Supplementary Document 1. Enhanced Methods

Genetics

The inbred lines were from two North American populations of D. *melanogaster*, a panel of 90 inbred lines derived from a vineyard near Davis, California, in 1996 (CA lines, courtesy of S. Nuzhdin) and a second set of 120 inbred lines established from a peach orchard near West End, North Carolina, in 2000 (NC lines). Single gravid females were collected from these populations, and their offspring were sib-mated for 15-20 (NC) or 50 generations (CA). This inbreeding program was sufficient to generate isogenicity at *EGFR* for greater than 95% of the lines as determined by comprehensive sequence analysis [29]. The remaining individual heterozygous sites were scored as missing data.

Sensitization crosses were performed using a gain-of-function allele of *Egfr* termed *Ellipse*^{1} (*Egfr*^{E1}) that is caused by an alanine to threonine amino acid replacement, A887T (14). In order to control for genetic background effects, the experiment was conducted with two strains harboring the *Egfr*^{E1} allele with the following genotypes: BL-1564, *nw*^{D} *Pu*^{2} *Egfr*^{E1} *Pin*^{Yt}/SM1; and BL-5144, P{w$^{+mW.hs}$=GawB}*elav*^{C155} *w** P{ry$^{+t7.2}$=neoFRT}19A; *Bc*^{1} *Egfr*^{E1}/CyO. Both strains were obtained from the Bloomington stock center and inbred by 3 generations of sib-mating immediately prior to the experiment. Two virgin females of each of the *Egfr*^{E1} stocks were crossed to two males of each inbred line. Each cross was conducted in replicate and according to a randomized block design, interleaving both *Egfr*^{E1} strains and

populations of inbred lines. All flies were grown at standard conditions 25ºC on standard cornmeal yeast medium under constant light/dark cycle.

Variation in eye roughness was scored by simple direct examination of 10 individuals of each sex from each of 2 replicate vials according to a previously described scale (15). A single experimenter assigned a numerical score to each specimen along the preset qualitative scale of eye roughness, ranging from complete suppression ("0") to extreme enhancement ("5").

Polymorphisms in *EGFR* were identified by sequencing 10.9 kb corresponding to the 6 exons and flanking non-coding regions of each of the inbred lines [29]. A total of 245 SNPs and 30 indels with the more rare allele at a frequency of 0.05 or higher were tested for their effect on eye roughness. In brief, sequencing was performed by PCR amplification of genomic DNA extracted from single flies. Gel purified PCR products were then used as templates for the sequencing reaction, using Big Dye terminator chemistry and ABI 3700 automated sequencers at the NCSU-GRL.

For the Case-Control and TDT tests, 1000 males from West End, NC, were collected in the summer of 2002 (WE2002). These males were immediately crossed to BL-1564 virgin females, bearing the $Egfr^{E1}$ allele. Eye phenotypes were determined for the non-*Cy* offspring (namely those that must carry the $Egfr^{E1}$ allele). Approximately 15% of crosses produced at least one male with apparently wild-type eyes, and 15% produced progeny with blistered eyes due to strong enhancement of the mutation. One male was selected from each of these case and control-producing crosses, along with its father, for subsequent genotypic analysis. These

flies were sequenced for ~700 bp using the following primers: sense, AGATTAACGTGCTCCACAGA or GGCCACTGGAATCTTGACG and anti-sense GAACAGGTGTGCTCCAAGTG . Heterozygous individuals could be easily scored from the sequence chromatographs. Any ambiguously called SNPs were resequenced from the opposite direction. Of the approximately 150 father-son pairs genotyped in each class, as expected between one quarter and one half were heterozygous in the father and hence contributed to the TDT calculations in Table 2.

Confirmation of the polymorphism for the pogo transposable element was made via PCR amplification of flanking regions with the following primers: sense, GGCCAACAACAGAGTGTGTG, and anti-sense CGATTAGCAACGAGCTTTCC. Presence of the polymorphism could be detected by size alone, and was unambiguous.

Statistical Analyses

All statistical analyses were performed using SAS Version 8.2 (Cary, NC), fitting models that included fixed effects of Population (*P*: NC or CA), Sex (*S*), Cross (*C*: BL1564 or BL5144) and Genotype for each SNP (*G*), and random effects of Line (*L*) and Replicate vial (*R*). The dependent variable, *y*, was the qualitative eye roughness score. Line means by sex, denoted *Y* below, were assessed for each $Egfr^{E1}$ strain separately, using the least square means option in PROC GLM with the model $y = L + S + L\times S + R(L\times S) + \varepsilon$.

Associations between sequence variants in *Egfr* and eye-roughness were tested using PROC MIXED for each SNP separately on least square mean phenotype estimates, with the model $Y = P + S + C + G + G\times C + G\times P + L(G\times P) + \varepsilon$. The random effect $L(G\times P)$ controls for the correlation between sexes and crosses due to line effects that are nested within genotype and population. A number of sub-models were also considered, including Cross separately and Sex and Cross separately using PROC GLM. Where possible, results were confirmed using a non-parametric permutation procedure in Tassel (www.maizegenetics.net), which allows for the control of population admixture [22,23]. In general, to correct for multiple comparisons we used a Sequential Bonferroni procedure [17].

To determine whether sites of significant interest were interacting in some fashion, an *a posteriori* approach was employed. First we assessed the significance of haplotypes based on the results from the initial association test using the same ANOVA approach as discussed above, but with multiple-state haplotype codes replacing biallelic SNPs. Second, epistatic contributions in the BL-1564 cross were assesed in PROC MIXED using a multi-site association test and a model of the form: $Y = P + S + G1 + G2 + G1\times G2 + G1\times G2\times P + L(\mathrm{G1}\times G2\times P) + \varepsilon$, where *G1* and *G2* are the two SNP genotypes.

For the case control test, approximately 150 individuals with the most extreme class of eye roughness score were treated as "cases", and the individuals without any visible eye roughening (i.e. the same phenotype as wild-type flies) were labeled "controls". The case control test essentially tests for differences between observed and expected allele frequencies between the two groups, using a G test statistic and the traditional "χ^2" which provided similar

results [18,28]. This test was only performed for four candidate SNPs, three of which are in strong linkage disequilibrium, and thus are non independent. To test for haplotype effects we used haplotypes that showed significant differences in eye roughness as our candidate haplotypes. As well, a log-linear analysis was performed to test for interactions between candidate SNPs and the phenotypic classes. Analysis was performed in SAS using PROC FREQ for the likelihood-ratio tests, and PROC CATMOD for the log-linear analysis.

The Transmission-Disequilibrium test was performed according to Spielman *et al*. [19], using the same SNP variants that were employed in the case-control tests. The TDT evaluates the significance of unequal SNP allele transmission from heterozygous fathers to their affected offspring. Given the sample size, we used both conventional "χ^2" and a exact binomial test for the equality of two proportions, which provided comparable results. PROC FREQ was used for both of these tests.