69
XR_group5	20.3789	740970	94.8644	702917	105668	105158	0.4537	19	21.71
4_group1	16.3376	5287126	90.9867	4810580	609460	610129	0.4451	16	16.38
4_group2	18.8393	1235759	89.9682	1111790	162242	162151	0.4547	18	20.14
4_group5	16.2932	2439919	88.1169	2149982	279051	278681	0.444	17	14.46
4_group3	17.8253	11685562	93.7048	10949928	1461704	1460165	0.4556	17	23.03
4_group4	17.2507	6594820	89.696	5915288	802786	803197	0.4491	17	22.14

Table S3: The expected fold reduction in polymorphism in *SR* relative to *ST* obtained from simple neutral coalescent simulations.

SR Frequency	***SR-ST Divergence: 2M generations*** ***ST-Miranda Divergence: 4M generations***	***SR-ST Divergence: 4M generations*** ***ST-Miranda Divergence: 8M generations***
30%	4.75 (95% CI: 4.73-4.78)	5.59 (95% CI: 5.52-5.65)
13.5%	10.68 (95% CI: 10.64-10.73)	12.44 (95% CI: 12.39-12.49)
1%	150.83 (95% CI: 150.05-151.61)	128.09 (95% CI: 127.21-128.97)

Table S4: Primers used to amplify intergenic regions for linkage disequilibrium analysis.

LD forward primer sequence	
XL1_F	CTTTTGCGTGGGTGTGTTGC
XL2_F	TGCAACCGCACTTGACCGTA
XR1_F	ATGAGGGCGTTCCGAAAACAC
XR2_F	GTGTTTGGGTCGGGAACAGC
XR3_F	TGTCCCAGTCCCCGTTCTGT
XR4_F	AGCTGCCATCCCATTCCAAA
XR5_F	GGGCGAGACATGGGACATTC
XR6_F	TGCCTCGACCCACGAATACA
XR7_F	GCTGTTGCTGGGCAAACTGA
LD reverse primer sequence	
XL1_R	CGGGGACTCCTGCATTATCG
XL2_R	CTCGGCCAGAACCACATGCT
XR1_R	GCATTGGCCCCGAAAAATCAAC
XR2_R	AGCCGAACAGAACCGCAAAG
XR3_R	GCGGATTCGAACCATTCCTG
XR4_R	TAACTTGCACAGCCCCGTCA
XR5_R	CGGCTTCGGGAAACTTTGTGG
XR6_R	ACGGAACAAAACGGCCAAGA
XR7_R	AGCATCGCTTCGCATCTGTG

Table S5: RNA-seq read mapping data.

Strain	*Reads*	*Assigned*
KB10	96835955	83070610
KB12	71249493	62675945
KB1	48891591	42325411
KB2	83033873	72656519
KB3	33197393	29127193
KB5	36001605	31805100
NPZ11	93166191	80719581
NPZ1	57669830	50083732
NPZ29	68329286	58163461
NPZ33	33991458	30042015
NPZ6	95164635	82682359
NPZ8	37619356	31948845

Table S6: Reads per kilobase per million mapped reads (RPKM) and differential expression statistics for all genes.

See attached file DpseSR.GeneExpression.Data.xlsx

Table S7: Counts of non-synonymous and annotated *cis*-regulatory changes in genes detected as differentially expressed on *XR*.

See attached file Dpse.TableS7.csv

Table S8. XR karyotype for tested male offspring from two ST/SR female strains.

