**Supplementary Materials for "Social environment influences the relationship between genotype and gene expression in wild baboons"**

**DE Runcie, R Wiedmann, EA Archie, J Altmann, GA Wray, SC Alberts, and J Tung**

**Supplementary Methods: SNP ascertainment and genotyping**

We used the Genome Analysis Toolkit (GATK) for variant discovery and genotyping of the target flanking regions for the 96 individuals. We first used PicardTools (MarkDuplicates) to remove potential PCR duplicate reads from the overall read set for each individual. Then, we performed five rounds using the following GATK tools to identify variable sites in the set of 35 target regions: 1) RealignerTargetCreator, 2) IndelRealigner, 3) CountCovariates, 4) TableRecalibration, 5) AnalyzeCovariates, 6) UnifiedGenotyper, 7) VariantFiltration. For all but 3 samples, Steps 1-5 were run on each sample individually; for the three samples with the fewest total reads, we merged the samples prior to running this pipeline for computational efficiency. Steps 6-7 were run on all samples at the same time. We repeated these procedures five times because GATK relies on a set of known SNPs for accurately recalibrating quality scores at each sequenced site. Relatively few SNPs have been identified in baboons, particularly in our target regions, which would lead to assignment of recalibrated quality scores that were too low relative to the actual quality of the sequencing (all SNPs would be assumed to be sequencing errors in the recalibration pipeline). Thus, for the first round, we ran the pipeline without base quality recalibration (steps 3-5). For each subsequent round, we treated the set of high quality SNPs inferred in the previous round as "known" SNPs. To define high quality SNPs, we used the following GATK filters: QD < 2.0, MQ < 40.0, FS > 60.0, HaplotypeScore > 13.0, MQRankSum < -12.5, ReadPosRankSum < -8.0. After five repetitions, the analysis converged to a consensus number of variable sites (i.e., we no longer identified additional SNPs beyond those treated as known variants).

This analysis resulted in a high quality set of SNPs containing 3527 high confidence segregating sites (among 250,333 total base pairs for which at least 10 individuals were sequenced at a coverage of at least 30x). Among these sites, which were the sites in which SNPs were most likely to be called with high confidence, we therefore identified a total of one variant per ~70 base pairs (including both rare and common variants). This large number of sites is due to our sample size (95 individuals, after dropping one individual for whom we did not obtain many sequencing reads, or 190 chromosomes): extensive resequencing studies in humans show that the number of newly discovered SNPs increases approximately linearly, for both common and rare SNPs, within this range of samples (the number of newly discovered rare SNPs continues to increase linearly with sample size even into sample sizes >2000 individuals: Tennessen et al 2012, *Science* 337: 64 - 69). Two additional diagnostics also indicate that the SNP data set is high quality. First, SNPs in the data set exhibit a transition-transversion ratio (Ti/Tv) of 2.26, close to the genome-wide Ti/Tv ratio for highly filtered Illumina data on humans, with SNPs called at a false discovery rate of 1% (Ti/Tv = 2.15: DePristo et al 2011, *Nature Genetics* 43: 491 - 498). Poorer quality SNP data sets (i.e., pre-filtering) exhibit substantially lower Ti/Tv ratios. Second, the overall data set yields an estimate of pi, a measure of population genetic diversity, within the range of pi values for gene regulatory regions in this population based on Sanger sequencing data. Here, we estimate pi = 0.003; from Sanger sequencing of 22 regulatory regions, pi = 0.002 +/- 0.002 s.d. (Tung et al 2009, *Nature* 460: 388-391).

For missing genotypes in the data set, we used the Bayesian imputation program BEAGLE to infer missing data, which occurred at individual-site combinations with too low coverage to call genotypes with high confidence. Local linkage disequilibrium sometimes allows these sites to be imputed with high confidence, if other linked sites in the same individual have been successfully genotyped. To perform imputation, we used the genotype likelihoods produced by GATK as input into BEAGLE, and then filtered the post-imputation results to include only genotypes with posterior probability >0.98. Target regions for each gene were analyzed separately, as genes are not expected to be in high LD with one another (and are therefore mutually uninformative with respect to imputing missing sites). Even after imputation, very few genotypes were called with high confidence for one individual (FACE) for whom we had the fewest number of reads.

**Supplementary Methods: Potential confounds in the GEI analysis**

GEIs are indicated by cases in which the magnitude of ASGE is correlated with environmental variation, which suggest that environmental effects influence the expression levels of the two alleles of a gene differently (i.e., the environment interacts with genetic variation). We hypothesized that, in these cases, individuals heterozygous for a functional *cis*-regulatory variant (indexed by the SNP that best partitioned overall variance in the ASGE data) would exhibit different levels of ASGE depending on environmental conditions, provided that the variant was also linked to the ASGE assay SNP. This hypothesis assumes that the social environmental variables of interest are not correlated with genetic background. If this is not the case, then apparent GEIs could actually represent cases of gene-gene interactions instead of gene-environment interactions. To exclude the likelihood that the GEIs we detected are in fact due to gene-gene interactions, we considered two major sources of genetic background in this population: *i)* admixture-related genetic background, as some baboons in Amboseli are hybrids between yellow baboons and anubis baboons (Tung et al 2008, *Molecular Ecology* 17: 1998 – 2011); and *ii)* kinship.

*Admixture.* Three lines of evidence suggest that our results are not confounded by admixture-related genetic background, assessed using the microsatellite marker-based approach in Tung et al (2008). First, there is no correlation between genetic estimates of admixture proportions and any of the environmental variables significantly involved in GEIs (p-values for these variables range from 0.06 – 0.85, and the maximum $R^2$ for all of these correlations, for female social connectedness, is small: $R^2 = 0.042$, $p = 0.15$). Second, including these estimates as a covariate in our GEI models does not qualitatively change the results. Finally, admixture estimates do not correlate with ASGE for any gene in our analysis ($p > 0.38$).

