Supplemental Data

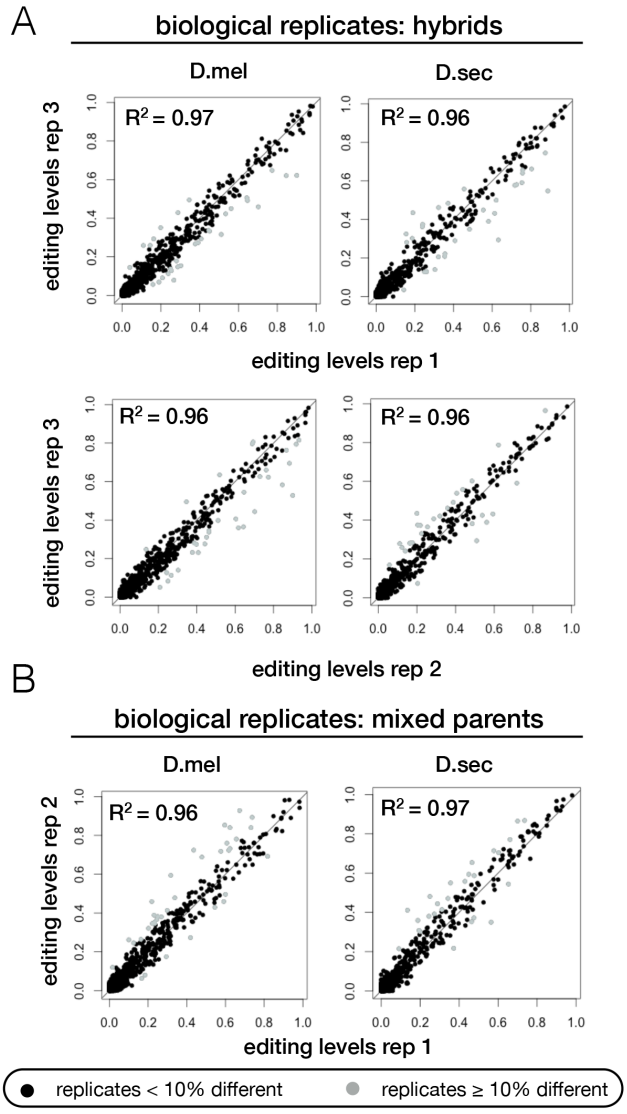

Figure S1, Related to Figure 1. Reproducibility of editing measurements in biological replicates of F1 hybrid and mixed parental mapping control. (A) Scatter plots comparing editing levels of third F1 hybrid biological replicate in both *D. melanogaster* and *D. sechellia* alleles to replicates 1 and 2 (see Fig 1E). (B) Scatter plots comparing editing levels in biological replicates of mixed parental control. Gray dot, replicates differed by ≥ 10% editing and site was excluded from further analyses.