XR Karyotype (female gamete/male gamete)	KBPN2 No.	Freq	AO4_No.	Freq
Par 1 ST_1 ST_2 ST_3/ ST_1 ST_2 ST_3	50	0.467	56	0.583
Par 2 ST_1 ST_2 ST_3/ SR_1 SR_2 SR_3	57	0.533	40	0.417
CO 1a ST_1 ST_2 ST_3/ ST_1 SR_2 SR_3	0	0	0	0
CO 1b ST_1 ST_2 ST_3/ SR_1 ST_2 ST_3	0	0	0	0
CO 2a ST_1 ST_2 ST_3/ ST_1 ST_2 SR_3	0	0	0	0
CO 2b ST_1 ST_2 ST_3/ SR_1 SR_2 ST_3	0	0	0	0
DCO 1 ST_1 ST_2 ST_3/ ST_1 SR_2 ST_3	0	0	0	0
DCO 2 ST_1 ST_2 ST_3/ SR_1 ST_2 SR_3	0	0	0	0
Total	107		96	

Par, parental; CO, cross over; DCO, double cross over;

Table S9. Sex ratio (%female) variation across 107 and 96 F1 male offspring of a ST/SR female.

ST/SR Strain	ST	SR
KBPN2		
Range %female	0.41 – 0.68	0.87 - 1.00
Mean %female ± SD	0.53 ± 0.06	0.99 ± 0.02
Mean No. offspring ± SD	121.9 ± 51.2	125.2 ± 47.1
AO2		
Range % female	0.47 – 0.84	0.80 - 1.00
Mean % female ± SD	0.56 ± 0.06	0.96 ± 0.05
Mean No. offspring ± SD	156.6 ± 39.9	156.6 ± 32.2

Table S10: Counts from the recombination experiment. The reported recombination fraction is accompanied by exact 95% confidence intervals for the binomial distribution.

	Visible Marker Classes				Recombination Fraction
Gene Arrangement	+ +	+ *sh*^{1}	*se*^{1} +	*se*^{1} *sh*^{1}	(95% Confidence Interval)
SR Chromosome Isolate 1	1459	0	3	1557	0.0010 (0.0002 - 0.0029)
SR Chromosome Isolate 2	1657	0	6	1914	0.0017 (0.0006 - 0.0036)
SR Chromosome Isolate 3	1464	0	3	1745	0.0009 (0.0002 - 0.0027)
ST Chromosome	321	212	253	297	0.4294 (0.3996 - 0.4595)

Figure S1. *D. pseudoobscura* male and female larvae (Anterior to Posterior, left to right). The male larva has an obvious gonad about a third of the distance from the posterior end of the larva.

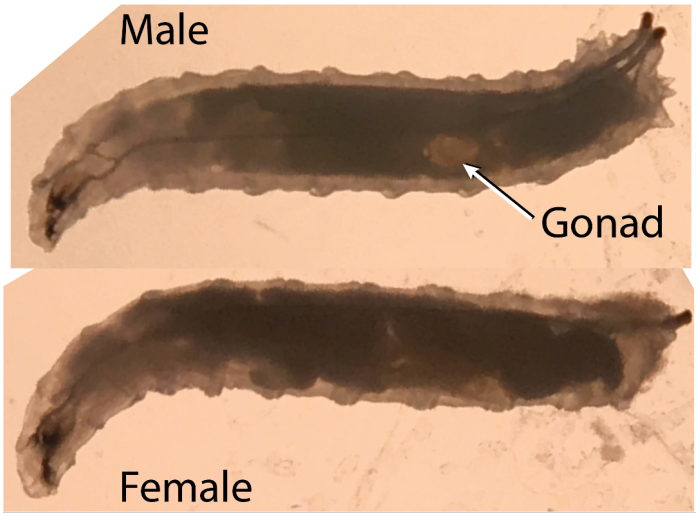


Figure S2. Genetic mapping crosses of KBPN2 and AO4 ST/SR females to detect cross overs among the three non-overlapping inversions that comprise the SR chromosome. Heterozygous ST/SR females were crossed to marked ST hemizygous males. F1 male offspring from each parental cross were individually crossed to virgin ST/ST females. Female larvae from each male were karyotyped at the three inversion loci. The F2 offspring were sexed and counted.