*Kinship.* Within social groups, females tend to be similar in rank to their mothers, potentially producing a close relationship between kinship and female dominance rank. However, we sampled individuals from seven different social groups, and females with similar ranks in different social groups are not expected to be closely related. Thus, although females with more similar ranks in the overall data set do tend to be slightly more closely related (correlation between pairwise relatedness derived from known pedigrees and pairwise rank similarity for females: $p = 0.0004$), rank similarity explains only 1% of variation in relatedness, and is therefore unlikely to confound our results. Similarly, relatedness explains less than 1% of variation in maternal dominance rank ($p = 0.0003$; $R2 = 0.0064$) and female social connectedness ($p = 0.083$; $R2= 0.002$), and there is no relationship between pairwise relatedness and pairwise similarity in male rank ($p = 0.582$) or male social connectedness ($p = 0.573$).

Social groups themselves could also contribute to population structure in our sample, which could affect social environmental measures that involve group-level characteristics. In our analysis, this condition could apply to group size, if groups that are genetically more similar also tend to be of similar size. However, our sample does not exhibit substantial population structure explained by group identity (Figure S1). Additionally, there is no strong relationship between the major axes of population structure and group size (Figure S2): among the top 10 principal components of the SNP data, none are significantly correlated with group size after multiple hypothesis correction. PC1 is correlated with group size at a nominal $p = 0.02$, but group size explains only 6% of the variation in PC1, which itself explains only 7.3% of the variation in the overall data set (i.e., group size explains <1% of genetic variation in the sample). This lack of strong social group-mediated population structure in the sample likely arises from the high rate of migrant exchange between social groups due to (repeated) male dispersal.

## Table S1. Number of individuals assayed and SNPs tested for each gene

| Gene | # Males[1] | # Females[1] | Total SNPs[2] | Best SNP[3] | Model R2[4] |
|---|---|---|---|---|---|
| ***AIM2*** | 7 | 9 | 5 | AssaySNP | 0.8831643 |
| ***APOBEC3G*** | 24 | 20 | 78 | AssaySNP | 0.853127 |
| ***CCR1*** | 16 | 17 | 7 | AssaySNP | 0.4296628 |
| ***CD5*** | 9 | 12 | 2 | S_12611661 | 0.318564 |
| ***CD58*** | 7 | 3 | 0 | AssaySNP | 0.6864284 |
| ***CD8A*** | 22 | 22 | 18 | S_85851418 | 0.5543711 |
| ***CD9*** | 9 | 13 | 11 | S_6281260 | 0.8139922 |
| ***CLC*** | 9 | 9 | 5 | AssaySNP | 0.7199155 |
| ***CXCR4*** | 21 | 24 | 28 | AssaySNP | 0.3709335 |
| ***CYP17A1*** | 8 | 10 | 0 | NONE | 0.0792132 |
| ***DAP*** | 17 | 23 | 27 | S_10531885 | 0.6178548 |
| ***GBP1*** | 18 | 21 | 15 | AssaySNP | 0.7546709 |
| ***GIMAP2*** | 23 | 19 | 6 | AssaySNP | 0.3707347 |
| ***GZMA*** | 12 | 25 | 13 | S_51139696 | 0.5538139 |
| ***HLA-F*** | 24 | 21 | 8 | S_29346355 | 0.3599964 |
| ***IFNAR1*** | 16 | 23 | 1 | AssaySNP | 0.8647966 |
| ***IL10*** | 14 | 12 | 3 | AssaySNP | 0.5763804 |
| ***IL4R*** | 19 | 23 | 12 | AssaySNP | 0.3019124 |
| ***IL6*** | 6 | 4 | 3 | AssaySNP | 0.7829775 |
| ***IRF2*** | 12 | 21 | 22 | AssaySNP | 0.3198682 |
| ***KLF6*** | 10 | 17 | 8 | AssaySNP | 0.6517723 |
| ***LGALS3*** | 17 | 17 | 17 | S_111537912 | 0.3394394 |
| ***MNDA*** | 14 | 17 | 8 | S_132726692 | 0.795599 |
| ***MSR1*** | 18 | 16 | 11 | S_14106376 | 0.6950351 |
| ***MSRA*** | 13 | 19 | 19 | S_8851908 | 0.5362489 |
| ***NMI*** | 13 | 21 | 5 | AssaySNP | 0.2747337 |
| ***OAS2*** | 18 | 21 | 10 | S_112177655 | 0.5412198 |
| ***OASL*** | 19 | 15 | 3 | AssaySNP | 0.2213378 |
| ***PHF11*** | 12 | 14 | 4 | AssaySNP | 0.7464097 |
| ***PRF1*** | 7 | 11 | 3 | S_59137422 | 0.5318094 |
| ***PRNP*** | 13 | 20 | 6 | AssaySNP | 0.5421846 |
| ***RNASE2*** | 18 | 15 | 7 | S_77517141 | 0.786341 |
| ***S100A9*** | 7 | 13 | 9 | AssaySNP | 0.4413103 |
| ***SLA*** | 11 | 20 | 12 | AssaySNP | 0.5917604 |
| ***SLAMF7*** | 16 | 12 | 0 | AssaySNP | 0.4248022 |

[1]In some cases, the sample sizes for each gene were reduced in GEI analyses due to missing environmental data.

[2]Number of SNPs in the resequenced region with complete genotypes for all individuals and at least 5 heterozygous and 5 homozygous genotypes. A model incorporating genotype at the pyrosequencing assay SNP (AssaySNP) was also tested for all genes.

[3]SNP that best partitioned ASGE data by heterozygous/homozygous genotype. AssaySNP is the SNP used in the ASGE assay and was required to be heterozygous. Other SNPs were in the resequenced regions and are named according to their chromosomal location in *Panu2.0*, except for *CYP17A1*, for which we identified no SNPs that partitioned the ASGE data well.

