**Supplementary File Captions**

**File S1: Supplementary Methods and Results.**

**File S2:** **Candidate SNPs and genes.** SNPs identified as candidates in each control and desiccation generation 13 replicate with respect to the mass-bred population. SNP effect and gene annotation (SNPEff results) and candidate gene lists for each pool are reported. For the desiccation-selected replicate line gene lists, whether each gene was present in the putative lab adaptation gene list (genes occurring in at least one control line replicate) is also indicated.

**File S3:** **Allele counts and SNP frequency results** for candidate SNPs identified in each control and desiccation-selected replicate. See included Readme file for details on data format.

**File S4: Category enrichment for candidate genes.** Gene enrichment test results for gene lists significantly differentiated between the mass-bred and generation 13 pools, for all control and desiccation-selected replicates.

**Figure S1: Number of SNPs called in each mass bred—generation 13 comparison.** Results are presented for different levels of minimum coverage depth (x-axis) and minimum count of minor allele across both pools (panels). The combination of 0 minimum coverage depth and 3 minimum minor allele count was used for further analysis.

**Figure S2:** **Chromosome level plots of genomic regions enriched in SNPs that showed significant differentiation between generation 0 and generation 13 pools**. Grey line shows the proportion of SNPs that were significantly-differentiated in 100 Kbp windows; black line shows the proportion in 1 Mb windows. Horizontal black line segments highlight regions containing 100 Kbp windows with a proportion of significantly-differentiated SNPs in the 80th percentile (genome-wide). Where such windows occurred within 1 Mbp of each other, they were combined into a single region (see Methods for analysis performed on these regions). The pale blue line shows the proportion of SNPs with allele frequency <0.1 at generation 0 (‘low-frequency SNPs’) per 100 Kbp window; the dark blue line shows the same parameter calculated over 1 Mbp windows. Asterisks indicate cases where the proportion of low-frequency SNPs showed a significant positive correlation with the proportion of significantly-differentiated SNPs (see Figure S6) after excluding windows without significantly-differentiated SNPs.

**Figure S3: Distribution of sequencing coverage depth in the generation 13 pool** for candidate SNPs (solid coloured line) and 100,000 randomly-selected SNPs (coloured dashed line) for each replicate. The distribution is also shown for the same candidate SNPs (solid black line) and randomly-selected SNPs (dashed grey line) for the generation 0 mass-bred pool.

**Figure S4:** **Boxplots of frequency change (from generation 0 to generation 13 of the selection regime) in each replicate line for candidate SNPs**. Boxes display the range of observations (whiskers), the median value (thick line), first and third quartiles (box ends), and outliers (points).

**Figure S5**: **Correlation between the number of true associations and the number of false associations** detected 10Kbp (left panel) and 1Mbp (right panel) away from a causal variant explaining 1% of broad-sense heritability. The results are based on 2 million independent simulations under the model presented in File S1**.** The densities are colour-coded on a log-scale. Excluding the simulations where no association was found, we observed a positive correlation between number of true associations and short-range false associations and a negative correlation between number of true associations and long-range false associations, suggesting (rare) short-range false positives could be due to linkage while long-range false positive could be due to drift alone.

**Figure S6:** **Scatterplots of the proportion of significantly-differentiated SNPs versus the proportion of low-starting-frequency SNPs (allele frequency < 0.1) across 100 Kbp genomic windows for each chromosome and each replicate line.** Between 235 and 320 windows were examined per chromosome. Dashed pink line and solid red line show linear relationships for the full dataset, and excluding windows without significantly-differentiated SNPs, respectively. Excluding windows without significantly-differentiated SNPs left between 3 and 230 datapoints. R-squared and significance values are reported where *P* < 0.05. Note y-axis scale differs between panels.

**Figure S7:** **Observed lab adaptation candidate SNP overlap compared to expected.** Histograms of the distribution of overlap (no. shared SNPs) among all replicate combinations of the simulated control-replicate gene lists. The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of genes shared between the observed control candidate lists, labelled as a percentile of the normal distribution.