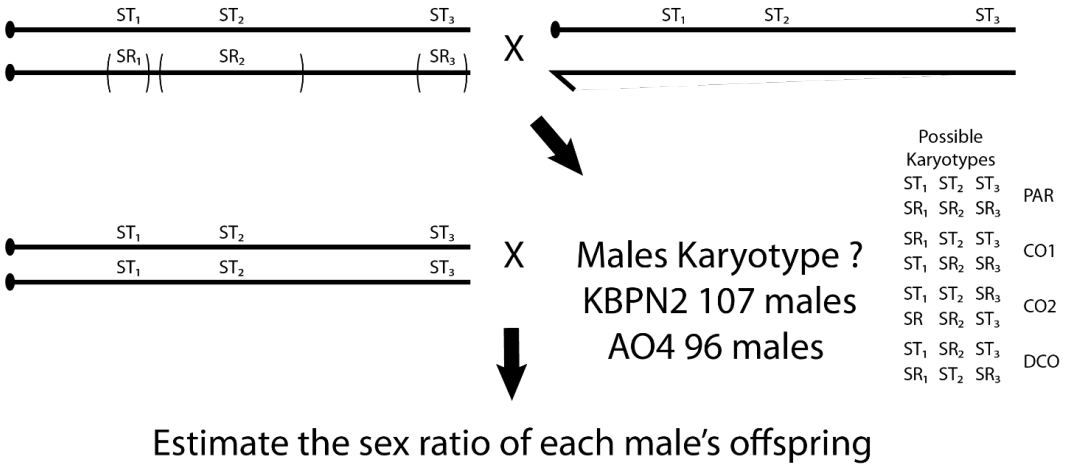


Figure S3: Patterns of polymorphism across chromosome *XR* in non-overlapping 10kb windows, considering all sites. A) Pairwise nucleotide diversity measured as π for *ST* (blue), *SR* (green), and across both chromosome types jointly (black). The lines are the loess smoothed trend lines and dots represent each window. The boxplot on the right summarizes diversity in different regions of the chromosome. Shaded regions in purple represent inverted regions. B) The same plots, but for the site frequency spectrum summarized with Tajima's *D*. C) Tajima's *D* summarized for chromosomal regions in non-overlapping 10kb windows considering all sites for *SR* (green) and for sites with shared polymorphisms masked for *SR* (light green).