[4]Fraction of variance in ASGE levels explained by the model that partitioned the data on heterozygosity versus homozygosity at the best SNP. All models assume that homozygotes have mean ASGE = 0 and heterozygotes have mean ASGE $\neq$ 0. $R^2$ is calculated as the fraction change in variance relative to a model with no partition and global mean = 0.

## Table S2. SureSelect sequencing targets, by gene

| Gene | Chromosome[1] | Start[1] | End[1] | Length | Median depth[2] | Median % coverage[3] | Median coverage length[4] |
|---|---|---|---|---|---|---|---|
| *AIM2* | chr1 | 132955593 | 132964498 | 8905 | 26428 | 0.71089275 | 6330.499939 |
| *APOBEC3G* | chr10 | 79180057 | 79300153 | 120096 | 44114 | 0.2169223 | 26051.50054 |
| *CCR1* | chr2 | 89160123 | 89169348 | 9225 | 18998.5 | 0.4441734 | 4097.499615 |
| *CD5* | chr14 | 12610915 | 12612817 | 1902 | 3597.5 | 0.8267613 | 1572.499993 |
| *CD58* | chr1 | 117335732 | 117370301 | 34569 | 74809.5 | 0.4198849 | 14515.00111 |
| *CD8A* | chr13 | 85849748 | 85858937 | 9189 | 35535.5 | 0.94912395 | 8721.499977 |
| *CD9* | chr11 | 6276833 | 6284714 | 7881 | 17203 | 0.69337645 | 5464.499802 |
| *CLC* | chr19 | 33978585 | 34069409 | 90824 | 18985.5 | 0.0921893 | 8373.000983 |
| *CXCR4* | chr13 | 106907441 | 106916614 | 9173 | 43677 | 0.9824485 | 9012.000091 |
| *CYP17A1* | chr9 | 94642524 | 94644509 | 1985 | 7188.5 | 0.82493705 | 1637.500044 |
| *DAP* | chr6 | 10527992 | 10537175 | 9183 | 29481 | 0.9560601 | 8779.499898 |
| *GBP1* | chr1 | 89753327 | 89815458 | 62131 | 56381 | 0.25832515 | 16049.99989 |
| *GIMAP2* | chr3 | 171952919 | 171962210 | 9291 | 16912.5 | 0.4336993 | 4029.500196 |
| *GZMA* | chr6 | 51130538 | 51139709 | 9171 | 26794 | 0.77679645 | 7124.000243 |
| *HLA-F* | chr4 | 29337813 | 29347221 | 9408 | 14933 | 0.5085034 | 4783.999987 |
| *IFNAR1* | chr3 | 13197137 | 13206645 | 9508 | 11005.5 | 0.42306475 | 4022.499643 |
| *IL10* | chr1 | 157147958 | 157157149 | 9191 | 39402 | 0.9652921 | 8871.999691 |
| *IL4R* | chr20 | 24996265 | 25005057 | 8792 | 10666 | 0.5444154 | 4786.500197 |
| *IL6* | chr3 | 88209583 | 88218789 | 9206 | 52410.5 | 0.9994569 | 9201.000221 |
| *IRF2* | chr5 | 173549606 | 173556760 | 7154 | 21244.5 | 0.87517475 | 6261.000162 |
| *KLF6* | chr9 | 3931870 | 3941043 | 9173 | 21298.5 | 0.6321269 | 5798.500054 |
| *LGALS3* | chr7 | 111534356 | 111542889 | 8533 | 33693 | 0.9697644 | 8274.999625 |
| *MNDA* | chr1 | 132719928 | 132729183 | 9255 | 22194.5 | 0.53922205 | 4990.500073 |
| *MSR1* | chr8 | 14099697 | 14109244 | 9547 | 18311.5 | 0.5683461 | 5426.000217 |
| *MSRA* | chr8 | 8851858 | 8861084 | 9226 | 26998 | 0.7217646 | 6659.0002 |
| *NMI* | chr12 | 13558857 | 13567522 | 8665 | 8067 | 0.3369302 | 2919.500183 |
| *OAS2* | chr11 | 112176742 | 112185686 | 8944 | 14174.5 | 0.5380702 | 4812.499869 |
| *OASL* | chr11 | 120304339 | 120313512 | 9173 | 13571 | 0.40711875 | 3734.500294 |
| *PHF11* | chr17 | 27796800 | 27806083 | 9283 | 26405 | 0.84175375 | 7814.000061 |
| *PRF1* | chr9 | 59130679 | 59139400 | 8721 | 13692 | 0.495356 | 4319.999676 |
| *PRNP* | chr10 | 29256070 | 29264681 | 8611 | 24881.5 | 0.8659273 | 7456.49998 |
| *RNASE2* | chr7 | 77516298 | 77570255 | 53957 | 135740 | 0.48084585 | 25944.99953 |
| *S100A9* | chr1 | 127176726 | 127185598 | 8872 | 27742.5 | 0.8321686 | 7382.999819 |
| *SLA* | chr8 | 128026290 | 128035625 | 9335 | 32572.5 | 0.8619175 | 8045.999863 |
| *SLAMF7* | chr1 | 134666586 | 134674540 | 7954 | 14467.5 | 0.56675885 | 4507.999893 |

[1]Chromosome, start, and end coordinates are relative to the current baboon genome assembly: Panu2.0.
[2]Median depth refers to the median number of reads mapped within each genomic region across samples.
[3]Coverage percentage is the percentage of bases in each region covered by at least 1 read in a single sample. Median % coverage is the median of this value across samples.
[4]Coverage length is the number of bases in each target region covered by at least 1 read in a single sample. Median coverage length is the median of this value across samples.