**Figure S8: Observed desiccation candidate SNP overlap compared to expected.** Histograms of the distribution of overlap (no. shared SNPs) among all replicate combinations of the simulated desiccation-replicate gene lists. The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of genes shared between the observed desiccation candidate lists, labelled as a percentile of the normal distribution.

**Figure S9:** **Enrichment of candidate SNPs in genomic location categories.** The arrows show the difference in frequency of SNPs mapping to a genomic location type, between the list of all SNPs and the list of candidate SNPs only. Black arrows indicate this difference was significant by a χ2 test (*P* < 0.05), grey arrows indicate no significant difference. Separate calculations were performed for each control and desiccation-line replicate comparison between the mass-bred and generation 13.

**Figure S10:** **Enrichment of candidate SNPs in genetic effect categories.** The arrows show the difference in frequency of SNPs mapping to a genomic effect type, between the list of all SNPs and the list of candidate SNPs (only those within 1000 bp of an annotated gene in each case). Low: either a synonymous variant, a variant producing a premature start codon, a variant located near but not within a splice site (within 1-3 bases of an exon or 3-8 bases of an intron), or a synonymous or non-synonymous mutation affecting a start or stop codon; Moderate: a non-synonymous coding variant; High: variant causing loss of a start codon, loss or gain of a stop codon or loss or gain of a splice donor or acceptor site (2 bases after coding exon start or 2 bases before coding exon end); Modifier: any other variant, including upstream, downstream, 5’- and 3’-UTR locations. Black arrows indicate this difference was significant by a χ2 test (*P* < 0.05), grey arrows indicate no significant difference. Separate calculations were performed for each control and desiccation-line replicate.

**Figure S11:** **Observed eQTL/veQTL overlap compared to expected.** Histograms of the null distribution of the number of candidate SNPs expected to hit known eQTL/veQTL loci (Huang *et al.* 2015) for each control (A-E) and desiccation-selected (F-J) replicate line. The histogram shows the distribution of eQTL/veQTL SNP numbers for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of known eQTL/veQTL loci hit in the real candidate SNP lists, labelled as a percentile of the normal distribution.

**Figure S12:** **Observed eQTL/veQTL target gene overlap compared to expected.** Histograms of the distribution of overlap (no. shared genes targetted by candidate eQTL/veQTL SNPs) among all replicate combinations of the simulated desiccation-candidate SNP lists. The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of genes shared between the observed desiccation-candidate eQTL/veQTL target lists, labelled as a percentile of the normal distribution.

**Figure S13:** **Heatmap of lab adaptation gene overlap among replicates.** This shows the overlap in putative lab-adaptation genes among the control (C1-C5) and desiccation-selected (D1-D5) replicates. Red indicates that gene was significantly differentiated between mass-bred and generation 13 pool for that line; grey indicates no significant differentiation.

**Figure S14:** **Observed lab adaptation gene occurrence in desiccation replicates compared to expected.** Histograms of the distribution of overlap (no. shared genes) among each simulated desiccation-candidate gene list and the corresponding simulated putative lab-adaptation gene list. The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of genes shared between the observed desiccation-candidate list and the putative lab-adaptation list, labelled as a percentile of the normal distribution.

**Figure S15:** **Observed lab adaptation candidate gene overlap compared to expected.** Histograms of the distribution of overlap (no. shared genes) among all replicate combinations of the simulated control-replicate gene lists. The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of genes shared between the observed control candidate lists, labelled as a percentile of the normal distribution.

**Figure S16**: **Observed desiccation candidate gene overlap compared to expected.** Histograms of the distribution of overlap (no. shared genes) among all replicate combinations of the simulated desiccation-candidate gene lists. The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the number of genes shared between the observed desiccation-candidate lists, labelled as a percentile of the normal distribution.

**Figure S17:** **Heatmap of desiccation candidate gene overlap among replicates.** Heatmap showing the overlap in putative desiccation-resistance candidate among the desiccation-selected (D1-D5) replicates. Red indicates that gene was significantly differentiated between mass-bred and generation 13 pool for that replicate; grey indicates no significant differentiation.