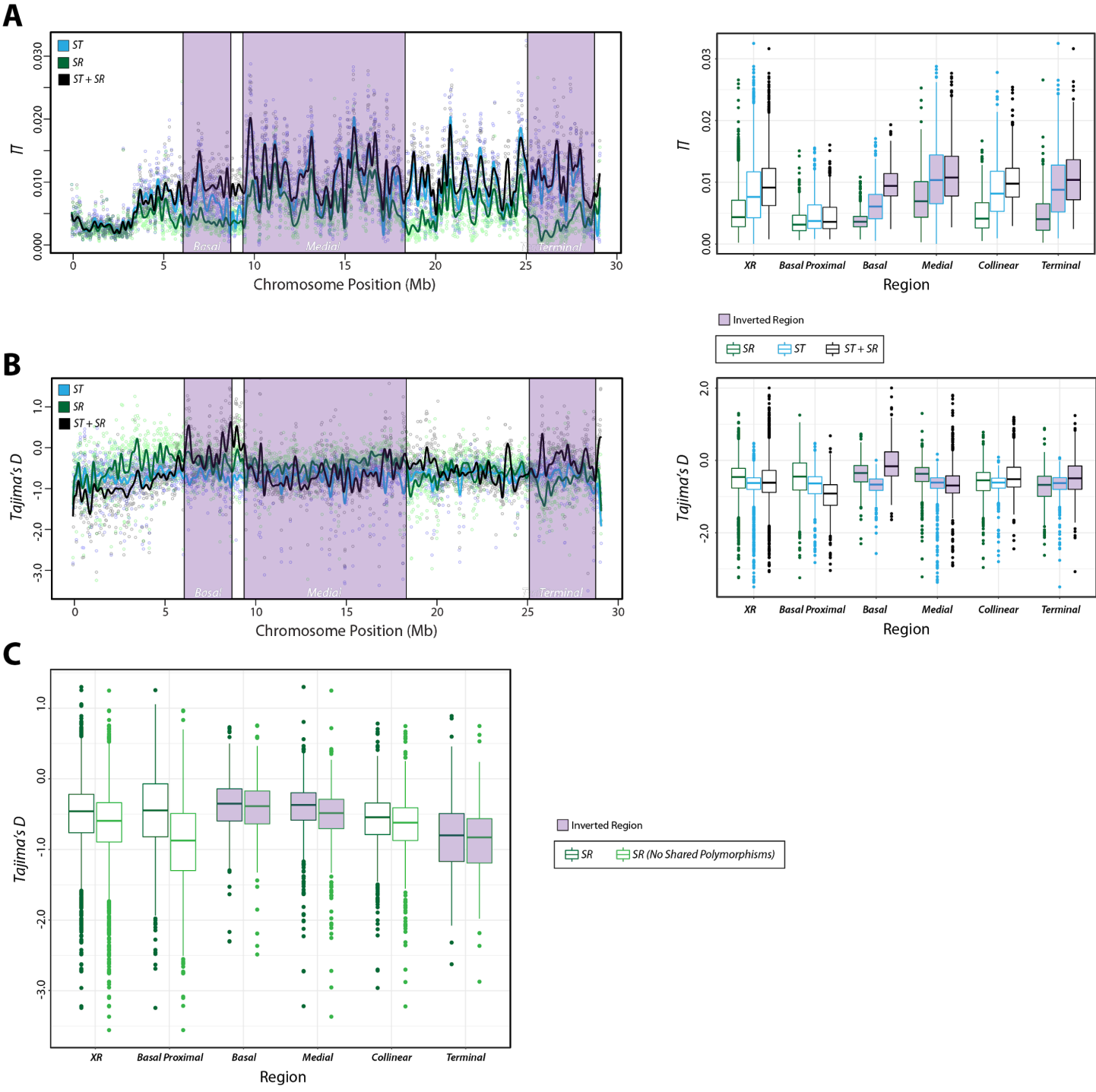


Figure S4: Polymorphism measured as the proportion of pairwise differences per-synonymous (π_S; A) and per-nonsynonymous (π_N; B) site for *ST* (blue) and *SR* (green). Each boxplot represents a region on the chromosome. The only significant differences (by Mann-Whitney *U* test) are detected as reduced π_S for *SR* within the terminal inversion and intervening collinear regions.