## Table S3. Summary of sequencing reads generated for each individual.

| Individual | Total | Total mapped | MAPQ >= 20 | Uniquely mapped | MAPQ >= 20 in target regions | Percent MAPQ >= 20 | Percent MAPQ >= 20 in target region |
|---|---|---|---|---|---|---|---|
| ABBY | 1495197 | 1284435 | 1263648 | 1284435 | 905249 | 0.845 | 0.716 |
| ADRIANO | 1798096 | 1477988 | 1445153 | 1477988 | 1071758 | 0.804 | 0.742 |
| ALEX | 2001812 | 1707333 | 1678424 | 1707333 | 1212607 | 0.838 | 0.722 |
| AMOK | 2329111 | 1892345 | 1833991 | 1892345 | 1181222 | 0.787 | 0.644 |
| APOLLO | 2121523 | 1559974 | 1464098 | 1559974 | 686488 | 0.690 | 0.469 |
| ARAB | 3409925 | 2598523 | 2452965 | 2598523 | 1175920 | 0.719 | 0.479 |
| ARSENAL | 3779403 | 3124554 | 3046763 | 3124554 | 2115252 | 0.806 | 0.694 |
| BAGINE | 2490863 | 1702171 | 1553332 | 1702171 | 484762 | 0.624 | 0.312 |
| CABANA | 937243 | 770628 | 749770 | 770628 | 534322 | 0.800 | 0.713 |
| CADET | 1529456 | 1281390 | 1252588 | 1281390 | 915867 | 0.819 | 0.731 |
| DAGGER | 1258604 | 835379 | 750164 | 835379 | 183762 | 0.596 | 0.245 |
| DASTON | 1583926 | 1333211 | 1305471 | 1333211 | 980207 | 0.824 | 0.751 |
| DRONGO | 1904545 | 1596975 | 1564779 | 1596975 | 1124503 | 0.822 | 0.719 |
| DUBU | 1939102 | 1640573 | 1603934 | 1640573 | 1104390 | 0.827 | 0.689 |
| DUNLIN | 1891468 | 1580926 | 1546046 | 1580926 | 1168150 | 0.817 | 0.756 |
| EAGLE | 1701029 | 1434898 | 1405878 | 1434898 | 1031456 | 0.826 | 0.734 |
| EVA | 1221561 | 1020398 | 998907 | 1020398 | 757724 | 0.818 | 0.759 |
| FACE | 4711 | 3929 | 3825 | 3929 | 2729 | 0.812 | 0.713 |
| GABRIEL | 2738398 | 2302724 | 2257486 | 2302724 | 1651088 | 0.824 | 0.731 |
| GANJA | 1812979 | 1535843 | 1506154 | 1535843 | 1139160 | 0.831 | 0.756 |
| HAPPY | 2208938 | 1880251 | 1843191 | 1880251 | 1297384 | 0.834 | 0.704 |
| HIBISCUS | 1883451 | 1580385 | 1547114 | 1580385 | 1235748 | 0.821 | 0.799 |
| HOLLAND | 1892393 | 1594988 | 1562216 | 1594988 | 1093883 | 0.826 | 0.700 |
| HONEY | 1750707 | 1384873 | 1322520 | 1384873 | 550086 | 0.755 | 0.416 |
| HYMN | 1742364 | 1424267 | 1367784 | 1424267 | 384530 | 0.785 | 0.281 |
| JAGGER | 2133118 | 1725098 | 1674126 | 1725098 | 1170823 | 0.785 | 0.699 |
| KATHRYN | 1734508 | 1456590 | 1429459 | 1456590 | 1027413 | 0.824 | 0.719 |
| KELLY | 1598776 | 1249200 | 1193960 | 1249200 | 696801 | 0.747 | 0.584 |
| KERNEL | 1695791 | 1429004 | 1399530 | 1429004 | 1025653 | 0.825 | 0.733 |
| KOLA | 1911984 | 1642751 | 1609226 | 1642751 | 1082691 | 0.842 | 0.673 |
| LANCASTER | 1793612 | 1512642 | 1483932 | 1512642 | 1129860 | 0.827 | 0.761 |
| LAOS | 1585401 | 1348401 | 1324626 | 1348401 | 1035298 | 0.836 | 0.782 |
| LAX | 1651447 | 1381654 | 1351408 | 1381654 | 1015551 | 0.818 | 0.751 |
| LAZA | 1544048 | 1296925 | 1263368 | 1296925 | 826936 | 0.818 | 0.655 |
| LEBANON | 1641208 | 1396178 | 1368273 | 1396178 | 1049222 | 0.834 | 0.767 |
| LIBERTY | 2310008 | 1925813 | 1886916 | 1925813 | 1461683 | 0.817 | 0.775 |
| LIMAU | 1347576 | 1168852 | 1150994 | 1168852 | 831181 | 0.854 | 0.722 |
| LIWAZA | 2113135 | 1819200 | 1790803 | 1819200 | 1359548 | 0.847 | 0.759 |
| LIZZY | 1729591 | 1479438 | 1454425 | 1479438 | 1028346 | 0.841 | 0.707 |
| LOBO | 1848783 | 1515929 | 1476402 | 1515929 | 1105407 | 0.799 | 0.749 |
| LOCUST | 1889743 | 1612426 | 1581762 | 1612426 | 1233640 | 0.837 | 0.780 |
| LOFTY | 1954771 | 1654565 | 1624983 | 1654565 | 1259010 | 0.831 | 0.775 |
| LOGAN | 1552940 | 1278863 | 1251634 | 1278863 | 903802 | 0.806 | 0.722 |
| LOLLIPOP | 1540787 | 1288010 | 1260253 | 1288010 | 913853 | 0.818 | 0.725 |
| LOZENGE | 2136115 | 1710896 | 1654864 | 1710896 | 1077508 | 0.775 | 0.651 |
| LUIGI | 2024469 | 1671928 | 1637009 | 1671928 | 1202670 | 0.809 | 0.735 |
| LUNA | 1325864 | 1110374 | 1090682 | 1110374 | 762428 | 0.823 | 0.699 |
| LURCH | 2336220 | 1862254 | 1799145 | 1862254 | 1197373 | 0.770 | 0.666 |
| LUTHER | 1857327 | 1521234 | 1475632 | 1521234 | 1020158 | 0.794 | 0.691 |