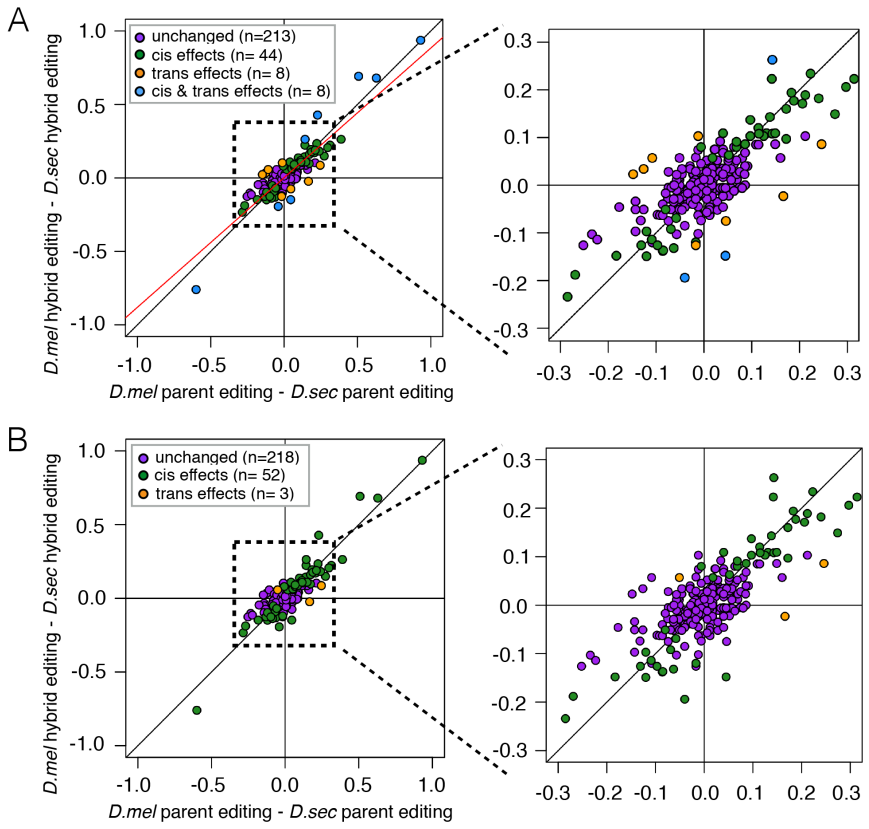

Figure S2, Related to Figure 2. Significant *cis* and *trans* regulation with alternative statistical analyses for *trans* sites. Scatter plots comparing the difference in editing between parents and the difference in editing between hybrid species-specific alleles. *Cis* regulated sites were determined as in **Fig 2C**. (A) *Trans* sites were determined as outliers of standardized residuals from the linear regression (in red), as described in **Experimental Procedures**, without controlling for FDR. (B) *Trans* sites were determined as sites where the 95% confidence intervals of the odds ratios from Fisher's exact tests between parents and hybrids (see Fig 2A, 2B) did not overlap. Purple dot, unchanged between the two species and hybrids. Green dot, evidence of *cis* divergence. Orange dot, evidence of *trans* divergence. Blue dot, evidence of *cis* and *trans* divergence. Right plots, magnification of points for clarity.

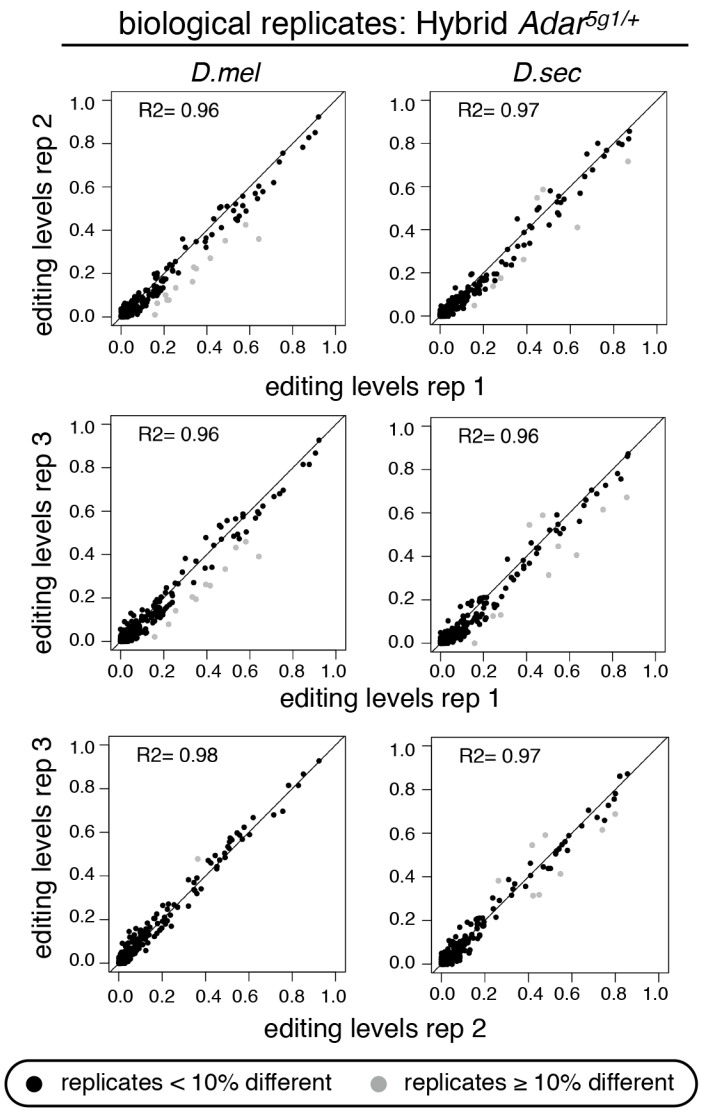

Figure S3, Related to Figure 4. Reproducibility of editing measurements in biological replicates of individual *Adar*$^{5G1/D.sec+}$ hybrid heads. Scatter plots comparing editing levels in *D. melanogaster* and *D. sechellia* alleles from three individual heads of *Adar*$^{5G1/D.sec+}$ hybrid flies. Gray dot, replicates differed by ≥ 10% editing and site was excluded from further analyses.