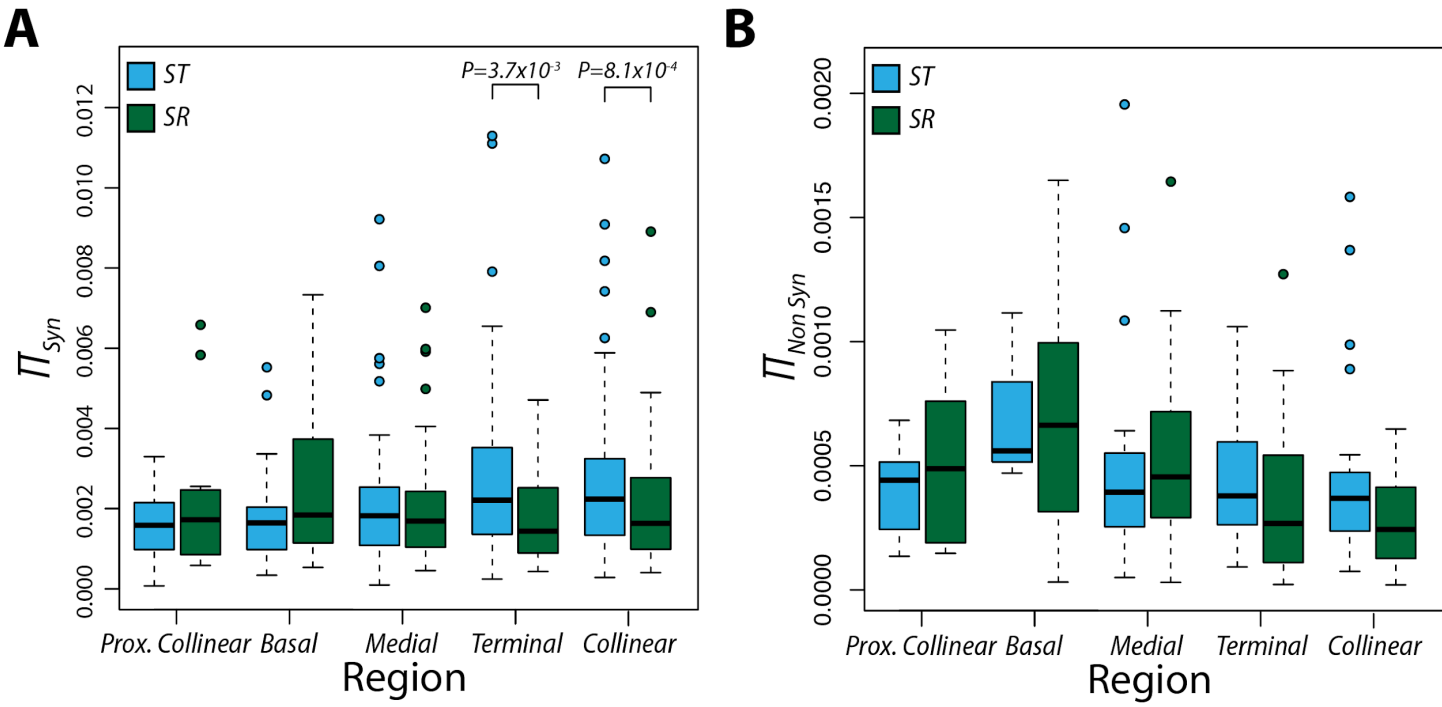


Figure S5: Structure of genetic variation on the autosomes in the RNA-seq data. A principal components analysis (PCA) was performed on genotypes called in all autosomal transcripts. The samples strongly cluster by *ST/SR X*-chromosome status, suggesting the presence of structured genetic variation present on the autosomes introduced by the crossing scheme used to generate the strains. As a result, only *X*-chromosome transcripts are analyzed further for differential expression.