| LYENA | 2515672 | 2087527 | 2041782 | 2087527 | 1505985 | 0.812 | 0.738 |
| LYME | 3037665 | 2509487 | 2455186 | 2509487 | 1816432 | 0.808 | 0.740 |
| MBEGU | 2490597 | 2076865 | 2026238 | 2076865 | 1501781 | 0.814 | 0.741 |
| MONGOOSE | 1597737 | 1359409 | 1336468 | 1359409 | 1004937 | 0.836 | 0.752 |
| MORRIS | 130063 | 109486 | 107222 | 109486 | 80988 | 0.824 | 0.755 |
| NARASHA | 1760026 | 1452741 | 1418760 | 1452741 | 1012058 | 0.806 | 0.713 |
| NAWA | 1652478 | 1391371 | 1363432 | 1391371 | 927442 | 0.825 | 0.680 |
| NIKE | 2984474 | 2435007 | 2360916 | 2435007 | 1582831 | 0.791 | 0.670 |
| NJUGU | 1951141 | 1644186 | 1612064 | 1644186 | 1159629 | 0.826 | 0.719 |
| NOBEL | 1981065 | 1606228 | 1554598 | 1606228 | 997006 | 0.785 | 0.641 |
| NOODLE | 2514187 | 2103848 | 2059431 | 2103848 | 1494440 | 0.819 | 0.726 |
| NOZZLE | 2140135 | 1813680 | 1771657 | 1813680 | 1219592 | 0.828 | 0.688 |
| NUTTY | 2512710 | 2019505 | 1951954 | 2019505 | 1271117 | 0.777 | 0.651 |
| NYUKI | 1998866 | 1650531 | 1605480 | 1650531 | 1025602 | 0.803 | 0.639 |
| OCTAGON | 2070496 | 1746428 | 1705866 | 1746428 | 1194235 | 0.824 | 0.700 |
| ODESSA | 1707076 | 1418214 | 1382300 | 1418214 | 973885 | 0.810 | 0.705 |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| OJUANG | 2328275 | 1922423 | 1870298 | 1922423 | 1236535 | 0.803 | 0.661 |
| OPHELIA | 2176368 | 1792728 | 1748752 | 1792728 | 1222978 | 0.804 | 0.699 |
| OXYGEN | 2607572 | 2204115 | 2158514 | 2204115 | 1546290 | 0.828 | 0.716 |
| PAISLEY | 1524350 | 1260078 | 1229222 | 1260078 | 848766 | 0.806 | 0.690 |
| PITTSBURGH | 2749209 | 2196233 | 2108937 | 2196233 | 1196407 | 0.767 | 0.567 |
| PLATO | 2481620 | 2079009 | 2030144 | 2079009 | 1389591 | 0.818 | 0.684 |
| QUASI | 2826516 | 2298516 | 2228635 | 2298516 | 1386191 | 0.788 | 0.622 |
| RAJA | 1959767 | 1608845 | 1571956 | 1608845 | 1109879 | 0.802 | 0.706 |
| RHODA | 1566052 | 1304426 | 1273013 | 1304426 | 885832 | 0.813 | 0.696 |
| RWANDA | 1575757 | 1319767 | 1293848 | 1319767 | 998054 | 0.821 | 0.771 |
| SCENIC | 2455244 | 2038224 | 1986074 | 2038224 | 1399561 | 0.809 | 0.705 |
| SORGHUM | 1426866 | 1214169 | 1193727 | 1214169 | 895523 | 0.837 | 0.750 |
| TALBOT | 922880 | 768311 | 753614 | 768311 | 580537 | 0.817 | 0.770 |
| VAIN | 1828717 | 1498088 | 1461161 | 1498088 | 1048856 | 0.799 | 0.718 |
| VAPOUR | 2141347 | 1832473 | 1799016 | 1832473 | 1385872 | 0.840 | 0.770 |
| VEIL | 2309459 | 1954522 | 1915038 | 1954522 | 1346214 | 0.829 | 0.703 |
| VET | 1660238 | 1120578 | 1017076 | 1120578 | 391977 | 0.613 | 0.385 |
| VIBRANT | 778148 | 663835 | 651911 | 663835 | 511036 | 0.838 | 0.784 |
| VOTE | 1713897 | 1477187 | 1452205 | 1477187 | 1123113 | 0.847 | 0.773 |
| VOYAGE | 1184658 | 1003416 | 981830 | 1003416 | 746491 | 0.829 | 0.760 |
| WABASH | 1964590 | 1666075 | 1636388 | 1666075 | 1169307 | 0.833 | 0.715 |
| WASP | 1214645 | 1025543 | 1004707 | 1025543 | 774576 | 0.827 | 0.771 |
| WENDY | 3632203 | 2934107 | 2836423 | 2934107 | 1841581 | 0.781 | 0.649 |
| WIPER | 951874 | 815014 | 799850 | 815014 | 626898 | 0.840 | 0.784 |
| WIRE | 1929046 | 1622437 | 1591236 | 1622437 | 1197723 | 0.825 | 0.753 |
| WIVU | 1790159 | 1551534 | 1524955 | 1551534 | 1109327 | 0.852 | 0.727 |
| WIZARD | 1536865 | 1285554 | 1254572 | 1285554 | 878552 | 0.816 | 0.700 |
| WRINKLE | 2582734 | 2163015 | 2115896 | 2163015 | 1490519 | 0.819 | 0.704 |
| WYNNE | 1455187 | 1239105 | 1214818 | 1239105 | 936456 | 0.835 | 0.771 |
| YAI | 1522068 | 1303177 | 1278714 | 1303177 | 954547 | 0.840 | 0.746 |
| YOGI | 1742380 | 1503305 | 1477320 | 1503305 | 1097120 | 0.848 | 0.743 |
| | | | | | | | |
| median | 1853055 | 1543688.5 | 1495043 | 1543688.5 | 1080099.5 | 0.819 | 0.717 |
| sd | 602151.4965 | 487903.969 | 474198.3193 | 487903.969 | 349100.5869 | 0.044 | 0.103 |
| min | 4711 | 3929 | 3825 | 3929 | 2729 | 0.596 | 0.245 |
| max | 3779403 | 3124554 | 3046763 | 3124554 | 2115252 | 0.854 | 0.799 |