Table S1, Related to Figure 1. Multiplex PCR primer sequences. Principal amplicon target site and PCR primers used in mulitplex PCR. Each pool of primers is separated by a space.
(Table S1.xlsx)

Table S2, Related to Figure 2. Editing levels in *D. melanogaster* and *D. sechellia* parents and hybrids. Editing site annotations and editing levels measurements for *D. melanogaster* and *D. sechellia* from separate and mixed parent samples and *D. melanogaster* and *D. sechellia* alleles in hybrids. See **Experimental Procedures** for description of *cis* and *trans*-regulation p-value calculations.
(Table S2.xlsx)

Supplemental Experimental Procedures

Fly husbandry

D. melanogaster wildtype flies (iso-1 : y^1; $Gr22b^1$ $Gr22d^1$ cn^1 $CG33964^{R4.2}$ bw^1 sp^1; $LysC^1$ $MstProx^1$ $GstD5^1$ $Rh6^1$) were obtained from the Bloomington *Drosophila* Stock Center (BL2057). *D. sechellia* wildtype flies (14021-0248.25) were obtained from the UC San Diego *Drosophila* Stock Center. Female *D. melanogaster* were crossed to male *D. sechellia* to obtain F1 hybrids. $Adar^{5G1}$ mutants were generated in (Palladino et al., 2000) and generously provided by Liam Keegan. All flies were raised at 23°C on molasses-based food (Fly Media Center at Stanford University).

Multiplex PCR primer design considerations

To best dissect *cis* and *trans* regulation of editing using mmPCR-seq, we wanted to target sites with high editing levels in either *D. melanogaster* or *D. sechellia*, such that we might observe large differences in editing between the species. We chose to quantify editing in *D. melanogaster* and *D. sechellia* and their hybrids because, despite having diverged just ~1.2 million years ago (Cutter, 2008), they may have functional differences in potential *trans* regulators because *D. sechellia* derives from an isolated population with a small effective population size resulting in an increase in slightly deleterious mutations (Kliman et al., 2000). We utilized existing RNA-seq data (McManus et al., 2010) to identify editing sites that might be highly edited (>10%) in either species. Requiring a minimum coverage of 10 reads, we measured editing levels by determining the percentage of G reads at known editing sites (Ramaswami and Li, 2014) that are conserved across at least two *Drosophila* species (R.Z. and P.D. et al., unpublished data). We did not specifically target sites that are lowly edited in both species as they are more likely to only show subtle, undetectable differences in editing levels between the species, although we amplified many of these sites as they were nearby sites predicted to be highly edited. We designed 493 pairs of PCR primers, which were pooled into 48 groups of 8-12

primer pairs each and used in mmPCR-seq. To design primers that would amplify alleles from both species, we required high conservation within primer sequences, but at least one mismatch within each amplicon so that we could distinguish sequencing reads derived from *D. melanogaster* versus those from *D. sechellia*. These primers amplified regions spanning a total of 1,036 sites in *D. melanogaster* or *D. sechellia* parents. Of these sites, 638 are found in previous studies (Graveley et al., 2010; Ramaswami et al., 2013; Rodriguez et al., 2012; St Laurent et al., 2013), 105 were identified using standard protocols (Ramaswami et al., 2013) in at least two *Drosophila* species from RNA-seq data available through modENCODE (Celniker et al., 2009)(R.Z. and P.D. et al., unpublished data), and an additional 293 sites were called in this study by requiring at least 2% A-to-G or T-to-C variant frequency across all samples for *D. melanogaster* or *D. sechellia* parents. The origin of each of the final 273 sites analyzed for *cis* or *trans* effects is listed in **Table S2**.

Processing parent, mixed parent and hybrid samples for mmPCR-seq

Fly heads from 0-2 day old female flies were dissected in PBS and flash frozen in liquid nitrogen. Total RNA was extracted using Trizol (Life Technologies) followed by Qiagen RNeasy column cleanup, and cDNA synthesis was performed using iScript Advanced cDNA synthesis kit (BioRad) following standard protocols. Multiplex PCR was performed using the Fludigim Access Array system using KAPA 2G multiplex PCR mix (Kapa Biosciences) following our previously published protocols (Zhang et al., 2014). Amplicons were sequenced using the Illumina MiSeq platform for paired-end 145bp reads.