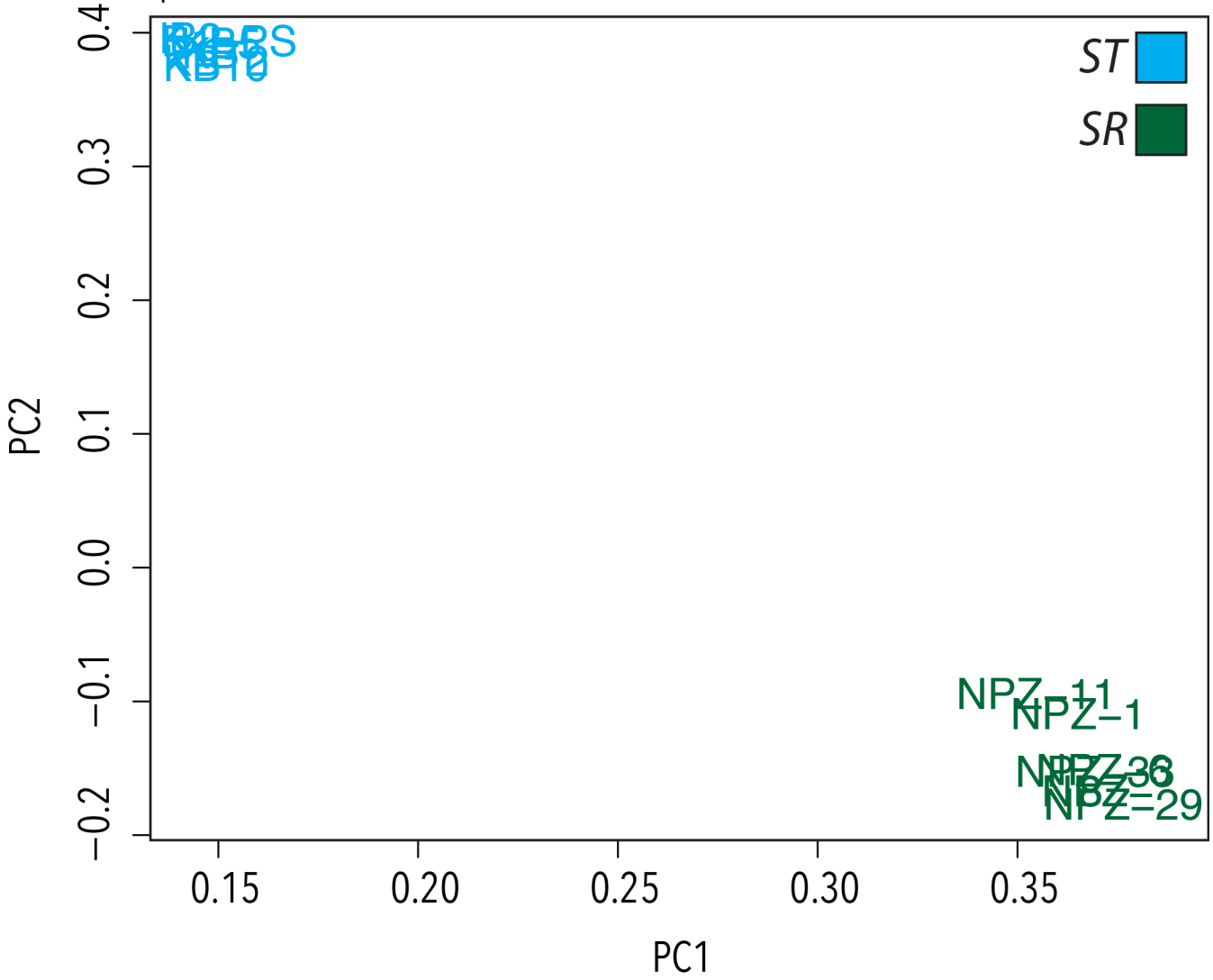


Figure S6: Diagram of the drive-selection balance model. The frequencies of the recombinant chromosomes are modeled in gametic and genotypic stages (both pre and post selection). Because *SR* is *X*-linked the irregular transmission between male and female gametic pool must also be incorporated. Finally, the sex-specific effects of drive (males only) and recombination (females only) are modeled. Given this model of allele frequency change, decay of linkage disequilibrium can be calculated in the traditional manner.