## Table S4. GEI model results.[1]

| Sex | Environment | Df_N[2] | DF_R[3] | Likelihood ratio | p-value | BH q-value[4] |
|---|---|---|---|---|---|---|
| Males | age at darting | 30 | 329 | 28.54372633 | 0.541647258 | 0.541647258 |
| Males | maternal rank | 24 | 205 | 40.32814378 | 0.019714451 | 0.043809891 |
| Males | group size at birth | 24 | 205 | 24.51213285 | 0.432645591 | 0.480717323 |
| Males | group size at darting | 30 | 300 | 55.79755812 | 0.002880249 | 0.014401247 |
| Males | male dominance rank | 30 | 300 | 33.90377159 | 0.284696327 | 0.355870409 |
| Males | male social connectedness | 28 | 278 | 56.44851553 | 0.001133818 | 0.01133818 |
| Males | maternal social connectedness | 24 | 193 | 28.57051048 | 0.23671417 | 0.315618894 |
| Females | age at darting | 31 | 381 | 58.04568673 | 0.002277703 | 0.014401247 |
| Females | maternal rank | 31 | 355 | 31.04701475 | 0.46384505 | 0.488257947 |
| Females | group size at birth | 31 | 381 | 63.78119868 | 0.000472557 | 0.00945114 |
| Females | group size at darting | 31 | 381 | 52.53741093 | 0.009199003 | 0.030663344 |
| Females | female dominance rank | 31 | 381 | 39.80571732 | 0.13348055 | 0.1906865 |
| Females | female social connectedness | 31 | 381 | 50.7289888 | 0.014147561 | 0.035368904 |
| Females | maternal social connectedness | 30 | 336 | 39.14795398 | 0.122448536 | 0.188382363 |
| Both | sex | 34 | 787 | 57.8813826 | 0.00650066 | 0.02600264 |
| Both | age at darting | 34 | 753 | 45.04296071 | 0.097551753 | 0.177366823 |
| Both | maternal rank | 34 | 603 | 37.68510893 | 0.3043452 | 0.358053177 |
| Both | group size at birth | 34 | 629 | 52.14243613 | 0.024070279 | 0.048140558 |
| Both | group size at darting | 34 | 722 | 55.21693698 | 0.012152455 | 0.034721299 |
| Both | maternal dominance rank | 34 | 569 | 44.43609912 | 0.108542801 | 0.180904668 |

[1]Each row of the table provides details for a model that tested the hypothesis that GEIs influenced gene expression for one or both sexes (specified in column 1), involving the environment specified in column 2.
[2]Model degrees of freedom
[3]Residual degrees of freedom
[4]Benjamini-Hochberg corrected p-values.

**Table S5. Evidence for GEIs involving tested interaction effects, for genes (n = 13) in which ASGE was best explained by a putative regulatory SNP in the resequencing data[1]**

| Interaction effect | p-value (females)[2] | p-value (males) | p-value (both sexes) |
|---|---|---|---|
| ***social environments*** | | | |
| dominance rank | | | |
| early (maternal) | 0.313 *(12)* | 0.225 *(5)* | 0.187 *(13)* |
| adult | 0.080 *(12)* | 0.047 *(9)* | NA[3] |
| social connectedness | | | |
| early (maternal) | 0.126 *(11)* | 0.643 *(5)* | 0.098*(13)* |
| adult | 0.030 *(12)* | $1.20 \times 10^{-3}$ *(7)* | NA |
| group size | | | |
| early | $3.63 \times 10^{-4}$ *(12)* | 0.377 *(5)* | 0.027 *(13)* |
| adult | 0.003 *(12)* | 0.074 *(9)* | 0.635 *(13)* |
| | | | |
| ***non-social effects*** | | | |
| age | 0.049 *(12)* | 0.331 *(9)* | 0.181 *(13)* |
| sex | NA | NA | 0.06 *(13)* |

[1]This table recapitulates Table 1 in the main manuscript, but focuses on the smaller subset of genes for which the SNP that best explained ASGE variation was a SNP in the resequenced putative regulatory region (i.e., not the assay SNP)

[2]P-values from likelihood ratio tests comparing a model with GEIs to a model without GEIs are provided in each cell. Values in parentheses provide the number of genes included in each test.

[3]Social connectedness and dominance ranks in adulthood could not be evaluated for both sexes combined, as SCI values were calculated separately for males and females.