Processing *Adar*$^{5G1/D.sec+}$ hybrid samples for mmPCR-seq

Adar$^{5G1/D.sec+}$ hybrids were collected as individual heads. The *Adar*^{5G1} mutation is balanced over FM7 on the X chromosome, but the eye markers associated with this balancer are not visible in the hybrid. Therefore, we PCR amplified and Sanger sequenced the *Adar* locus in the hybrids to verify the

absence of *D. melanogaster* variants that were found in *FM7/D.sec* flies. The samples from individual heads determined to be null for *D. melanogaster Adar* were then processed using the same protocols as the parent and hybrid samples and sequenced using the Illumina NextSeq 500 platform for single-end 150bp reads.

Mapping pipeline

RT-PCR primer sequences were trimmed from the beginning of reads. Custom python scripts were used to assign reads to a species and calculate editing levels. For parent, mixed parent and hybrid samples, paired-end 124 bp sequencing reads were mapped using the Burrows-Wheeler algorithm (BWA) (Li and Durbin, 2010) allowing 10 mismatches per read to the *D. melanogaster* genome (BDGP R5/dm3, Apr. 2006) (Adams et al., 2000; Celniker et al., 2002) and to the *D. sechellia* genome (Broad/droSec1, Oct. 2005)(Drosophila 12 Genomes Consortium et al., 2007) plus splice junctions assembled at slightly shorter lengths than the read length to avoid duplicate mapping. Gene models were obtained from the University of California Santa Cruz (UCSC) genome browser (Karolchik et al., 2014; 2004) (http://genome.ucsc.edu/): FlyBase Genes (*D. melanogaster*), Genscan genes (*D. sechellia*). For each mapped read, base calls at known variants between species (homologous sites determined by the UCSC liftOver tool: http://genome.ucsc.edu/cgi-bin/hgLiftOver with strand information obtained from pairwise alignment files) were compared to both the *D. melanogaster* genome and the *D. sechellia* genome in order to assign it to a species. Reads were discarded if no variants were covered or if multiple variants were present in the read that did not agree on the species match. Sites were discarded if more than 15% of the reads were called the incorrect species in the single parent samples. A and G counts at editing sites from species-assigned reads were used to calculate editing levels. For $Adar^{5G1/D.sec+}$ mutant hybrid heads, the custom python scripts were modified to account for single-end sequencing reads.

Editing site annotation

Editing sites were annotated using RefSeq gene annotations and ANNOVAR software (Wang et al., 2010).

Comparison of Adar protein sequences from two species

Adar protein sequences were obtained from Flybase (Santos et al., 2014) for *D. melanogaster* Adar (Accession: FBpp0070286) and *D. sechellia* GM19166 (Accession: FBpp0200643) and aligned using protein BLAST (Altschul et al., 1990).

Quantitative real time PCR (qPCR)

Primers used in qPCR experiments were designed by Fly Primer Bank (Hu et al., 2013), and primer efficiency was tested on 5-fold dilutions of cDNA from heads of both *D. melanogaster* and *D. sechellia* female flies. Primer sequences are as follows: *Gapdh* (F: 5'-TAAATTCGACTCGACTCACGGT, R: 5'-CTCCACCACATACTCGGCTC), *Adar* (F: 5'-TACACGCACCTCTATTCACGA, R: 5'- AGAACTGCTCCATCCTTAAACTG), *period* (F: 5' - GGCACAGAGCTTTCCGATTC, R: 5'-CTCCCAGGTCTTGTCGAACT), *Fmr1* (F: 5' - TCAGAAGCGAGCCAGTATGC, R: 5' – CAAACCCATGAGATCGTCACG). qPCR was performed with the BioRad CFX96 Real-Time System.

Measuring editing levels from Sanger sequencing

Editing levels at the *Adar* editing site were measured from Sanger sequencing traces using the height of G and A peaks as representative of edited and unedited reads, respectively, and editing levels were calculated as G/(A+G). Peak heights were determined using the ab1PeakReporter from Life Technologies (https://apps.lifetechnologies.com/ab1peakreporter/).

Supplemental References

Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., et al. (2000). The genome sequence of *Drosophila melanogaster*. Science *287*, 2185–2195.

Altschul, S.F., Gish, W., Miller, W., and Myers, E.W. (1990). Basic local alignment search tool. J Mol Biol *215*, 403–410.

Celniker, S.E., Dillon, L.A.L., Gerstein, M.B., Gunsalus, K.C., Henikoff, S., Karpen, G.H., Kellis, M., Lai, E.C., Lieb, J.D., MacAlpine, D.M., et al. (2009). Unlocking the secrets of the genome. Nature *459*, 927–930.