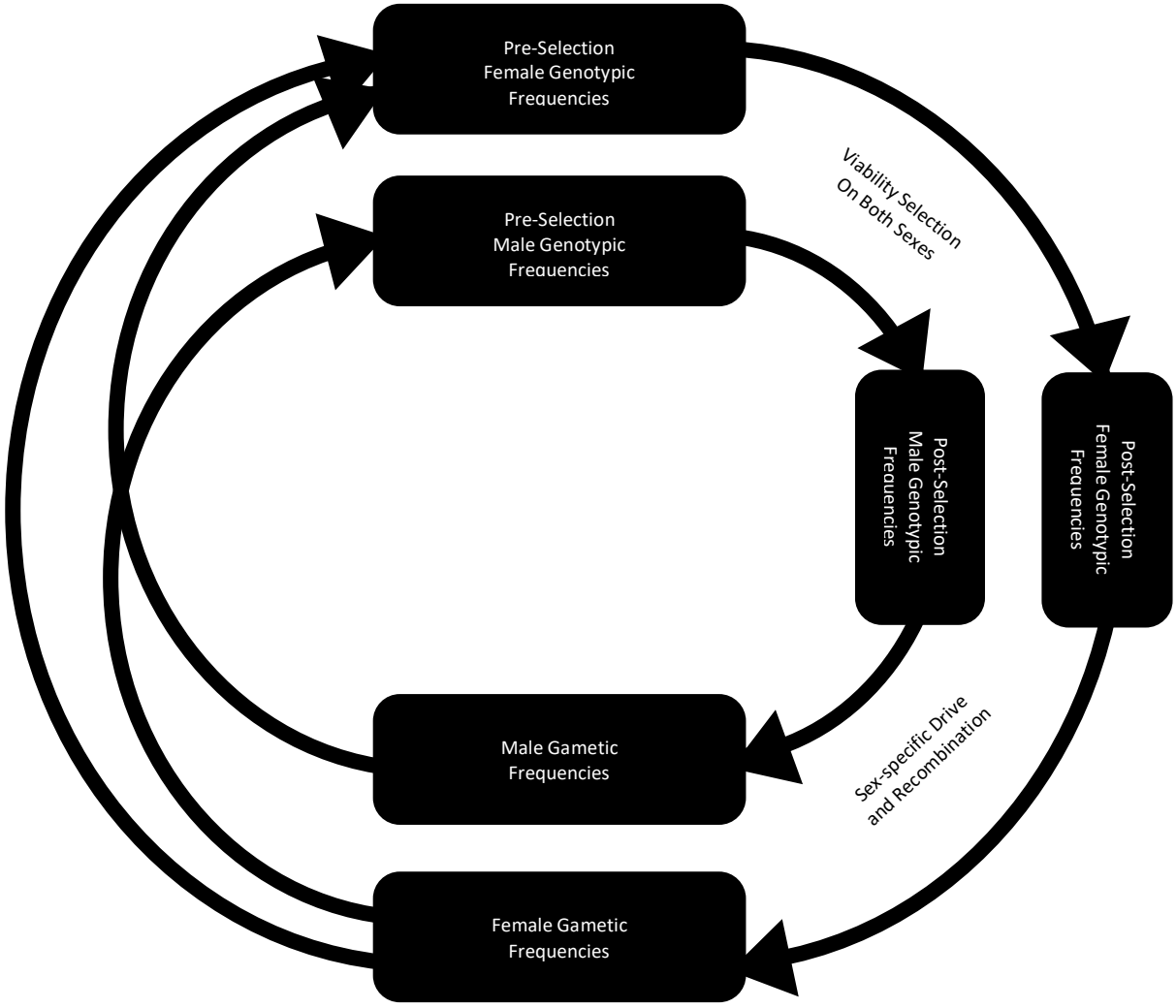
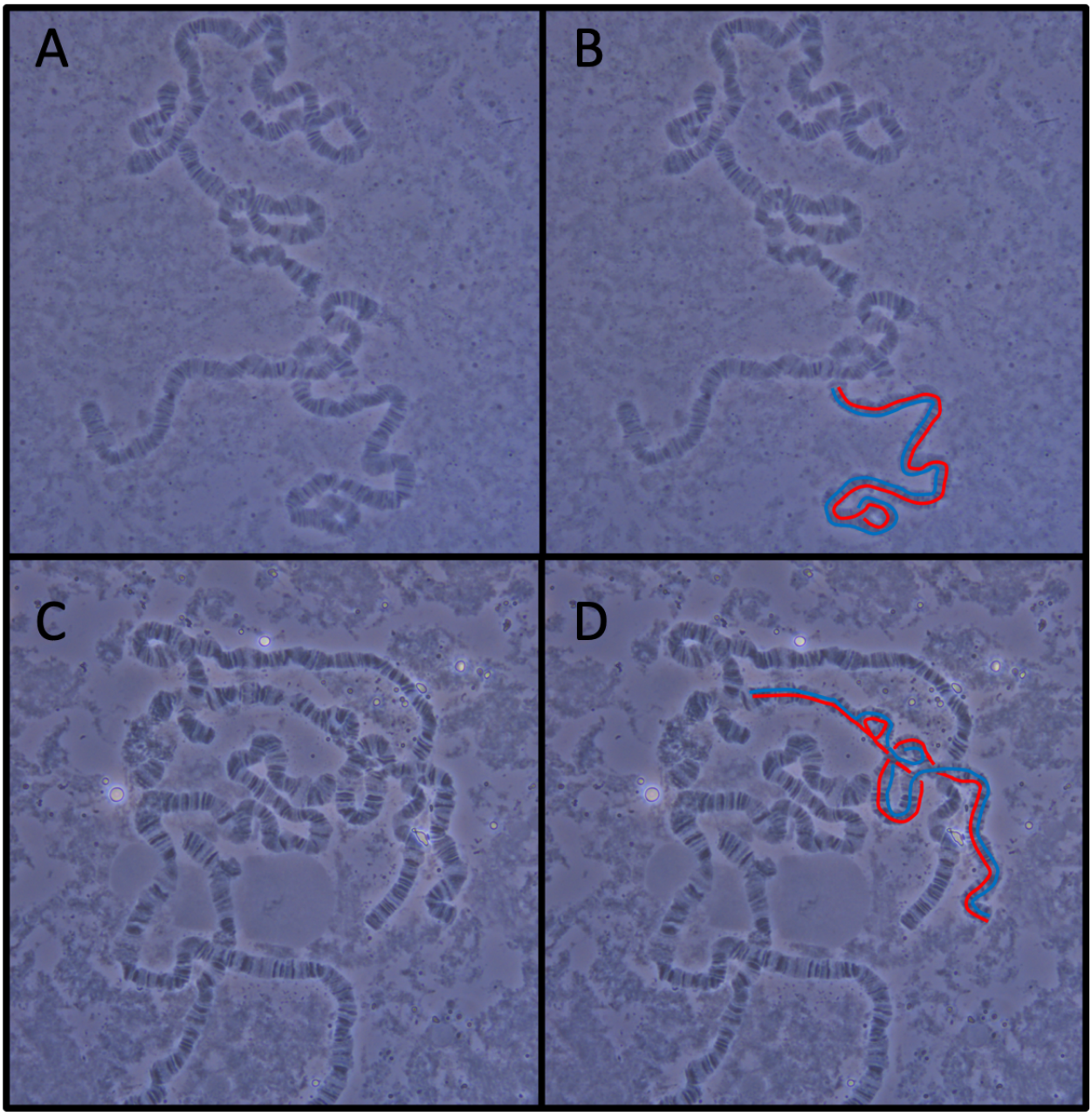