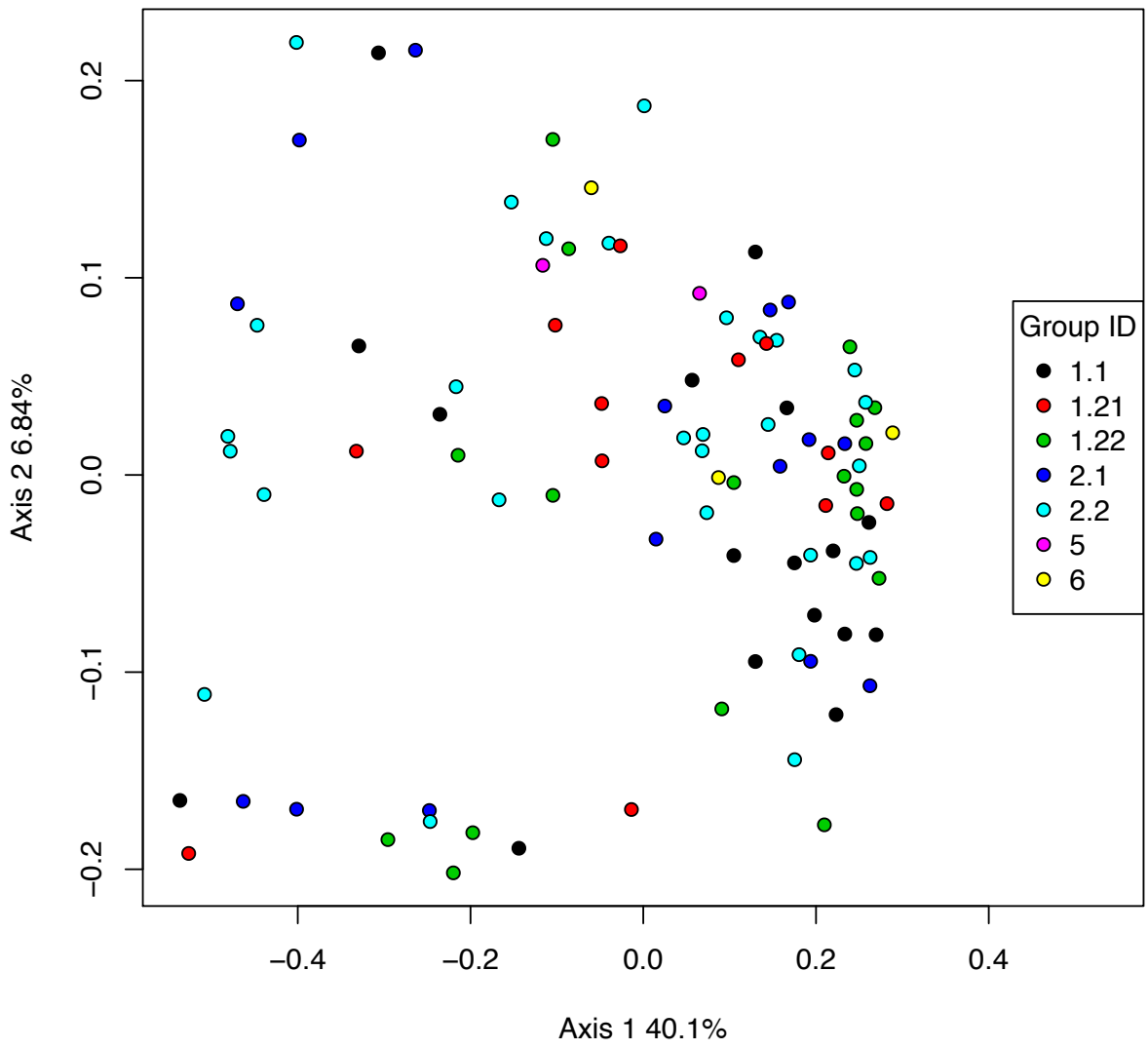


**Figure S1. Population structure in the Amboseli study sample largely does not reflect social group.** Classical (metric) multidimensional scaling was used to project a genetic distance matrix among the 96 individuals in the study sample onto two axes. The distance matrix was calculated as $d_{ij} = 1\text{-}R^2_{ij}$ where $R^2_{ij}$ is the squared correlation coefficient of genotypes (coded as 0,1,2) for individuals i and j. Only SNPs that passed all GATK quality filters were considered. The percent of variation in *d* (i.e., the percentage of overall genetic variation in the data set) accounted for by each axis is reported. Dots represent individual study subjects, and are coded according to the group that they belonged to at the time of darting. Note that subjects do not cluster by social group, reflecting the weak contribution of social group to overall population genetic structure in this population.