Celniker, S.E., Wheeler, D.A., Kronmiller, B., Carlson, J.W., Halpern, A., Patel, S., Adams, M., Champe, M., Dugan, S.P., Frise, E., et al. (2002). Finishing a whole-genome shotgun: release 3 of the *Drosophila melanogaster* euchromatic genome sequence. Genome Biology *3*, RESEARCH0079.

Cutter, A.D. (2008). Divergence times in *Caenorhabditis* and *Drosophila* inferred from direct estimates of the neutral mutation rate. Mol. Biol. Evol. *25*, 778–786.

Drosophila 12 Genomes Consortium, Clark, A.G., Eisen, M.B., Smith, D.R., Bergman, C.M., Oliver, B., Markow, T.A., Kaufman, T.C., Kellis, M., Gelbart, W., et al. (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. Nature *450*, 203–218.

Graveley, B.R., Brooks, A.N., Carlson, J.W., Duff, M.O., Landolin, J.M., Yang, L., Artieri, C.G., van Baren, M.J., Boley, N., Booth, B.W., et al. (2010). The developmental transcriptome of *Drosophila melanogaster*. Nature *471*, 473–479.

Hu, Y., Sopko, R., Foos, M., Kelley, C., Flockhart, I., Ammeux, N., Wang, X., Perkins, L., Perrimon, N., and Mohr, S.E. (2013). FlyPrimerBank: an online database for *Drosophila melanogaster* gene expression analysis and knockdown evaluation of RNAi reagents. G3 (Bethesda) *3*, 1607–1616.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2014). The UCSC Genome Browser database: 2014 update. Nucleic Acids Research *42*, D764–D770.

Karolchik, D., Hinrichs, A.S., Furey, T.S., Roskin, K.M., Sugnet, C.W., Haussler, D., and Kent, W.J. (2004). The UCSC Table Browser data retrieval tool. Nucleic Acids Research *32*, D493–D496.

Kliman, R.M., Andolfatto, P., Coyne, J.A., Depaulis, F., Kreitman, M., Berry, A.J., McCarter, J., Wakeley, J., and Hey, J. (2000). The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics *156*, 1913–1931.

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics *26*, 589–595.

McManus, C.J., Coolon, J.D., Duff, M.O., Eipper-Mains, J., Graveley, B.R., and Wittkopp, P.J. (2010). Regulatory divergence in *Drosophila* revealed by mRNA-seq. Genome Res. *20*, 816–825.

Palladino, M.J., Keegan, L.P., O'Connell, M.A., and Reenan, R.A. (2000). A-to-I Pre-mRNA Editing in

Drosophila Is Primarily Involved in Adult Nervous System Function and Integrity. Developmental Biology *102*, 437–449.

Ramaswami, G., and Li, J.B. (2014). RADAR: a rigorously annotated database of A-to-I RNA editing. Nucleic Acids Research *42*, D109–D113.

Ramaswami, G., Zhang, R., Piskol, R., Keegan, L.P., Deng, P., O'Connell, M.A., and Li, J.B. (2013). Identifying RNA editing sites using RNA sequencing data alone. Nat Meth *10*, 128–132.

Rodriguez, J., Menet, J.S., and Rosbash, M. (2012). Nascent-seq indicates widespread cotranscriptional RNA editing in *Drosophila*. Molecular Cell *47*, 27–37.

Santos, dos, G., Schroeder, A.J., Goodman, J.L., Strelets, V.B., Crosby, M.A., Thurmond, J., Emmert, D.B., Gelbart, W.M., and Consortium, T.F. (2014). FlyBase: introduction of the *Drosophila melanogaster* Release 6 reference genome assembly and large-scale migration of genome annotations. Nucleic Acids Research *43*, gku1099–D697.

St Laurent, G., Tackett, M.R., Nechkin, S., Shtokalo, D., Antonets, D., Savva, Y.A., Maloney, R., Kapranov, P., Lawrence, C.E., and Reenan, R.A. (2013). Genome-wide analysis of A-to-I RNA editing by single-molecule sequencing in *Drosophila*. Nat Struct Mol Biol *20*, 1333–1339.

Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research *38*, e164–e164.

Zhang, R., Li, X., Ramaswami, G., Smith, K.S., Turecki, G., Montgomery, S.B., and Li, J.B. (2014). Quantifying RNA allelic ratios by microfluidic multiplex PCR and sequencing. Nat Meth *11*, 51–54.