Figure S7: Polytene chromosomes of recombinant *SR* chromosomes in the heterozygous state with the standard arrangement. The terminal inversion only inversion is depicted in panel A with traces of homologous strands in panel B. The basal and medial inversion carrying recombinant is depicted in panel C with traces of homologous strands in panel D.



Literature Cited

Beckenbach, A. T. (1996). Selection and the "Sex-Ratio" polymorphism in natural populations of *Drosophila pseudoobscura*. *Evolution, 50*, 787-794.

Crow J. F., and M. Kimura, 1970 *An introduction to population genetics theory*. Harper & Row.

Ewens W. J., 2004 *Mathematical Population Genetics 1: Theoretical Introduction*. Springer Science & Business Media.

Foss, E., Lande, R., Stahl, F. W., & Steinberg, C. M. (1993). Chiasma interference as a function of genetic distance. *Genetics, 133*(3), 681-691.

Hilliker, A. J., Harauz, G., Reaume, A. G., Gray, M., Clark, S. H., & Chovnick, A. (1994). Meiotic gene conversion tract length distribution within the *rosy* locus of *Drosophila melanogaster*. *Genetics, 137*(4), 1019-1026.

Navarro, A., Betrán, E., Barbadilla, A., & Ruiz, A. (1997). Recombination and gene flux caused by gene conversion and crossing over in inversion heterokaryotypes. *Genetics, 146*, 695-709.

Nielsen R., and M. Slatkin, 2013 *An Introduction to Population Genetics: Theory and Applications*. Sinauer.

Painter, T. S. (1934). A new method for the study of chromosomal aberrations and the plotting of chromosomal maps in *Drosophila melanogaster*. *Genetics, 19*, 175-188.

Wallace, B. (1948). Studies on "sex ratio" in *Drosophila pseudoobscura*. I Selection and "sex ratio". *Evolution, 2*, 189-217.