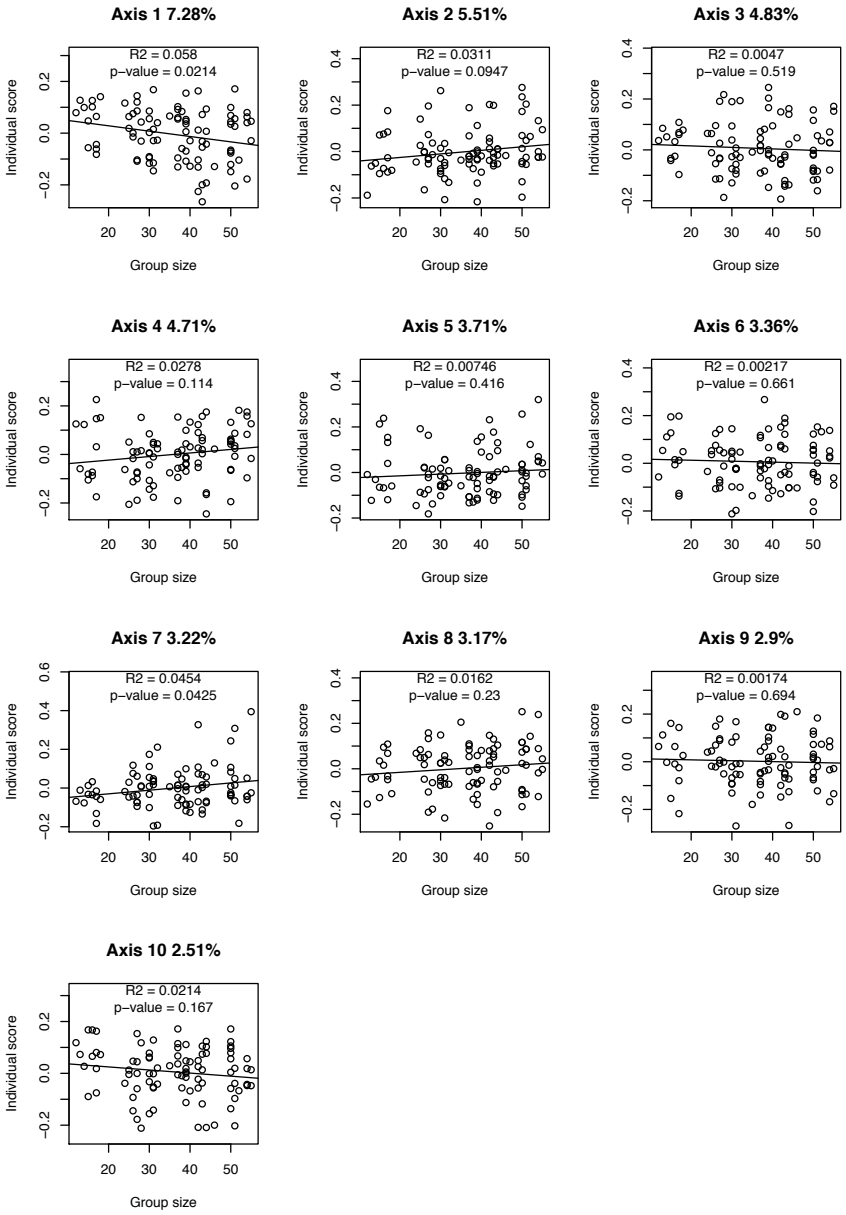


**Figure S2. Group size does not explain substantial variation in the major axes of population structure.** Principal components decomposition was performed on SNPs identified in the data set (using the R package *SNPRelate*) to identify the major axes of genetic structure in the sample. Each panel shows the correlation between one of these axes and group size at the time of darting of each individual. Panel titles provide the percentage of overall genetic variation accounted for by each axis, and inset text shows the $R^2$ and p-values for a linear regression of PCA axis score on group size. Note that even where these p-values are nominally significant, group size explains very little overall genetic structure in the Amboseli population.

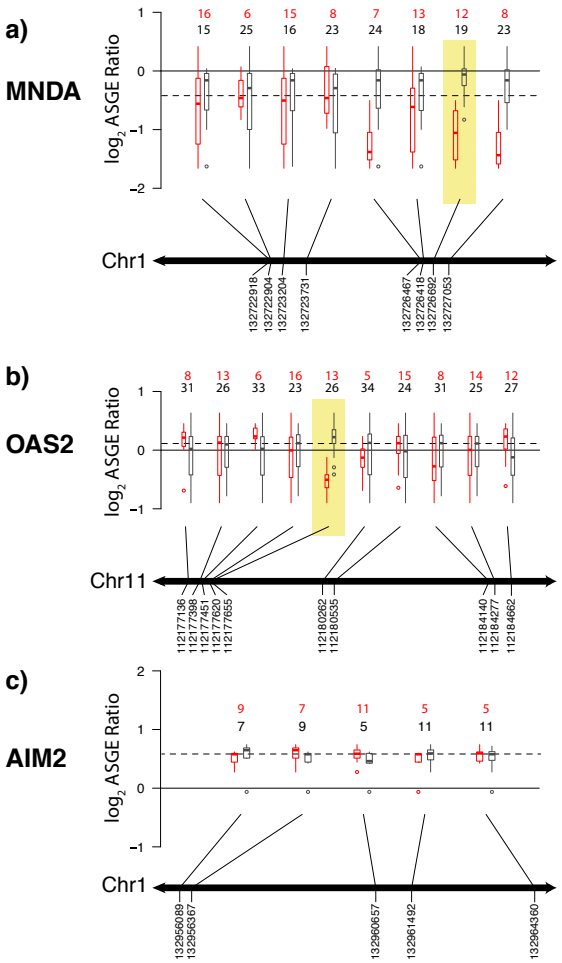


**Figure S3. Identification of SNPs for downstream GEI analysis.** Common ASGE implies the presence of functional *cis*-regulatory variation. Individuals who exhibit high (low) levels of ASGE are expected to be heterozygous (homozygous) for the functional *cis*-regulatory site or sites. Thus, genotype at the functional site, or sites in high linkage disequilibrium (LD) with the functional site, should explain substantial variance in ASGE levels. For each gene that exhibited common ASGE in the Amboseli samples, we identified the SNP for which genotype explained the most variance in ASGE levels, constraining homozygotes to have ASGE levels near zero, and heterozygotes to have non-zero ASGE levels. In each case, we considered SNPs in the putative upstream *cis*-regulatory region as well as the SNP used to assay ASGE itself. We calculated the variance in ASGE ($R^2$) explained by each SNP, and identified the model that explained the most variance in the ASGE data as focusing on the most suitable SNP for downstream GEI analysis. a) and b) show cases in which the best SNP was located in the putative *cis*-regulatory region resequenced for the two respective genes. c) shows a case in which all individuals in the sample exhibit consistent non-zero ASGE and variance in ASGE was best explained by the assay SNP (not shown); if the assay SNP is in high LD with a functional site, all individuals in the sample are expected to exhibit non-zero ASGE. In each panel, SNPs are shown in their relative positions on the chromosome (not to scale), with gene ASGE ratios partitioned by heterozygous (red) versus homozygous (black) genotype. Number of heterozygotes and homozygotes for each SNP are provided about the boxplots. Horizontal dashed lines in each plot show the median ASGE ratio of all samples.