**Figure S18**: **Comparison of gene list overlap** **due to single-mapping SNPs (SNPs that were mapped to only one gene) and overlap due to multi-mapping SNPs (SNPs that mapped to more than one gene).** Each point shows the level of overlap (number of overlapping genes) calculated for one possible pair, trio, quartet or the overlap among all 5 replicates in the control (panel A) or selected line (panel B) treatment. The x- and y-axis are log10-transformed in panel B for ease of visualisation. Grey dashed line shows the expected slope of 1, i.e. no difference between the single- and multi-mapping SNP results on the extent of overlap. The red line in panel B shows the line of best fit from observed points.

**Figure S19: Variance in allele frequency among replicates binned by starting allele frequency.** Mean standardized variance *f* versus starting allele frequency among replicate generation 13 control lines (blue) and desiccation-selected lines (red) measured at desiccation candidate loci (‘des’), lab-adaptation candidate loci (‘lab’) or a random subset of neutral loci (‘neutral’). Loci were assigned to bins of width 0.05 by starting (mass-bred) allele frequency, shown on the x-axis.

**Figure S20: Observed protein-protein interaction network overlap compared to expected based on SNP resampling.** Histograms of zero-order PPI network size, number of listed genes with known interactions, and network density for simulated networks. Size is represented as node count (A-E) and edge count (F-J) distributions simulated by resampling from the entire mapped *D. melanogaster* gene list (r6.01). Number of resampled genes with known interactions (i.e. genes that appear in the background *Drosophila* PPI network) is shown in panels K-O. Network density (the average connectivity across all nodes) is shown in panels P-T. For network density only, the background simulations were conducted by selecting subnetworks from the background network of the same size (same no. nodes) as the observed network. For all panels, the histogram shows the distribution of size measure for 1000 simulations. Distribution density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the size of the observed network, labelled as a percentile of the normal distribution.

**Figure S21:** **Observed protein-protein interaction network overlap compared to expected based on gene resampling.** Histograms of zero-order PPI network overlap node count (A-J, S-AD) and edge count (K-T, AE-AN) distributions simulated by resampling from the entire mapped *D. melanogaster* gene list (r6.01). The histogram shows the distribution of overlap measures for 1000 simulations. Density is shown by the black line and the corresponding normal distribution is shown by the blue line. The red vertical line shows the overlap observed among the real networks, labelled as a percentile of the normal distribution.

**Figure S22:** **Relationship between gene length and likelihood of representation in a candidate gene list.** For each panel, the number of cases (out of 1000) in which a gene appeared in the simulated candidate list is plotted against the gene’s length (in base pairs, including 1000 bp up- and down-stream). Grey points are genes not hit in the observed candidate gene list; red points are genes observed in the real candidate gene list. Results are shown for all individual control (A-E) and desiccation-selected replicate lines (F-J) as well as for the lists of genes overlapping among replicates. Putative lab adaptation genes were removed from desiccation-selected gene lists in each simulation and in the observed data.

**Table S1:** **Summary statistics for each replicate line.** Total raw read count, mean depth of genome coverage, number of SNPs called, Bonferroni-corrected *P*-value threshold, number of candidate SNPs, number of candidate genes and zero-order PPI network sizes for each replicate line.

**Table S2**: **Power analysis.** Posterior probability of identifying at least one SNP association with desiccation tolerance in a simulated causal gene (True Positive Rate), a neutral gene located ~10Kbp from the causal variant (short-range False Positive Rate), or a neutral gene located ~1Mbp from the causal variant (long-range False Positive Rate). The results are based on 2 million independent simulations under the model presented in File S1.

**Table S3:** **Inversion frequency results.** Frequencies of diagnostic SNPs assayed for various chromosomal inversions in the mass-bred and all generation 13 replicates.

**Table S4:** **Desiccation candidate SNPs with putative high impact**. SNPs identified as putative high-impact by SNPEff (Cingolani *et al.* 2012) that increased significantly in frequency between the mass-bred and generation 13 in the desiccation replicates. These SNPs caused the loss or gain of a stop codon, splice acceptor or donor site and hence may impact protein function. Gene function is presented where this is known, using information from Flybase (Santos *et al.* 2015) and references cited.